# Information sharing, credit market competition and loan performance

Jaap Bos[*], Ralph De Haas[†] and Matteo Millone[‡]

Preliminary version: December 30, 2013. Please do not cite or redistribute. Comments welcome.

**Abstract**

We use detailed data on over 200,000 loans granted by a large microfinance institution in Bosnia and Herzegovina to assess the impact of the introduction of a credit registry in 2009 on loan quality. We find that the introduction of mandatory information sharing among lenders had a substantial positive impact on the quality of new loans. In line with theory, this impact was the highest in areas with intense credit market competition and for loans to first-time borrowers. While the introduction of the credit registry did not lead to shifts in the average borrower profile along observable characteristics, we find that loans became smaller, shorter, and more expensive. This suggests that at least part of the improvement in loan quality was due to more conservative lending practices at the intensive margin.

*Keywords:* Information sharing, credit market competition, hazard models

*JEL Codes:* C58, G18, G21

[*]Maastricht University (j.bos@maastrichtuniversity.nl).
[†]European Bank for Reconstruction and Development (dehaasr@ebrd.com).
[‡]Maastricht University (m.millone@maastrichtuniversity.nl).

# 1 Introduction

Agency problems in banking remain rife and this holds in particular for emerging markets, where information asymmetries tend to be high, screening and monitoring is costly, and legal institutions are weak. This makes it difficult for lenders to successfully apply mitigation mechanisms that are tried and tested in more benign lending environments: requiring borrowers to post collateral [Bester, 1987], acquiring private information about borrowers via relationship lending [Bhattacharya and Chiesa, 1995] and employing contingency contracts [Stiglitz and Weiss, 1983]. On the demand side, potential borrowers may want to gradually reveal their quality by building up a reputation with their bank [Diamond, 1989], but may simply find it too difficult to credibly signal initial quality in order to get a first 'entry' loan. Some of these issues become even more salient as the number of lenders increases.

Perhaps not surprisingly then, various countries have recently introduced credit reporting institutions, either private credit bureaus or public credit registries, in an attempt to increase transparency and reduce information asymmetries in credit markets and, ultimately, to improve access to finance among households and entrepreneurs. The empirical evidence on the impact of such formalized systems of information sharing remains limited and stems either from broad cross-country comparisons or stylized experiments in a laboratory setting.

This paper presents direct evidence of what happens when lenders are required to share borrower information in a competitive credit market. We carry out our analysis using a unique data set built around the loan portfolio of one of the largest providers of small business loans in Bosnia and Herzegovina (BiH). We match over 200,000 loans to both borrower and branch characteristics. We also measure the level of local credit market competition, either in the form of standard concentration measures or by culling information on loan officers' own perceptions of the intensity of local competition from a survey. Our sample period runs from 2002 to 2012, with the introduction of the credit bureau taking place in 2009. The result is a setting where we can exploit both time variation in information sharing and a cross-sectional variation in competition. This allows us to identify the impact of a mandatory credit register on loan quality. In addition, we analyze whether the improvement in information sharing has had an impact on loan conditions such as size, maturity and interest rates.

Our paper contributes to, and for the first time connects, two main strands of the literature. First, we add to the evidence on the effects of information sharing. On the theoretical side, Pagano and Jappelli [1993] model the conditions under which information sharing arises endogenously in credit markets. They find that size of the market and borrower mobility and heterogeneity all increase the incentives for lenders to share information. Most importantly, under severe adverse selection an information sharing system facilitates lending. Moral hazard is mitigated too as information sharing also serves as a disciplining device that increases borrower effort [Padilla and Pagano, 1997, 2000]. Karapetyan and Stacescu [2008] prove that improvements in the distribution of hard information may encourage lenders to invest more in soft information in order to gain a competitive advantage, and this will lead to even better lending decisions. These papers therefore tend to predict that (mandatory) information sharing will increases the volume of lending while also improving loan quality.

The available empirical evidence appears to broadly confirm that countries with a well-functioning information sharing system exhibit a higher level of bank lending to the private sector, with lower default rates [Jappelli and Pagano, 2002, Pagano and Jappelli, 1993] and lower interest rates. This beneficial impact is particularly strong in developing countries [Djankov et al., 2007] and for opaque firms [Brown et al., 2009]. Evidence also suggests that information sharing may reduce bank risk taking, increase bank profits and lower the risk of financial crises [Houston et al., 2010]. Using lab experiments, De Janvry et al. [2010] show that the introduction of a credit bureau increases lending efficiency and particularly so when borrowers are aware of its existence. Our contribution is to use loan-level evidence from before and after the introduction of a mandatory credit registry to assess at the microlevel, where we can adequately control for both borrower and loan characteristics, how information sharing impacts lending at the intensive margin.

Second, we contribute to the work on the relationship between bank competition and credit availability. This literature has long been characterized by two opposing views. On the one hand, there is theory - the market-efficiency view, cf. Pagano and Jappelli [1993] - as well as evidence to suggest that bank competition alleviates credit constraints as more loans become available at better terms. This can in turn positively influence local economic growth and entrepreneurship [Guiso et al., 2004].

Other contributions, however, suggest that less bank competition may benefit firms, especially more opaque ones, as market power allows banks to forge long-term lending relationships (Berger and Udell [1995], Ongena and Smith [2001], Petersen and Rajan [1994]). Petersen and Rajan [1995] show theoretically how in a concentrated banking market lenders subsidize early loans by extracting rents from later ones. Banks will only be willing to assist firms in the beginning of a relationship if these firms can commit not to leave the bank in the future. This will be impossible in highly competitive markets, thus ruling out the intertemporal smoothing of interest rates that is needed to give opaque borrowers a chance. Interestingly, a small branch of this literature suggests that even in a relationship-lending setting more competition may be beneficial to access to credit [Boot and Thakor, 2000]. If competition incentivizes banks to invest more in generating 'soft' information about borrowers, then competition may benefit small and opaque firms in particular [Dell'Ariccia and Marquez, 2004, Hauswald and Marquez, 2006].

Empirically, Degryse and Ongena [2007], Elsas [2005] find that Belgian and German firms, respectively, indeed enjoy stronger credit relationships in more competitive markets. Recently, attempts have been made to reconcile both opposing views. Bonaccorsi di Patti and Dell'Ariccia [2004] use Italian data to show that while bank market power boosts firm creation, in particular in opaque industries, after some optimum additional market power starts to have a negative impact on firm creation. Likewise, Cetorelli and Gambera [2001] - with a cross-country, industry-level dataset - show that bank concentration promotes the growth of sectors that depend on external finance. Yet, overall there is a negative association between banking-industry concentration and economic growth.

While our paper speaks to both of these streams of the literature, our main contribution lies at their intersection, an as yet unexplored territory in particular among empiricists. On the theoretical side, Hoff and Stiglitz [1998] posit that without information sharing between lenders, borrowers perceive a lower cost of default when competition increases as defaulting borrowers can now more easily get a loan from another lender. Impatient borrowers have an incentive to take multiple smaller loans from different lenders instead of applying for one larger (more expensive) loan from one lender. If they can hide their outstanding debt, these loans will be considered less risky and cheaper. This will increase average debt levels, eventually leading to worse repayment rates and higher interest rates. In this view, increased competition in the absence of information sharing may lead to overindebtedness and a gradually worsening of loan repayment. The introduction of an effective information sharing mechanism may put an end to this as borrowers can no longer hide multiple loans and will also realize that default will cut them off from future credit from *all* lenders.

McIntosh and Wydick [2005] come to a similar conclusion by taking the lenders' perspective. They demonstrate that under imperfect information sharing, competition between for-profit and non-profit lenders can leave poor borrowers worse off as it becomes more difficult for a non-profit lender to cross-subsidize among borrowers. At the same time, as in Hoff and Stiglitz [1998], impatient borrowers may obtain multiple loans and become overindebted. Finally, Hauswald and Marquez [2006] show that as competition increases, banks invest less in proprietary information. The result is a drop in interest rates but also a loss of efficiency in lending, resulting in higher default rates. If adverse selection contributes to overindebtedness, then the latter is expected to be particularly troublesome in markets with a high level of competition. Sharing information may then be an important way to overcome adverse selection problems and to prevent the resulting overindebtedness, in particular in highly competitive credit markets. Our contribution to this literature is to provide the first empirical evidence on the interaction between changes in information sharing and local credit market competition and to directly test some of the main conjectures that theory has put forward over the last couple of years.

Our identification relies on cross-sectional and time-series variation in competition, which we combine with the introduction of a mandatory national credit register. Further facilitating our analysis, we can distinguish between first-time customers, for whom information sharing can result in large decreases in asymmetric information and repeat customers, who have already built up a credit history with their bank (but whose behavior may still change in line with the dynamic incentive mechanisms outlined by Hoff and Stiglitz [1998]). To assess what happens when information sharing becomes mandatory, we fully exploit the richness of our data by modeling the repayment performance of borrowers using hazard function estimators or survival analysis. This class of estimators is particularly suited for our dataset as it corrects for right censoring, which would lead to biased estimates in static logistic models [Shumway, 2001]. Furthermore by taking the duration of the loan into account it will not only provide information on whether but also on when a loan is most likely to default.

To preview our results, we find that the introduction of mandatory information sharing leads to a significant decrease in defaults. In line with the theoretical work outlined above, the effect of the credit register introduction on loan performance is particularly striking in markets with high levels of competition and for first-time borrowers. Moreover, mandatory information sharing results in smaller, shorter and more expensive loans, indicating that information sharing led to more conservative lending behavior at the intensive margin. Our results are robust to a large set of specification choices. Moreover, a placebo analysis confirms that we correctly identify treatment effects from the introduction of the credit register.

The article is organized as follows. In Section 2 we give some more background on our empirical setting in Bosnia and Herzegovina, while Section 3 describes our data and identification strategy. In Section 4 we then present our empirical results, after which Section 5 discusses the implications of our analysis and concludes.

# 2 Information sharing and small business lending in Bosnia

Loan portfolio data have been obtained from EKI Mikrokreditina Fondcija (henceforth EKI), a local Bosnian micro lender. EKI was founded in 1996 in Zenica as a part of World Vision International. It is now the fourth largest provider of individual liability microloans in Bosnia, it operates via 14 branches and 52 regional offices. In May 2012 it had an outstanding loan portfolio of BAM 93 million (USD 65.5 million) and around 34,000 active borrowers. According to EKI's website its mission is to "provide financial services and technical support to those who have no access to bank services or to businesses that are creating and sustaining jobs in the whole of BiH, wherever there is a need and opportunity."

EKI went through a number of regulatory changes following the evolution of the Microfinance sector in the country. It changed its status first in 2001 from NGO to Microfinance organization (MCO) and again in 2008 when it became a Microfinance Foundation (MCF). EKI now falls under the supervision of the Bosnian Central Bank, it is allowed to give loans but cannot accept deposits. It funds its operations mainly through loans from Microfinance Investment Vehicles (MIVs) or Development Financial Institutions (DFIs).

## 2.1 Microcredit in Bosnia and Herzegovina

Microfinance plays a central role in the provision of small business loans in Bosnia and Herzegovina (BiH). The Bosnian microfinance market is the largest in South East Europe and has reached penetration rates second only to Bangladesh. From their initial status as independent NGOs, Bosnian MFIs have evolved into financially sustainable and sometimes highly profitable micro-lenders. These institutions have grown in number and outreach and in certain cases compete directly with national and international commercial banks in the supply of loans to micro entrepreneurs Woodworth [2006].[1]Nevertheless a repayment crisis in 2008 re-dimensioned the growth prospects of the sector.

In 1995, when BiH emerged devastated by the Yugoslavian civil war, microfinance was one of the main tools adopted by the international community to aid the economic recovery and development of the country. The industry grew quickly, but did so with somewhat different characteristics from the microfinance model pioneered by Mohammed Yunus in Bangladesh. Unlike in South East Asia and Africa, MFIs in BiH were allowed to lend to a highly educated entrepreneurial middle class who had lost its livelihood due to the civil war [Hartarska and Nadolnyak, 2008, Demirgüç-Kunt et al., 2007]. As a result the microfinance business model in BiH quickly moved away from joint liability lending and focused on providing individual loan products tailored to the needs of able and experienced micro-entrepreneurs.

In the year 2006, the sector underwent a large regulatory overhaul, when limits to interest rates and loan size were lifted [Welle-Strand et al., 2010]. The new regulatory framework in combination with good financial performance generated an increase in investment by development financial institutions and microfinance investment vehicles [Chen et al., 2010]. Abundant capital allowed MFIs to significantly increase lending and led to growth rates of 38% p.a. up to 2009. In a small country of 3.8 million people, penetration rates quickly doubled from 10% to 20% prior to the crisis [Lützenkirchen et al., 2012].

Intensive growth combined with high competition levels and decreasing lending standards led to a sharp increase in over-indebtedness, as 40% of borrowers had more than one microfinance loan at the same time. The signs of market saturation where already clear in early 2008 when at least 16% of borrowers were late on repaying their loans. The already precarious situation collapsed at the onset of the financial crisis when PAR30 went from 1% to 11% in one year and return on asset turned negative [Lützenkirchen et al., 2012]. MFIs reacted with an immediate reduction in lending and by aggressively writing off non-performing loans. Despite high losses and significant contractions in portfolio size, the industry did not collapse and since 2010 lending has started to grow again [Glisovic et al., 2012].

## 2.2 Information sharing system in Bosnia and Herzegovina

One of the causes of the spike in non-performing loans in 2008 was the fact that lenders were unable to check whether borrowers already had a loan outstanding

[1]International banks like Intesa San Paolo and Raiffeisen Bank are heavily involved in the provision of loans to micro entrepreneurs.The difference in loan size between MFIs and commercial banks has been converging since the early 2000s [Berryman and Pytkowska, 2003].

with another institution. Multiple lending was in fact quite possible despite the existence of a privately owned and operated credit bureau (the LRC) since the year 2000. The problem with the LRC was that participation was voluntary and expensive, making it incomplete and therefore ineffective. Experimental evidence by Brown and Zehnder [2010] shows that in highly competitive markets lenders do not have incentives to share information via a private credit bureau.

The Bosnian Central Bank introduced a public credit registry (Centralni Registrar Kredita, CRK) in 2006, but it was not until July 2009 that the coverage was extended and participation was made mandatory to all formal lenders, including MFIs. The most important feature of the CRK is that the submission of credit history and current balance is mandatory, granting complete coverage of the entire lending market [Lyman et al., 2011]. Interviews with local stakeholders confirm that the July 2009 introduction of the mandatory and full CRK led to a sudden change in the amount of information about prospective borrowers that both banks and MFIs had access to. According to one manager at a large Bosnian MFI: "Before that date, we were basically blind". Importantly, no other regulations in the area of banking or financial regulation more broadly were undertaken in the Summer or Autumn of 2009 that may have confounded the effect of the CRK introduction in July 2009.

The strict requirements imposed on lenders by the introduction of the CRK resulted in a significant improvement in the degree of information sharing. In particular, the Bosnian Central Bank requires every lender to submit a report to the CRK every time a loan is disbursed, repaid in full, late or written off. A credit score is given to each borrower, based on the information on current debt and past performance. The CRK also includes information on whether the borrower has a guarantor or whether she is a guarantor himself. Lenders are required to introduce a condition in each loan contract to be signed by the borrower, in which the latter agrees on a credit check via the CRK and on the submission of his or her credit history. Borrowers are therefore completely aware of the fact that their repayment performance will be recorded and shared with other lenders. Submitting updates is obligatory, but checking the data is voluntary and subject to a small fee (BAM 0.15).

# 3 Data and Methodology

## 3.1 Data

We obtain data from a number of different sources. Our main dataset consists of individual loan information from a large Bosnian microlender. This institution has been operating since 1996 and is active in 15 geographically distinct branches. This data contains information on the characteristics and repayment schedule for more than 200,000 microloans disbursed from June 2002 until December 2012. This information is particularity useful as we do not only observe loan characteristics such as size, maturity, interest rate, collateral and loan purpose, but we also have precise information on whether and when there was a late repayment, whether the loan was written off and how much of the principal and interest was recovered. Furthermore, for each loan we can match characteristics of close to 130,000 clients, including income, education, gender, employment status and family size.

Using this information we construct our measure of loan quality which is a dummy equal to 1 if repayment on a loan is at least 30 days late and the loan is subsequently written off. Around 97% of loans that are late once end up being written-off. The time between disbursement and late repayment is our main variable of interest for the estimation of the hazard rate. We do not use the write-off date as our variable of interest because its timing would be more dependent on the lender rather than on the borrower.

Our main independent variable is the local level of competition. We construct a branch-based HHI index by collecting information on the amount of branches in each area of operation. Given the incompleteness of most available datasets on microlenders we need to collect this data from a number of different sources. We conduct a survey where we ask two loan officers for each branch a number of questions relating to their perception of the competitive environment. One of these questions requires loan officers to list the main competitors operating in the area. We then cross-check this list with the list of microlenders operating in Bosnia as reported by the MIX dataset. Once we have a complete picture of the universe of Bosnian microlenders, we proceed to compile a panel dataset of the geographical distribution of branches. In order to do so we ask each MFI for a list of their branches and year of operation. We have a response rate of around 50%. Therefore we complete the dataset manually by extracting lists of branches from annual reports of each MFI. We are then able to produce a year-location branch-based measure of competition which is equal to $1 - HHI_{by}$ where $b$ indicates branch and $y$ the year.

We also construct a perception-based measure of competition based on the loan officer survey. We ask the loan officers how much they agree, on a scale from one to seven, with the following statement: "In the last ten years I think that there has been an increase in competitive pressure in my area of operation". This measure of competition, or rather increase in competition, is time invariant and available only for 14 branches as one branch had closed down by the time the questionnaire was sent out.

For both measures of competition we construct a dummy equal to one for values above median. We do so for two reasons. First, Stepanova and Thomas [2002]show that grouping observations into cohorts improves the performance and the interpretation of hazard models. Second, we are aware of the limitations of using concentration as a measure of competition , and do not want to attach too much weight to interpreting a marginal decrease in concentration, but we are more interested in a measure that allows us to rank branches.

In order to make sure that we are estimating the effect of competition and information sharing on repayment behavior we need to make sure that we control for differences in economic activity both geographically and over time. Unfortunately, it is impossible to find conventional measures of local economic activity. At the country level the highest frequency data on GDP is yearly. We control for local economic activity by using local yearly night light data from 2003 to 2010 as proposed by Henderson et al. [2011]. We use the Bosnian Investment Fund Index published by the Sarajevo stock exchange as a high frequency proxy for national economic activity.

## 3.2 Identification strategy

For our identification strategy, we fully exploit the rich data that are at our disposal to identify the effects of mandatory information sharing on the loan conditions of opaque borrowers.

In the context in which we study it, mandatory information sharing requires the co-occurrence of two things. The first is a means for sharing information, i.e., a well-functioning credit register. The second is for that credit register to result in mandatory information sharing: after all, in markets with a low degree of competition, there is less of a need to share information and less of a possibility to share information, since there are fewer banks. Put simply, in a market with only one bank, there is still the introduction of the credit register, but there is no information sharing as a result of that introduction. However, in a market with a high degree of competition the introduction of a credit register truly introduces mandatory information sharing in that market.

Now that we have started to describe, at least informally, the treatment effect we aim to measure, it is time to take a closer look at the treated. Most importantly, mandatory information sharing can have an effect, if there is an information advantage from doing so. If a bank has a long credit history with a borrower, it is impossible to argue that this borrower is, to this bank, opaque. In that case, mandatory information sharing is a treatment for the healthy: we cannot expect any serious effects. However, for first-time borrowers who do not have a credit history with their bank, mandatory information sharing can be the treatment that cures the negative effects the large degree of asymmetric information may otherwise have on the lending process.

More formally, we can best describe our identification process as follows:

$$\begin{aligned} E[L_{it}|MIS &= 1] - E[L_{it}|MIS = 0] = \\ &E[L_{1it}|MIS = 1] - E[L_{1it}|MIS = 0] + \\ &E[L_{0it}|MIS = 1] - E[L_{0it}|MIS = 0], \end{aligned} \tag{1}$$

where $L_{it}$ is a loan characteristic for a borrower $i$ at time $t$, for example loan quality or loan amount, and if $MIS = 1$, there is mandatory information sharing. Assuming for now that we can perfectly measure $MIS$ and identify the treated, $E[L_{1it}|MIS = 1] - E[L_{1it}|MIS = 0]$ measures the average treatment effect, since for each treated borrower $E[L_{1it}|MIS = 1]$ is the expected (or average) value for the loan characteristic after treatment, and $E[L_{1it}|MIS = 0]$ is the expected value of the same loan characteristic before treatment. The remaining selection bias is captured by $E[L_{0it}|MIS = 1] - E[L_{0it}|MIS = 0]$, since $E[L_{0it}|MIS = 1]$ is the expected value of the loan characteristic after the treatment, for those who did not receive the treatment, and $E[L_{0it}|MIS = 0]$ is the same value before treatment, for those who did not receive the treatment. Of course, a proper identification strategy minimizes the selection bias, so as to more accurately measure the average treatment effect.

In practice, of course, both measuring the existence of mandatory information sharing and those that are 'treated' by it (rather than just undergoing it) is not necessarily straightforward. In Table 1 we explain how we can still identify the impact of mandatory information sharing on opaque borrowers.

Table 1: Identification

| | | *Asymmetric information* | |
|---|---|---|---|
| | | High | Low |
| *Competition* | High | $A$. Treated, $MIS = 1$ | $B$. Control for asymmetric information, $MIS = 0$ |
| | Low | $C$. Control for competition, $MIS = 0$ | $D$. Control, $MIS = 0$ |

First, we observe that the borrowers we are interested are those that operate in markets with a high degree of asymmetric information and a high level of competition. In Table 1, these borrowers are located in $A$. For these borrowers, the introduction of the credit register results in the introduction of mandatory information sharing. Borrowers in $B$ also exist in markets with a high degree of competition, but because they have already reduced the degree of asymmetric information with a lender, the introduction of the credit register does not 'treat' them quite in the same way. To separate these borrowers from those in $A$, we have to carefully control for the degree of asymmetric information. Likewise, borrowers located in $C$ are 'eligible' for treatment, as they too face a high degree of asymmetric information (i.e., they are first-time borrowers), but since they operate in a market with a low level of competition, there is little to share once the credit register is installed. To separate these borrowers from those in $A$, we have to carefully control for the degree of competition. Finally, borrowers in $D$ constitute the best control group: they differ the most from the borrowers in $A$, since they operate in markets with low competition and have a low degree of asymmetric information.

For our interpretation of Table 1, three remaining matters are important. First, we require that first-time borrowers do not *select* into treatment. Put differently, we will verify that - especially in high-competition markets - first-time borrowers' characteristics are not significantly different before and after the introduction of the credit register. Second, in practice mandatory information sharing of course also occurs for returning borrowers. And whereas we can control for the level of competition rather well, controlling for the exact amount of asymmetric information is more difficult. So, in interpreting the impact of the mandatory information sharing on first-time borrowers, we will use the impact on returning borrowers as a benchmark, so as to not overestimate the former effect. Third, we wish to rule out that strategic timing of the introduction of the credit register drives our results. Therefore, in the second part of our analysis we run placebo analyzes where we either postpone or bring forward the introduction of the credit register.

Our objective, then, is to exploit both the cross-sectional and time-series variation in local competition and asymmetric information to separate treatment effects from selection bias to the best our abilities. In the next subsections, we explain how we do this, both for effects of mandatory information sharing on loan defaults and for its impact on loan characteristics.

### 3.3 Modelling loan defaults

In the first part of our analysis, we wish to identify the impact of mandatory information sharing on defaults. In line with the strategy outlined above, we especially want to investigate the impact of the introduction of the credit register in high-competition markets, for first-time borrowers. As also explained above, by also identifying the other cases (low competition markets, returning borrowers), we can more precisely extract this impact.

We do so using a hazard model, where the hazard rate is defined as the probability of a borrower being late on its repayment in time $t$ conditional on the fact that she has repaid regularly up to that point. Hazard functions allow us to model not only whether a loan is going to default but how the probability of default changes over time. This is particularly important for studying loan repayments as the reasons to default might change during the life of the loan (i.e. strategic default).

Aside from their economically intuitive interpretation, the main advantage of hazard models is their ability to deal with censoring. Censoring occurs when the loan is repaid or when the life of the loan extends beyond the sample period (right censoring). Given that most loans are repaid successfully, the effects of censoring in estimating the probability of default of a loan will be particularly severe. In fact not correcting for it will yield static biased and inconsistent estimates in static probability models Ongena and Smith [2001]. The semi-parametric model developed by Cox and Lewis [1966], Cox [1972] is able to deal with right censoring as the log-likelihood function takes into account the ratio of completed vs. non-completed loans.[2]

With the hazard model, we can compare the hazard rates for each of the four cases outlined in Table 1, before and after the introduction of the credit register. If mandatory information sharing indeed results in a better allocation of credit,

[2]Left censoring can cause biased estimates as well, but it is not an issue in our case as we only observe new loans.

then we expect a large drop in the hazard rate for all borrowers with loans in $A$. For borrowers in $B$, we may also observe a drop in the hazard rate, but it should be considerably smaller, as there is less need to share information. Likewise, for borrowers in C, although there is a need to share information, there is less information to be shared, as there are fewer competitors, so we expect a smaller drop in the hazard rate. Finally, borrowers in $D$ constitute our baseline: even if mandatory information sharing lowers strategic defaults, bad luck ensures that there is still a proportion of borrowers that are not able to repay their loans.

To estimate the baseline hazard, as well the different reference cases, let $T$ measure the amount of time before the first late repayment of the loan. The hazard function can be used to describe the distribution of $T$ and it is defined as:

$$h(t) = \lim_{\Delta t \to 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\}. \tag{2}$$

The hazard function $h(t)$is the probability of repayment on a loan being late in time $t$ conditional on regular repayments until then. Alternatively we can model the distribution of time until first late repayment as its survivor function:

$$\begin{gathered} S(t) = P(T \\ \geq t). \end{gathered} \tag{3}$$

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

$$h(t) = \frac{-dlogS(t)}{dt}. \tag{4}$$

Using the non-parametric [Kaplan and Meier, 1958] estimator we can plot the survival function for different groups of loans. This estimator is easily adjustable for right censoring.

Following Cox [1972] we can estimate the effect of a set of covariates $X_t$ and the distributions of time to late repayment with the proportional hazard model:

$$\begin{aligned} h(t) &= \lim_{\Delta t \to 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t, X(t), \beta)}{\Delta t} \right\} \\ &= h_0(t) \exp(\beta' X_t), \end{aligned} \tag{5}$$

where $h_0$ represents the baseline hazard when $X = 0$. Therefore the hazard rate for each individual with characteristics $X_t$ is proportional to $h_0$. The marginal effect $X_t$ on the log of the covariates hazard rate is represented by the estimated $\beta$ coefficients. 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 $\beta$ parameters via maximum likelihood.

The Cox proportional hazard model relies on two assumptions. First, it assumes continuous time as the presence of tied events in discrete time would make ranking impossible. Practically late repayments are only observed at intervals so we deal with tied events with the approximation by Breslow [1974]. Second, it assumes proportionality, which implies time fixed $\beta$ coefficients. We relax this assumption by estimating a model where the effect of covariates $X_t$ can change over the life of the loan.

We check the robustness of our results to the functional form of the hazard rate by estimating two common parametric specifications: the exponential and the Weibull distributions. The exponential distribution is widely used, it is easy to interpret and is characterized by a constant hazard rate. It is therefore considered memory less as the probability of late repayment is constant over time [Kiefer, 1988]. The exponential distribution is a special case of the Weibull distribution when $\alpha$ is equal to 1. The Weibull distribution is expressed as:

$$h(t) = \gamma \alpha t^{\alpha - 1} \tag{6}$$

The coefficient $\alpha$ is particularly interesting as it measures duration dependence. If $\alpha>1$ the hazard rate increases with time and vice versa, giving us an indication of the shape of the baseline hazard which is unobserved in the Cox specification.

## 3.4 Loan characteristics

In the second part of our analysis, we study the impact of mandatory information sharing on loan and client characteristics using both univariate analysis and multivariate regressions. To assess the consequences for credit supply, we study the impact on the loan amount, maturity and interest rate.

Here, our identification strategy translates into two elements. The first is the distinction between first-time and returning borrowers. In all analyses in this part, we focus on first-time borrowers, for whom mandatory information sharing is the most important. We start by testing whether these borrowers themselves are significantly different after the introduction of the credit register. In addition, in our multivariate analyses, we control for as many borrower characteristics as possible.

The second element is testing the *joint* effect of the introduction of the credit register and the fact that a borrower is in a highly-competitive market. After all, according to our earlier description, it is this joint effect that results in the mandatory information sharing we aim to measure. In our univariate analyses, we therefore split our data along both the competition dimension (high and low) and based on whether a loan is granted before or after the introduction of the credit register. In our multivariate analyses, we introduce both dimensions individually, but test the importance of their joint effect, as measured by their interaction.

Throughout, in our multivariate analyses, we of course control not only for other loan and client characteristics, but also for economics conditions. We now turn to our measurement of each of these.

# 4 Results

## 4.1 Information sharing, lender competition and loan quality

### 4.1.1 Non-parametric results

Figure 1 provides a first non-parametric look at our data in the form of a Kaplan-Meier survival analysis over the sample period June 2002 to December 2012. The graphs show how the probability that a borrower has not (yet) defaulted on her loan changes over time (horizontal axis, in quarters). At time of disbursement (t=0) the probability of survival is by definition 1 but then gradually erodes over time.[3] Panel (a) compares, for the whole sample period, the survival probability of borrowers in the branches that face below median competition versus those that are confronted with an above median level of competition. Competition is here measured as 1 minus the local Herfindahl-Hirschman Index (HHI), where local market shares are expressed in number of branches operated by a bank or MFI in the locality where the EKI branch is based.[4]

The key point to take away from Panel (a) is that there are is a limited difference in the survival behavior among borrowers in high versus low competition areas. After a year, the cumulative survival probability in the high competition areas is slightly *higher*. The statistical significance of this difference between both curves is confirmed by a logrank test (p-value=0.00). In economic terms, the difference is minimal however, an amounts to only 0.4 percentage points after a year. We come back to this result when we discuss Panel (c) below.

In Panel (b), we start to compare for the first time the survival behavior of loans provided before and after the introduction of the CRK. In this context right censoring will affect disproportionately the more recent group of 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].[5] The results show a substantial difference in repayment behavior, with loans after the CRK displaying a significantly higher survival probability compared to loans provided before mandatory information sharing was introduced. This is the first piece of evidence we bring to bear that points to a positive impact of the CRK introduction on loan quality. A striking aspect of Panel (b) is that the difference between both loan types seems to be mainly concentrated in the first couple of quarters after loan disbursement. Indeed, the probability of a loan not being late in the first six months after disbursement was increased from 94.6 percentage points before the CRK to 98.6 percentage points after the CRK introduction. Over time this difference is gradually reduced and becomes insignificant. This suggests that one of the mechanisms through which the CRK has impacted default is by affecting borrowers' behavior. Default in the first

[3]In effect, the graph thus shows the inverse of the cumulative default probability.

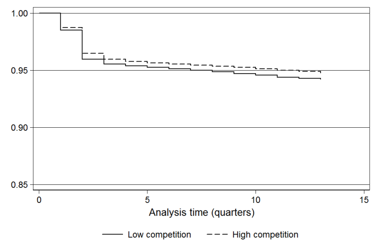
[4]Appendix Table A.1 (A.2) contains variable definitions and sources (summary statistics).

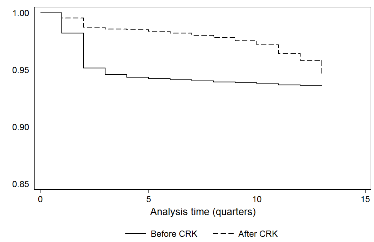
[5]Without correcting for right censoring, the hazard rate would be calculated as the ratio of all loans dropping from the dataset over remaining loans.

Figure 1: Information sharing, credit-market competition and loan quality: *Kaplan-Meier survival analysis*

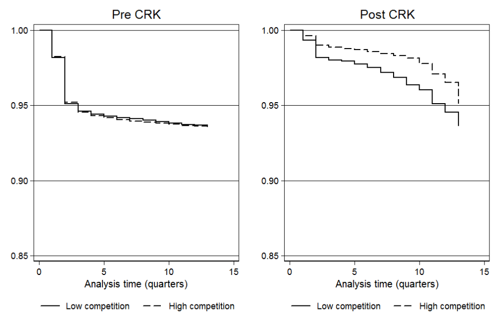
These three graphs show Kaplan-Meier survival estimates over the sample period Jun 2002-Dec 2012. Logrank test statistics for differences between the curves: Panel 1: $\chi2(1){=}27.35$ (p-value=0.00). Panel 2: $\chi2(1){=}1667.5$ (p-value=0.00). Panel 3 (left): $\chi2(1){=}0.53$ (p-value=0.47); Panel 3 (right): $\chi2(1){=}105.57$ (p-value=0.00).

(a) High versus low competition branches (b) Before versus after CRK introduction




(c) Interaction of credit-market competition and CRK introduction



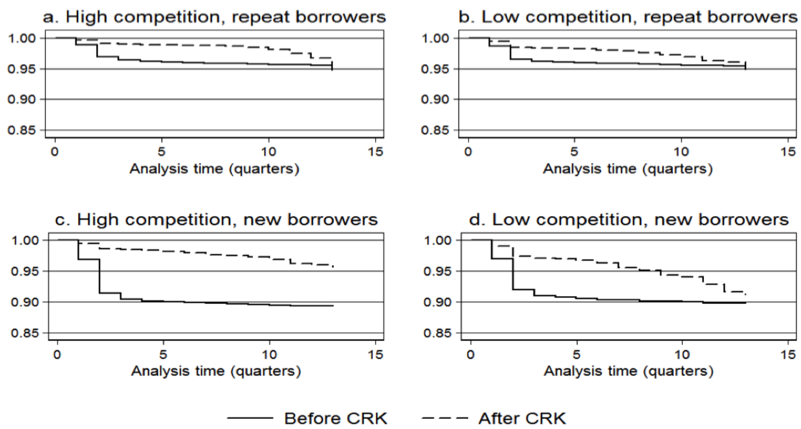

three months after disbursement is unlikely to stem from 'real' repayment problems and may to a large extent be due to strategic defaults. In line with some of the theories set out in the introductory section, one way in which a credit registry may improve repayment behavior is by disciplining borrowers through tougher dynamic incentives.

Finally, in Panel (c) we look for the first time at the interaction of the CRK introduction and the level of local competition. The graph on the left shows that before the CRK introduction there was no difference in survival probability between low and high competition areas (p-value of a logrank test is 0.47). However, after the CRK introduction, a significant gap opens up between repayment behavior in high versus low competition areas (and this is what drove the more limited overall difference over the whole sample period in Panel (a)).

With the CRK in place, repayment performance is now significantly higher in the high-competition areas (p-value =0.00). This difference amounts to 1 percentage points after 12 months, 2 percentage points after 24 months and remains significant throughout. Therefore, the CRK seems to have had most impact in high-competition areas where repayment rates before information sharing was in place were significantly below the level attainable in the presence of information exchange.

Figure 2: Information sharing and credit-market competition: First-time vs. repeat borrowers, *Kaplan-Meier survival analysis*

These four graphs show Kaplan-Meier survival estimates over the sample period Jun 2002-Dec 2012. Logrank test statistics for differences between the curves: Panel a: $\chi 2(1)$=523.90 (p-value=0.00). Panel b: $\chi 2(1)$=151.47 (p-value=0.00). Panel c: $\chi 2(1)$=797.98 (p-value=0.00); Panel d: $\chi 2(1)$=202.04 (p-value=0.00).




In Figure 2 we take this analysis one step further and now split our sample between first-time borrowers (that is, borrowers that were completely new clients to EKI) and repeat borrowers. One the one hand, we expect the impact of the CRK to be concentrated among first-time borrowers as the information asymmetry between lender and prospective borrower is largest. On the other hand, however, to the extent that the CRK introduction (also) had an impact on borrower behavior, we would also expect to observe an improvement in repayment behavior among repeat borrowers as these now realize that a default will 'cost' them more in terms of foregone future borrowing opportunities. As before, we also slice our data by competition level, leading to four separate quadrants in Figure 2. In Panels (a) and (b) we first focus on repeat borrowers. Independent of the level of competition, we see that the CRK introduction was accompanied by an upward shift of the survival function: at each point in time, repeat borrowers were less likely to have defaulted, suggesting that increased borrower discipline was indeed an mechanism through which the CRK had an impact. However, while in both graphs the differences between the 'before' and 'after' lines are statistically significant (p-value is 0.00 in both cases), this difference is limited and, moreover, fizzles out over time. This result links back to our earlier observation that part of the impact on the borrower side was in the form a reduction in ex ante moral hazard and strategic default.

Next, in Panel (c) and (d) we present a similar comparison, but now on the sample of first-time borrowers only. Compared to the two upper panels discussed before, there is now a striking difference. The impact of the CRK introduction appears to be much larger for new borrowers, suggesting that the CRK mainly 'worked' through the lender side. Comparing the low-competition areas (right) to the high-competition areas (left), we now also see clearly that the difference between both survival functions is widest and most persistent in the high-competition areas, exactly as theory would suggest. It is in these highly-competitive areas, where adverse selection problems are rife, that the CRK introduction had the most bite and EKI loan officers were able to put the borrower information that was hitherto unavailable to them to the best use. As a result, in these areas the survival probability for first time borrowers after 12 months went up from 94.3 to 98.7 percentage points.

#### 4.1.2 (Semi-)parametric results

In Table 2, we follow up on Figures 1 and 2 by providing semi-parametric and parametric evidence. As discussed in Section 3.2, an important advantage of using hazard models - where the hazard rate is defined as the probability of a borrower being late on its repayment in time $t$ conditional on having repaid regularly up to that point - is that it allows to deal properly with right censoring. A second advantage is that the specifications in Table 2 also allow us to control for a large battery of standard borrower and loan level covariates. Throughout the table we cluster the standard errors conservatively at the individual loan officer level. Results remain quantitatively and qualitatively unchanged when we do not cluster or cluster by EKI branch.

In columns 1-5 we present the results of our semi-parametric Cox proportional hazard model, while in columns 6 and 7 we show equivalent specifications using a parametric exponential and a Weibull model, respectively. In the first three columns we limit our sample to loans to first-time borrowers only, whereas in the following columns we use all loans while including a dummy *First loan* for first-time borrowers. We also interact this dummy with the CRK time dummy to test whether the impact of mandatory information sharing was larger for first-time than for repeat borrowers (as Figure 2 would suggest).

The results in the first three rows show, robustly across all specifications, that the CRK introduction is associated with a statistically significant reduction in the hazard rate and that this impact is almost twice as large in the high-competition areas. This confirms the evidence of Figure 1 and is line with the literature that we discussed before. The third line shows that the level of lender competition as such does not have an impact on the hazard rate. This confirms the visual evidence of Figure 1, Panel (c).

In the lower part of the table we show the estimated coefficients for a broad set of control variables. All of these have the expected sign and have in most cases a statistically significant relationship with the hazard rate. We find that more productive loan officers, as measured by *Loans/officer*, are characterized by better performing loan portfolios. On the borrower side, older and more educated clients pose less risk. Urban borrowers tend to be more risky, a result likely reflecting less tight social networks in urban as compared to rural areas.

As regards loan characteristics, we find that longer and larger loans tend to have higher repayment risks whereas - in line with various earlier contributions Rice and Strahan [2010], Roszbach [2004] - the presence of various types of collateral is an indicator of higher borrower risk.

We find a positive correlation between compliance with our MFI's collateral requirement and late payment though not with actual default. This is an interesting indication of adverse selection: to be a marginal client despite having collateral reveals other strong negative characteristics relating to repayment capacity.

As expected, columns 4-7 show that the interaction term between *First loan* and the CRK dummy is significantly negative, indicating that the CRK introduction reduced default risk in particular for these relatively opaque borrowers. The coefficient for *First loan* itself is positive, underlining the significantly higher risk associated with lending to borrowers previously unknown to the lender.

In column 5, we relax the proportionality assumption of the Cox model and allow the effect of the covariates to change over the life of the loan. This is accomplished by estimating another set of coefficients that change linearly over time since disbursement (not reported). We see that even without proportionality assumption the model still yields practically identical estimates.

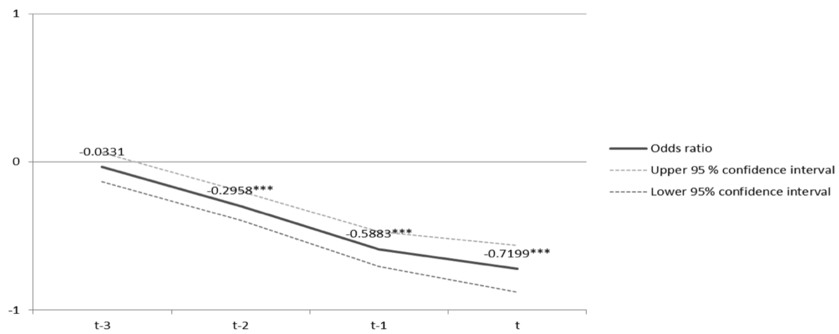
Finally, the Weibull model in column 7 produces a positive *alpha* of 0.525 (note that the exponential model shown in column 6 is a special case of the Weibull distribution where $\alpha$ is equal to 1). The $\alpha$ coefficient is particularly interesting as it measures duration dependence. The fact that alpha is smaller than one means that the hazard rate decreases with time, suggesting that a substantial part of the borrower risk if 'front loaded' and may reflect to a certain extent strategic default considerations.

In Appendix Table A.3, we provide further evidence on the robustness of these findings by estimating similar models while now allowing covariates to change over the life of the loan. In order to include time-varying covariates we need to modify the structure of our dataset. The previous models are estimated using a cross section of loans where every loan has one observation or line of data. In order to include time varying coefficients we expand the dataset so that we have a loan-period dataset. In the loan-period dataset each loan has multiple observations equal to the number of periods between disbursement and either repayment or default Singer and Willett [1993]. In this way the hazard rate does not depend only on the characteristics of the loan and the client at the time of disbursement, but also on a set of other variables that change during its life.

The results in Table A.3 are in line with those in Table 2: there is less default risk after the CRK introduction, in particular in more competitive areas and in particular for first-time borrowers. An interesting difference is that *Local competition* now enters negative and is estimated precisely. This reflects that in this more flexible specification we can actually exploit the time variation in our measure of local lender competition.

Figure 3: Cox proportional hazard model: Placebo test

This graph shows the odds ratio estimates (and a 95% confidence interval) for the variable Post CRK as used in column 3 of Table 1. The value at t shows the coefficient when using the actual timing of the CRK introduction. The values at t-1, t-2, etc. show the coefficient estimates when introducing the CRK 1 quarter, 2 quarters, etc. earlier than the real introduction date.




Finally, in Figure 3 we undertake a 'placebo' analysis in order to make sure that we the CRK variable we use really picks up the sudden introduction of the new credit registry and not some ongoing secular trend. The graph shows the odds ratio estimates (and a 95% confidence interval) for the variable *Post CRK* as used in column 3 of Table 2. The value at time *t* shows the coefficient when using the actual timing of the CRK introduction. The values at *t-1*, *t-2*, etc. show the coefficient estimates when introducing the CRK 1 quarter, 2 quarters, etc. earlier than the real introduction date. The results show that when we artificially bring the CRK introduction day forward, the placebo impact is quickly reduced in size and at three quarters before the actual introduction data becomes essentially zero. This gives us additional confidence that the significant and substantial effect that we detect exactly at the time the CRK was introduced is truly the impact of this shift in information sharing regime and not an ongoing longer-term trend.

Table 2: Information sharing and loan quality: Hazard analysis

| **Functional form** | | | **Cox** | | | **Exponential** | **Weibull** |
|---|---|---|---|---|---|---|---|
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] |
| Post CRK*local competition | -0.735*** | -0.550*** | -0.513*** | -0.343** | -0.341** | -0.375*** | -0.360*** |
| | (3.83) | (2.88) | (2.78) | (2.57) | (2.55) | (2.65) | (2.61) |
| Post CRK | -0.886*** | -0.624*** | -0.720*** | -0.502*** | -0.788*** | -0.372*** | -0.434*** |
| | (5.93) | (4.18) | (4.84) | (4.58) | (6.57) | (3.23) | (3.87) |
| Local competition | 0.0529 | -0.0518 | -0.0617 | -0.0920 | -0.0916 | -0.0674 | -0.0780 |
| | (0.40) | (0.43) | (0.55) | (0.76) | (0.76) | (0.50) | (0.61) |
| Local GDP | | -0.00671 | -0.00816 | -0.00925 | -0.00925 | -0.0116 | -0.0105 |
| | | (1.00) | (1.32) | (1.38) | (1.38) | (1.53) | (1.45) |
| Local stock index | | -0.00826 | 0.0405*** | 0.0602*** | 0.0595*** | 0.0596*** | 0.0611*** |
| | | (0.67) | (3.18) | (4.91) | (4.87) | (4.24) | (4.64) |
| Loans/officer | | -0.0144*** | -0.00827** | -0.0107** | -0.0129*** | -0.0111** | -0.0109** |
| | | (3.03) | (2.12) | (2.54) | (2.91) | (2.41) | (2.47) |
| Borrower education | | -0.241*** | -0.215*** | -0.251*** | -0.211*** | -0.253*** | -0.252*** |
| | | (4.91) | (4.35) | (5.99) | (4.69) | (5.52) | (5.74) |
| Borrower age | | -0.0148*** | -0.0130*** | -0.0111*** | -0.0142*** | -0.0126*** | -0.0118*** |
| | | (11.65) | (10.76) | (10.57) | (10.36) | (11.02) | (10.80) |
| Borrower female | | 0.0724** | 0.0934*** | -0.0146 | 0.0439 | -0.0273 | -0.0216 |
| | | (2.12) | (2.73) | (0.58) | (1.40) | (1.01) | (0.84) |
| Urban borrower | | 0.561*** | 0.155*** | 0.0777* | 0.0792** | 0.0839* | 0.0793* |
| | | (13.91) | (4.06) | (1.92) | (1.98) | (1.90) | (1.88) |
| Stable income | | -0.151* | -0.138* | 0.0254 | 0.0243 | 0.0671 | 0.0467 |
| | | (1.68) | (1.69) | (0.28) | (0.27) | (0.69) | (0.50) |
| Interest rate | | | 0.0159* | 0.00902 | -0.0116 | 0.0211*** | 0.0151** |
| | | | (1.85) | (1.26) | (1.35) | (2.76) | (2.05) |
| Loan maturity | | | 0.00631** | 0.0219*** | 0.0275*** | 0.00427 | 0.0147*** |
| | | | (2.16) | (8.13) | (9.68) | (1.30) | (4.96) |
| Loan/income ratio | | | 0.0426*** | 0.0282*** | 0.0273*** | 0.0301*** | 0.0292*** |
| | | | (3.77) | (3.27) | (3.20) | (3.08) | (3.18) |
| Loan immovable | | | -0.367** | -0.479*** | -0.482*** | -0.519*** | -0.498*** |
| | | | (2.29) | (3.96) | (4.03) | (3.81) | (3.87) |
| Loan movable | | | -0.0760 | -0.145 | -0.147 | -0.197 | -0.170 |
| | | | (0.51) | (1.22) | (1.25) | (1.46) | (1.33) |
| Loan stock | | | 0.113 | 0.0586 | 0.0564 | 0.0320 | 0.0463 |
| | | | (0.77) | (0.54) | (0.52) | (0.26) | (0.40) |
| Loan household | | | -0.129 | -0.0626 | -0.0665 | -0.136 | -0.0964 |
| | | | (0.79) | (0.48) | (0.51) | (0.93) | (0.70) |
| Personal collateral | | | 1.561*** | 1.593*** | 1.666*** | 1.438*** | 1.508*** |
| | | | (19.40) | (21.80) | (20.45) | (18.25) | (19.93) |
| Social collateral | | | 0.0428 | 0.290*** | 0.382*** | 0.160** | 0.223*** |
| | | | (0.57) | (3.73) | (4.40) | (1.98) | (2.82) |
| Third-party collateral | | | 1.895*** | 1.985*** | 2.114*** | 1.939*** | 1.955*** |
| | | | (22.70) | (25.90) | (26.25) | (20.96) | (23.18) |
| First loan | | | | 0.589*** | 0.648*** | 0.538*** | 0.564*** |
| | | | | (15.55) | (14.96) | (13.37) | (14.44) |
| Post CRK*First loan | | | | -0.196** | -0.200** | -0.247** | -0.222** |
| | | | | (2.04) | (2.07) | (2.44) | (2.26) |
| Constant | | | | | | -6.087*** | -4.942*** |
| | | | | | | (14.27) | (12.17) |
| Alpha | | | | | | | 0.525*** |
| | | | | | | | (115.03) |
| Competition measure | HHI | HHI | HHI | HHI | HHI | HHI | HHI |
| No. of obs | 70,804 | 55,034 | 55,034 | 185,935 | 185,935 | 185,935 | 185,935 |
| No. of branches | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| Log-likelihood ratio | -60,116.7 | -56,757.7 | -54,832.0 | -119,509.3 | -119,307 | -52,901.6 | -49,771.0 |
| Proportionality assumption | Yes | Yes | Yes | Yes | No | na | na |

*Notes:* This table shows the results of Cox proportional hazard models in column [1] to [4], a Cox non-proportional hazard model in [5], a parametric exponential hazard model in [6] and a parametric Weibull hazard model in [7]. 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 of by EKI. Sample period: June 2002-December 2010. We restrict the sample to first-time borrowers in columns [1]-[3]. Post CRK is a dummy variable that is "1" if the CRK was in place in a given month, zero otherwise. Local competition: Dummy variable that is '1' if local credit-market competition is above the median level of competition as measured by 1 minus the HHI index where local market shares are measured in number of branches. Robust standard errors are clustered by loan officer and z-statistics appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% significance level, respectively. Table A.1 in the Appendix contains all variable definitions.

## 4.2 Information sharing, lender competition and loan conditions

In Table 3, we proceed with our analysis of the change in lending conditions around the CRK introduction to gauge to what extent EKI loan officers reacted to the new credit registry by adjusting their lending at the intensive margin. In the left-hand side we use *Loan amount*, *Loan maturity*, and *Interest rate* as the key loan characteristics of interest. In the odd columns we include the CRK dummy by itself whereas in the even columns we also interact this dummy with our *Local competition* variable.

It becomes directly apparent that the introduction of mandatory information sharing was accompanied by a reduction in both loan amounts and loan maturities and an increase in the interest rate charged. All of these effects are statistically significant, stronger in the relatively competitive areas, and hold when including our standard set of borrower and other covariates.

More specifically, post CRK introduction loans showed an overall decrease in size of 6 percentage. With a much stonger reduction in size of 22 percentage points in high competition areas. The same pattern can be found looking at loan maturity, where loans were 6 percentage points shorter overall and 14.5 percentage points shorter in high competition areas. Smaller loans did not lead to lower interest rates as they were 0.7 percentage points higher overlall and 1 percentage point higher in competitive areas.

They suggest clearly that with the introduction of the CRK, loan officers tightened their lending conditions at the intensive margin. Interestingly, when we compare the (firs-time) borrower population before and after the introduction of the CRK along a number of observable characteristics, we do not find many substantive differences. From Table A.2 we see that even if differences are statistically significant, they are minor in economics terms. The only variables that see a moderate change are income (3% decrease) and the percentage of borrowers with stable employment (3% decrease). This indicates that EKI did not react to the CRK introduction by shifting its lending to a different type of borrower (nor did the number of loans decline). Instead, we find that, given certain borrower characteristics, loan amounts and maturities went down. At the same time, the improved borrower information from the CRK seems to have contributed to a relatively sharp improvement in borrower quality (again, this holds when we control for a large batch of standard observable borrower characteristics).

The covariate coefficients show that older, more highly educated, higher-income, and rural borrowers receive larger loans at lower interest rates. This squares with our earlier finding that these borrowers tend to have lower default rates too. In branches where the loan volume expands more rapidly, loan amounts tend to be higher and interest rates lower, most likely reflecting a residual competition effect not picked up by our competition variable.

In Table 4 we show a number of robustness tests. The same borrower covariates as in Table 3 are included but not shown for reasons of brevity. In columns 1, 2, 4, 5, 7 and 8 we re-estimate our results over two time windows that are

Table 3: Information sharing, credit-market competition, and loan characteristics

| **Dependent variable →** | **Loan amount** | | **Loan maturity** | | **Interest rate** | |
|---|---|---|---|---|---|---|
| | [1] | [2] | [3] | [4] | [5] | [6] |
| Post CRK intro | -0.145*** | -0.0582** | -0.104*** | -0.0601*** | 0.886*** | 0.723*** |
| | (7.25) | (2.00) | (7.32) | (3.25) | (12.66) | (7.49) |
| Post CRK intro*Local competition | | -0.162*** | | -0.0846*** | | 0.305** |
| | | (4.54) | | (3.48) | | (2.30) |
| Local competition | | -0.0283 | | -0.0488** | | 0.0813 |
| | | (1.09) | | (2.59) | | (1.02) |
| Borrower age | 0.00286*** | 0.00300*** | 0.00161*** | 0.00171*** | -0.0102*** | -0.0105*** |
| | (6.89) | (7.16) | (5.54) | (5.93) | (6.90) | (7.07) |
| Borrower female | 0.0167* | 0.0187** | -0.0234*** | -0.0219*** | -0.0976*** | -0.102*** |
| | (1.89) | (2.16) | (3.12) | (2.97) | (2.64) | (2.79) |
| Borrower education | 0.0314** | 0.0322** | -0.00728 | -0.00636 | -0.0389 | -0.0408 |
| | (1.99) | (2.10) | (0.63) | (0.56) | (0.77) | (0.81) |
| Borrower income | 0.400*** | 0.396*** | 0.0424*** | 0.0372*** | -0.171*** | -0.160*** |
| | (22.59) | (23.16) | (3.81) | (3.41) | (3.02) | (2.86) |
| Stable employment | -0.0511*** | -0.0631*** | -0.0364*** | -0.0493*** | 0.0244 | 0.0523 |
| | (2.75) | (3.42) | (2.66) | (3.46) | (0.43) | (0.95) |
| House members | -0.00165 | -0.00106 | 0.0178*** | 0.0182*** | -0.130*** | -0.131*** |
| | (0.29) | (0.19) | (4.42) | (4.56) | (6.57) | (6.57) |
| Number of dependents | 0.00961 | 0.0105 | -0.00262 | -0.00231 | -0.0217 | -0.0232 |
| | (1.42) | (1.55) | (0.55) | (0.49) | (0.96) | (1.01) |
| Minority | 0.0314 | 0.0314 | 0.0150 | 0.0138 | -0.118 | -0.117 |
| | (0.61) | (0.59) | (0.39) | (0.37) | (0.76) | (0.73) |
| Disabled | -0.0478 | -0.0525 | -0.0225 | -0.0248 | 0.0878 | 0.0965 |
| | (1.16) | (1.28) | (0.79) | (0.86) | (0.78) | (0.86) |
| Client urban | -0.0649*** | -0.0671*** | -0.0138 | -0.0156* | 0.370*** | 0.375*** |
| | (5.40) | (5.68) | (1.41) | (1.69) | (8.51) | (8.91) |
| Investment fund index | 0.0000769*** | 0.0000732*** | 0.0000465*** | 0.0000416*** | -0.000940*** | -0.000931*** |
| | (8.03) | (7.60) | (6.73) | (6.09) | (30.62) | (29.33) |
| Lending growth volume | 0.109*** | 0.101*** | 0.0640*** | 0.0555*** | -0.742*** | -0.724*** |
| | (5.18) | (4.78) | (4.05) | (3.50) | (9.93) | (9.73) |
| Personal collateral | 0.373*** | 0.372*** | 0.305*** | 0.305*** | -0.631*** | -0.628*** |
| | (29.72) | (29.36) | (31.52) | (31.15) | (18.11) | (17.98) |
| Social collateral | 0.239*** | 0.249*** | 0.110*** | 0.118*** | -0.555*** | -0.576*** |
| | (25.06) | (25.84) | (15.20) | (17.11) | (18.80) | (19.57) |
| Third-party collateral | 0.359*** | 0.358*** | 0.186*** | 0.186*** | -0.636*** | -0.635*** |
| | (16.41) | (16.45) | (12.03) | (12.09) | (11.95) | (11.88) |
| Local econ. control | Yes | Yes | Yes | Yes | Yes | Yes |
| No. of obs. | 17,035 | 17,035 | 17,035 | 17,035 | 17,035 | 17,035 |
| No. of branches | 15 | 15 | 15 | 15 | 15 | 15 |
| R-squared | 0.4337 | 0.4382 | 0.3129 | 0.3188 | 0.3124 | 0.3143 |

*Notes:* This table shows OLS regressions at the loan level to estimate the impact of the introduction of the Bosnian credit registry (CRK) on log of loan amounts ([1]-[2]); loan maturity ([3]-[4]); and interest rates ([5]-[6]) across branches that experience varying degrees of credit-market competition. Pre CRK introduction : July 2008-July 2009. Post CRK introduction: August 2009-August 2010. All specifications include a time-varying nigh-light measure of local economic activity. Constant not shown. Local competition: Dummy variable that is '1' if local credit-market competition is above the median level of competition as measured by 1 minus the HHI index where local market shares are measured in number of branches. Table A.1 in the Appendix contains all variable definitions. Sample contains first-time EKI borrowers only. Standard errors are robust and clustered by loan officer (t-statistics in parentheses). ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

either broader or narrower than our standard two-year window (in Table 3 the time window spanned July 2008-July 2009 (pre CRK) and August 2009-August 2010 (post CRK)). These are Feb 2009-Feb 2010 (Narrow window) and May 2008-Dec 2010 (Broader window). In columns 3, 6 and 9 we use the same sample as in Table 3 but we use an alternative competition measure, reflecting the perceptions of EKI loan officers as culled from a survey among loan officers in all EKI branches (see Table A.1 for the exact definition). The results in Table 4 show that our earlier findings are robust to these changes in window width as well as the competition measure used.

Lastly, in Table 5 we again perform a set of four placebo tests to carefully check whether our results are not driven by any secular data trends that hitherto remained undetected. We first rerun our analysis but now consider the pre-CRK period as the actual treatment period. If a secular trend was driving our results,

Table 4: Information sharing, credit-market competition, and loan characteristics: *Robustness tests*

| **Dependent variable →** | **Loan amount** | | | **Loan maturity** | | | **Interest rate** | | |
|---|---|---|---|---|---|---|---|---|---|
| | Narrow window | Broader window | Perceived competition | Narrow window | Broader window | Perceived competition | Narrow window | Broader window | Perceived competition |
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] |
| Post CRK intro*Local competition | -0.137*** | -0.137*** | -0.144*** | -0.0906*** | -0.0852*** | -0.0740** | 0.530*** | 0.239* | 0.482*** |
| | (3.30) | (3.93) | (3.46) | (3.06) | (3.50) | (2.45) | (3.51) | (1.91) | (3.50) |
| Borrower covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Local econ. control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| No. of loans | 7,357 | 21,391 | 16,523 | 7,357 | 21,391 | 16,523 | 7,357 | 21,391 | 16,523 |
| No. of branches | 15 | 15 | 14 | 15 | 15 | 14 | 15 | 15 | 14 |
| Adjusted R-squared | 0.4228 | 0.4262 | 0.434 | 0.2892 | 0.3062 | 0.309 | 0.1974 | 0.3225 | 0.316 |

*Notes:* This table shows robustness tests of our main results as reported in Table 2 (where the pre CRK period was July 2008-July 2009 and the post CRK period August 2009-August 2010). In columns [1], [2], [4], [5], [7] and [8] we re-estimate our results over two time windows that are either longer or shorter than our standard two-year window These are Feb 2009-Feb 2010 (Narrow window) and May 2008-Dec 2010 (Broader window). In columns [3], [6] and [9] we use the same sample as in Table 2 but we use an alternative perception based measure of competition. Robust standard errors are clustered by loan officer and t-statistics appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively. Table A.1 in the Appendix contains all variable definitions. The same borrower covariates as in Table 3 are included but not shown.

Table 5: Information sharing, credit-market competition, and loan characteristics: *Placebo tests*

| **Dependent variable →** | **Loan amount** | | | | **Loan maturity** | | | | **Interest rate** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pre period is post-CRK period | Similar period 2003 05 | Lehman Brothers as break | Structural break Oct-2010 | Pre period is post-CRK period | Similar period 2003 05 | Lehman Brothers as break | Structural break Sept-2006 | Pre period is post-CRK period | Similar period 2003 05 | Lehman Brothers as break | Structural break Feb-2009 |
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] | [10] | [11] | [12] |
| Post CRK intro*Local competition | 0.000992 | 0.0120 | -0.00277 | -0.00967 | -0.0255 | 0.00530 | -0.0285 | 0.0673* | -0.126 | -0.0490 | 0.195** | 0.0244 |
| | (0.02) | (0.22) | (0.10) | (0.32) | (0.74) | (0.24) | (1.40) | (1.91) | (0.96) | (0.56) | (1.98) | (0.20) |
| Borrower covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Local econ. control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| No. of loans | 7,803 | 8,000 | 19,596 | 23,642 | 7,803 | 8,000 | 19,596 | 15,898 | 7,803 | 8,000 | 19,596 | 20,916 |
| No. of branches | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| Adjusted R-squared | 0.3556 | 0.2201 | 0.4804 | 0.4693 | 0.2182 | 0.1492 | 0.3647 | 0.2844 | 0.1443 | 0.1513 | 0.2679 | 0.3459 |

*Notes:* In columns [1], [5] and [9] we show results for a placebo test where the control (pre) period is the actual treatment period. In columns [2], [6] and [10] we show results for a placebo test where we estimate over a period with comparable lending growth five years before the actual CRK introduction. In columns [3], [7] and [11] we show results for a placebo test where the placebo treatment starts in September 2008 (Lehman brother bankrupcy) and ends with the introduction of the CRK in July 2009. In columns [4], [8] and [12] we show results for a placebo test where the placebo treatment starts in October 2010 for loan size, September 2006 for loan maturity and February 2009 for interest rate (a Clemente-Montañés-Reyes unit-root test indicates a break point in that month for each dependent variable). In all regressions, the competition measure is (1-HHI). Robust standard errors are clustered by loan officer and t-statistics appear in parentheses. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively. Table A.1 in the Appendix contains all variable definitions. The same borrower covariates as in Table 3 are included but not shown.

we would expect to still find a result here. Second, we show results for a placebo test where we estimate over a period with comparable levels of credit growth exactly five years before the actual CRK introduction. Third, we perform a test where the placebo treatment starts in September 2008 - the collapse of Lehman Brothers - and ends with the introduction of the CRK in July 2009. If we would simply pick up a crisis effect, this should show up here. Fourth and finally, we show results for a placebo test where the placebo treatment starts in October 2010 for *Loan size*, September 2006 for *Loan maturity* and February 2009 for *Interest rate*. These placebo start times are chosen on the basis of a Clemente-Montañés-Reyes unit-root test which indicates a possible break point in that month for each dependent variable. The results in Table 5 show that throughout all these placebo tests, our original results disappear. This suggests that we are pick up the true CRK effect and not another trend or break-point such as the global financial crisis.

## 5 Conclusion

In this paper, we have presented evidence of what happens when lenders are required to share borrower information in a competitive credit market. In doing so, we have successfully connected the literature on the effects of information sharing to the literature on the relationship between bank competition and credit availability.

We have carried out our analysis using a unique data set built around the loan portfolio of one of the largest providers of small business loans in Bosnia and Herzegovina (BiH). We match over 200,000 loans to both borrower and branch characteristics, and use both standard concentration measures and loan officers' own perceptions to measure local competition. By exploiting the time variation in information sharing and cross-sectional variation in competition, we have been able to identify the impact of a mandatory credit register on loan quality. In addition, we have analyzed whether the improvement in information sharing has had an impact on loan conditions such as size, maturity and interest rates. In doing so, we have drawn attention to the role of competition in local markets and to the (in)existence of a credit history, by separating high-competition markets from low-competition markets and first-time borrowers from returning borrowers.

Our analysis has emphasized the importance of connecting the literature on competition and credit conditions with the literature on information sharing. When we limited our analysis to a comparison of high versus low competitition markets, we initially found that there was only a limited difference in the survival behavior among borrowers. Turning only to the impact of the CRK, we found that loans granted after the introduciton of the CRK displayed a significantly higher survival probability compared to loans provided before mandatory information sharing was introduced, as well a some evidence of strategic defaults prior to the introduction of the CRK. Combining both elements, however, we observed that with the CRK in place, repayment performance was significantly higher in high-competition areas, where repayment rates before information sharing was in place were significantly below the level attainable in the presence of information exchange.

In line with theory, our results appeared to be particularly pronounced for first-time borrowers, for whom the information asymmetry between lender and prospective borrower was the largest. Indied, for repeat borrowers, we found only limted evidence of increased borrower discipline after the introduction of the CRK. The impact of the CRK introduction was indeed be much larger for first-time borrowers, in particular in high-competition areas, suggesting that the CRK mainly 'worked' through the lender side.

Further analyses showed that the introduction of mandatory information sharing was accompanied by a reduction in both loan amounts and loan maturities and an increase in the interest rate charged, especially in high-competition areas.

While our approach has clear strengths and is the first to provide empirical micro evidence on the interaction between lender competition and a sudden change in the information sharing regime, it has drawbacks as well. First, while we have access to the complete loan portfolio of a large lender, this is just one lender in one country. This of course limits the external validity of our findings. Having said that, the Bosnian market for small business loans has much in common with other rapidly expanding (individual-liability) credit markets in other middle-income countries. Second, we have no information on loan rejections, which means that our analysis takes place at the intensive not the extensive margin. We note, however, that at the portfolio level the data from our lender shows no major shift in the type of client while at the same time there was no decline in overall amount of lending or number of loans. Instead, the improvement in loan quality appears to follow from changes in loan terms give to the same type of clients (that is, a tightening at the intensive margin) and a better client selection along unobservable characteristics based on the information drawn from the credit registry.

Our results point to a number of policy implications regarding the most effective design and implementation of information sharing regimes. The most important determinant of the success of a credit registry is coverage. Lack of information on a subset of borrowers or lenders will seriously hamper the effectiveness of any information sharing system. In Bosnia both the private credit bureau with voluntary participation and bank only public credit registry were unable to mitigate over borrowing, Participation should then be mandatory and extended to all types of lender.

# References


Allen N Berger and Gregory F Udell. Relationship lending and lines of credit in small firm finance. *Journal of business*, pages 351–381, 1995.

Mark Berryman and Justyna Pytkowska. A review of the bosnian microfinance sector: move to financial self-sufficiency. *Microfinance Centre for Central and Eastern Europe and the New Independent States*, 2003.

Helmut Bester. The role of collateral in credit markets with imperfect information. *European Economic Review*, 31(4):887–899, 1987.

Sudipto Bhattacharya and Gabriella Chiesa. Proprietary information, financial intermediation, and research incentives. *Journal of Financial Intermediation*, 4(4):328–357, 1995.

Emilia Bonaccorsi di Patti and Giovanni Dell'Ariccia. Bank competition and firm creation. *Journal of Money, Credit and Banking*, 36(2):225–51, April 2004.

Arnoud WA Boot and Anjan V Thakor. Can relationship banking survive competition? *The journal of Finance*, 55(2):679–713, 2000.

Norman Breslow. Covariance analysis of censored survival data. *Biometrics*, pages 89–99, 1974.

Martin Brown and Christian Zehnder. The emergence of information sharing in credit markets. *Journal of Financial Intermediation*, 19(2):255–278, 2010.

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

Nicola Cetorelli and Michele Gambera. Banking market structure, financial dependence and growth: International evidence from industry data. *The Journal of Finance*, 56(2):617–648, 2001.

Greg Chen, Stephen Rasmussen, and Xavier Reille. Growth and vulnerabilities in microfinance. *Focus Note*, 61, 2010.

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

DR Cox and PAWL Lewis. The statistical analysis of series of events. 1966.

Alain De Janvry, Craig McIntosh, and Elisabeth Sadoulet. The supply-and demand-side impacts of credit market information. *Journal of Development Economics*, 93(2):173–188, 2010.

Hans Degryse and Steven Ongena. The impact of competition on bank orientation. *Journal of Financial Intermediation*, 16(3):399–424, 2007.

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

Asli Demirgüç-Kunt, Leora Klapper, and Giorgios Panos. The origins of self-employment. *Development Research Group (February). Washington DC: World Bank*, 2007.

Douglas W Diamond. Reputation acquisition in debt markets. *The Journal of Political Economy*, pages 828–862, 1989.

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

Ralf Elsas. Empirical determinants of relationship lending. *Journal of Financial Intermediation*, 14(1):32–57, 2005.

Jasmina Glisovic, Henry González, Yasemin Saltuk, and Frederic Rozeira de Mariz. Volume growth and valuation contraction global microfinance equity valuation survey 2012. *CGAP and JP. Morgan. Washington, DC*, 2012.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Does local financial development matter? *The Quarterly Journal of Economics*, 119(3):929–969, 2004.

Valentina Hartarska and Denis Nadolnyak. An impact analysis of microfinance in bosnia and herzegovina. *World Development*, 36(12):2605–2619, 2008.

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

Vernon Henderson, Adam Storeygard, and David N Weil. A bright idea for measuring economic growth. *The American Economic Review*, 101(3):194–199, 2011.

Karla Hoff and Joseph E Stiglitz. Moneylenders and bankers: price-increasing subsidies in a monopolistically competitive market. *Journal of Development Economics*, 55(2):485–518, 1998.

Joel F Houston, Chen Lin, Ping Lin, and Yue Ma. Creditor rights, information sharing, and bank risk taking. *Journal of Financial Economics*, 96(3):485–512, 2010.

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

Edward L Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

Artashes Karapetyan and Bogdan Stacescu. Information sharing and information acquisition in credit markets. 2008.

Nicholas M Kiefer. Economic duration data and hazard functions. *Journal of economic literature*, 26(2):646–679, 1988.

Cédric Lützenkirchen, Christian Weistroffer, and Bernhard Speyer. Microfinance in evolution. *An industry between crisis and advancement, DB Research, September, under http://www. dbresearch. de/MAIL/DBR_INTERNE T_DE-PROD/PROD0000000000294314. pdf (accessed 20.09. 2012)*, 2012.

Timothy Lyman, Tony Lythgoe, and Xavier Reille Margaret Miller, and Shalini Sankaranarayan. Credit reporting at the base of the pyramid: Key issues and success factors. Technical report, FORUM 1. Washington, D.C.: CGAP and IFC., 2011.

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

Steven Ongena and David C Smith. The duration of bank relationships. *Journal of Financial Economics*, 61(3):449–475, 2001.

A Jorge Padilla and Marco Pagano. Endogenous communication among lenders and entrepreneurial incentives. *Review of Financial Studies*, 10(1):205–36, 1997.

A Jorge Padilla and Marco Pagano. Sharing default information as a borrower discipline device. *European Economic Review*, 44(10):1951–1980, 2000.

Marco Pagano and Tullio Jappelli. Information sharing in credit markets. *The Journal of Finance*, 48(5):1693–1718, 1993.

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

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

Tara Rice and Philip E. Strahan. Does credit competition affect small-firm finance? *The Journal of Finance*, 65(3):861–889, 2010. ISSN 1540-6261.

Kasper Roszbach. Bank lending policy, credit scoring, and the survival of loans. *The Review of Economics and Statistics*, 86(4):946–958, November 2004.

Tyler Shumway. Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74(1):101–124, 2001.

Judith D Singer and John B Willett. 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, 1993.

Maria Stepanova and Lyn Thomas. Survival analysis methods for personal loan data. *Operations Research*, 50(2):277–289, 2002.

Joseph E Stiglitz and Andrew Weiss. Incentive effects of terminations: Applications to the credit and labor markets. *The American Economic Review*, 73 (5):912–927, 1983.

Anne Welle-Strand, Kristian Kjøllesdal, and Nick Sitter. Assessing microfinance: The bosnia and herzegovina case. *Managing Global Transitions*, 8(2):145–166, 2010.

Warner Woodworth. Microcredit in post-conflict, conflict, natural disaster, and other difficult settings. *Global Microcredit Summit*, 2006.

# Appendix

Table A.1: Variable definitions and data sources

| | Definition | Source | Unit |
|---|---|---|---|
| *Dependent variables:* | | | |
| Loan amount | Loan amount at time of disbursement. | EKI | BAM |
| Loan/income ratio | Loan amount at time of disbursement divided by annual borrower income. Income includes primary plus secondary income. | EKI | Ratio |
| Loan maturity | Length (tenor) of the loan at time of disbursement | EKI | Months |
| Poor borrower | Dummy=1 if borrower belongs to the poorest 30 percent of the population as measured by annual income; 0 otherwise. | EKI | Dummy |
| Interest rate | Annual nominal interest rate on loan | EKI | Percentage |
| Problem loan | Dummy=1 if loan i is defaulted on in month t; 0 otherwise. A default event occurs when a borrower is at least 30 days late in making a payment and the loan was eventually written of by EKI. | EKI | Dummy |
| *Independent variables:* | | | |
| CRK introduction | Dummy=1 for all quarters after and including August 2009 (time of CRK introduction); 0 otherwise. | Central Bank of Bosnia | Dummy |
| Local competition: 1-HHI | 1 minus HHI index. The (time-varying) HHI index ranges between [0, 1] and measures microcredit market concentration in the locality where an EKI branch is based. Market shares are expressed as number of branches. | BEPS, MIX, Annual reports | [0, 1] |
| Local competition: Survey | Competition intensity as perceived by the two most senior loan officers in each branch. Average score on a 7-point Likert scale to the question: "Over the past ten years, I think that other microcredit providers have increased their competitiveness in my area". | Loan officer survey | 0.5 increments |
| Borrower age | Borrower age. | EKI | Years |
| Borrower female | Dummy= 1 if borrower is female; 0 otherwise. | EKI | Dummy |
| Borrower education | 1 = None, 2 = Primary, 3 = Secondary, 4 = Tertiary (College/University/Post Graduate). | EKI | 1 to 4 |
| Borrower income | Total annual borrower income (primary plus secondary income source). | EKI | BAM |
| Urban borrower | 0 = Rural; 1 = Urban. | EKI | Dummy |
| Stable income | 0 = unemployed or casually employed; 1 = stable employment (agricultural producer; full-time employed; own business; part-time employed) or pension. | 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 |
| Loan household | Loan purpose = Private (non-business related) expenses for the household. | EKI | Dummy |
| Personal collateral | Dummy=1 if borrower posted at least one type of personal collateral; 0 otherwise. Personal collateral includes mortgages, administrative bans on the borrower's salary and pledges of movable assets. | EKI | Dummy |
| Social collateral | Dummy=1 if loan was guaranteed by two or more guarantors; 0 otherwise. | EKI | Dummy |
| Third-party collateral | Dummy=1 if at least one third-party collateral (checks or bills of exchange issued by a guarantor company was posted); 0 otherwise. | EKI | Dummy |
| Stock index | Bosnia Investment Index (May 28th 2002=1). | Sarajevo Stock Exchange | Index |
| Local GDP | Time varying measure of local economic activity as proxied by the night-light intensity (derived from satellite images) in the locality where an EKI branch is based. Scale ranges from 0 to 63 where higher value indicate higher light intensity. | National Geophysical Datacenter; Henderson et al. (2011) | [0, 63] |
| Loans/officer | Monthly number of loans per loan officer. | EKI | Loans |
| Branch growth | Quarterly growth in total new lending volume (flow) per branch. | EKI | Percentage |

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

Table A.2: Summary statistics

| | Mean pre-CRK | Mean post-CRK | Obs. | Median | St. dev. | Min | Max |
|---|---|---|---|---|---|---|---|
| *Dependent variables:* | | | | | | | |
| Loan amount (BAM) | 3,567.80 | 3079.17*** | 191,956 | 3,000 | 2,871.91 | 500 | 15,000 |
| Loan/income ratio | 3.18 | 2.89*** | 191,956 | 2.50 | 2.34 | 0.44 | 11.93 |
| Loan maturity | 22.98 | 23.50*** | 191,956 | 19 | 11.38 | 6 | 60 |
| Interest rate | 18.56 | 21.66*** | 191,956 | 19 | 4.13 | 12 | 26 |
| Problem loan | 0.06 | 0.03*** | 191,956 | 0 | 0 | 0 | 1 |
| | | | | | | | |
| *Independent variables:* | | | | | | | |
| CRK introduction | 0 | 1 | 191,956 | 0 | 0.36 | 0 | 1 |
| Competition: 1-HHI | 0.81 | 0.80*** | 191,956 | 0.81 | 0.06 | 0.56 | 0.90 |
| Perceived competition | 4.98 | 5.13*** | 189,248 | 5.50 | 1.17 | 3 | 6.50 |
| Borrower age | 40.09 | 42.01*** | 110,294 | 39.50 | 11.92 | 20 | 68 |
| Borrower male | 0.59 | 0.60*** | 110,294 | 1 | 0.49 | 0 | 1 |
| Borrower education | 1.93 | 1.94*** | 110,225 | 2 | 0.39 | 1 | 3 |
| Borrower monthly income (BAM) | 1,216 | 1,168*** | 110,294 | 1,036 | 586.80 | 350 | 3,800 |
| Urban borrower | 0.39 | 0.35*** | 89,021 | 0 | 0.49 | 0 | 1 |
| Stable income | 0.86 | 0.83*** | 110,295 | 1 | 0.35 | 0 | 1 |
| Loan immovable | 0.08 | 0.11*** | 191,956 | 0 | 0.28 | 0 | 1 |
| Loan movable | 0.43 | 0.49*** | 191,956 | 0 | 0.50 | 0 | 1 |
| Loan stock | 0.41 | 0.20*** | 191,956 | 0 | 0.48 | 0 | 1 |
| Loan household | 0.07 | 0.17*** | 191,956 | 0 | 0.28 | 0 | 1 |
| Personal collateral | 0.25 | 0.52*** | 191,956 | 0 | 0.57 | 0 | 2 |
| Social collateral | 1.98 | 2.16*** | 191,956 | 2 | 1.11 | 1 | 6 |
| Third-party collateral | 0.04 | 0.09*** | 191,956 | 0 | 0.28 | 0 | 2 |
| Stock index (quarterly) | 4.36 | 1.68*** | 186,187 | 3.77 | 2.12 | 1.29 | 8.35 |
| Local GDP (night light measure) | 26.47 | 26.44 | 191,956 | 26.04 | 9.67 | 5.96 | 46.36 |
| Loans/officer | 21.39 | 16.27*** | 191,956 | 20 | 9.57 | 2 | 46 |
| Branch growth (quarterly) | 0.06 | 0.06 | 192,301 | 0.03 | 0.27 | -0.51 | 1 |

*Notes:* Sample period is June 2002-December 2010.

Table A.3: Information sharing and loan quality: Hazard model extensions and alternative specifications

| | [1] | [2] | [3] | [4] | [5] | [6] |
|---|---|---|---|---|---|---|
| **Functional form** | **Cox proportional** | | **Exponential** | | | **Weibull** |
| **Time structure** | **Time-varying predictors** | | | | | |
| Post CRK*local competition | -0.301** | -0.247** | -0.268* | -0.203** | -0.264* | -0.200* |
| | (1.97) | (2.44) | (1.72) | (1.98) | (1.71) | (1.96) |
| Post CRK | -1.332*** | -1.230*** | -0.789*** | -0.732*** | -0.922*** | -0.856*** |
| | (11.60) | (15.41) | (6.68) | (8.82) | (7.91) | (10.41) |
| Local competition | -0.0556* | -0.102*** | -0.0488 | -0.0960*** | -0.0690** | -0.119*** |
| | (1.85) | (4.57) | (1.42) | (3.87) | (2.09) | (4.96) |
| Local GDP | -0.00675*** | -0.00746*** | -0.0140*** | -0.0145*** | -0.0114*** | -0.0120*** |
| | (4.57) | (6.67) | (8.61) | (12.00) | (7.13) | (10.11) |
| Local stock index | 0.0000192** | 0.0000269*** | 0.0000510*** | 0.0000422*** | 0.000120*** | 0.000108*** |
| | (2.24) | (4.30) | (5.17) | (6.14) | (14.91) | (18.90) |
| Loans/officer | -0.00745*** | -0.00996*** | -0.00982*** | -0.0117*** | -0.00942*** | -0.0116*** |
| | (5.35) | (9.27) | (6.33) | (9.99) | (6.24) | (10.10) |
| Borrower education | -0.229*** | -0.250*** | -0.238*** | -0.251*** | -0.238*** | -0.253*** |
| | (6.21) | (9.00) | (5.73) | (8.27) | (5.92) | (8.55) |
| Borrower age | -0.0135*** | -0.0117*** | -0.0151*** | -0.0131*** | -0.0146*** | -0.0126*** |
| | (11.58) | (13.01) | (11.50) | (13.16) | (11.48) | (12.99) |
| Borrower female | 0.0967*** | -0.00291 | 0.0909*** | -0.0203 | 0.0920*** | -0.0146 |
| | (3.32) | (0.14) | (2.75) | (0.85) | (2.89) | (0.63) |
| Urban borrower | 0.189*** | 0.118*** | 0.216*** | 0.148*** | 0.226*** | 0.153*** |
| | (7.40) | (6.11) | (7.49) | (6.96) | (8.18) | (7.43) |
| Stable income | -0.162*** | 0.00546 | -0.119*** | 0.0471 | -0.132*** | 0.0299 |
| | (4.08) | (0.18) | (2.63) | (1.43) | (3.02) | (0.93) |
| Interest rate | 0.00102 | -0.00726* | 0.0111* | 0.00732* | -0.00238 | -0.00588 |
| | (0.17) | (1.77) | (1.68) | (1.67) | (0.38) | (1.40) |
| Loan maturity | -0.00681*** | 0.000831 | -0.0183*** | -0.00865*** | -0.0130*** | -0.00392*** |
| | (4.26) | (0.77) | (9.52) | (6.87) | (7.12) | (3.27) |
| Loan/income ratio | 0.0464*** | 0.0387*** | 0.0465*** | 0.0355*** | 0.0416*** | 0.0323*** |
| | (7.18) | (8.47) | (6.06) | (6.77) | (5.70) | (6.41) |
| Loan immovable | -0.460*** | -0.493*** | -0.550*** | -0.572*** | -0.570*** | -0.591*** |
| | (3.66) | (5.21) | (3.87) | (5.50) | (4.18) | (5.86) |
| Loan movable | -0.0909 | -0.0746 | -0.160 | -0.146 | -0.145 | -0.130 |
| | (0.77) | (0.83) | (1.19) | (1.48) | (1.12) | (1.36) |
| Loan stock | 0.105 | 0.138 | 0.0655 | 0.0922 | 0.0651 | 0.100 |
| | (0.88) | (1.52) | (0.48) | (0.93) | (0.50) | (1.04) |
| Loan household | -0.119 | 0.0207 | -0.197 | -0.0896 | -0.111 | 0.00596 |
| | (0.93) | (0.21) | (1.37) | (0.84) | (0.81) | (0.06) |
| Personal collateral | 1.687*** | 1.787*** | 1.661*** | 1.705*** | 1.731*** | 1.779*** |
| | (41.95) | (63.91) | (36.39) | (55.65) | (39.71) | (60.38) |
| Social collateral | 0.141*** | 0.467*** | 0.0783 | 0.383*** | 0.123** | 0.433*** |
| | (2.90) | (12.73) | (1.45) | (9.76) | (2.36) | (11.31) |
| Third-party collateral | 1.991*** | 2.129*** | 2.030*** | 2.119*** | 2.049*** | 2.134*** |
| | (41.18) | (64.11) | (34.44) | (54.60) | (36.52) | (57.42) |
| First loan | | 0.683*** | | 0.665*** | | 0.684*** |
| | | (30.80) | | (27.35) | | (28.85) |
| Post CRK*First loan | | -0.329*** | | -0.219*** | | -0.244*** |
| | | (4.30) | | (2.70) | | (3.06) |
| Constant | | | -3.737*** | -4.759*** | -3.192*** | -4.217*** |
| | | | (16.12) | (28.08) | (14.33) | (25.72) |
| Alpha | | | | | 0.624*** | 0.633*** |
| | | | | | (68.42) | (89.28) |
| No. of obs. | 356,131 | 1,119,122 | 356,131 | 1,119,122 | 356,131 | 1,119,122 |
| No. of branches | 15 | 15 | 15 | 15 | 15 | 15 |
| Log-likelihood ratio | -49,419.9 | -101,919.4 | -20,653.5 | -41,799.5 | -20,115.8 | -40,842.3 |

*Notes:* This table shows the results of semi-parametric and parametric hazard models. 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 in an earlier month. A default event occurs when a borrower is at least 30 days late in making a payment and the loan was eventually written of by EKI. The hazard function specifications are: Cox proportional in columns [1] and [2], exponential in [3]-[4], Weibull in [5]-[6]. Models [1] to [6] are estimated on the period-loan dataset in order to allow for time-varying predictors and mitigate bias due to tied time observations. Sample period: June 2002-December 2010. We restrict the sample to first-time borrowers in columns [1], [3], and [5]. Post CRK is a dummy variable that is "1" if the CRK was in place in a given quarter, zero otherwise. Robust standard errors are clustered by loan, and z-statistics appear in parentheses. Local competition: Dummy variable that is '1' if local credit-market competition is above the median level of competition as measured by 1 minus the HHI index where local market shares are measured in number of branches. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively. Table A.1 in the Appendix contains all variable definitions.