DOES LOAN MATURITY MATTER IN RISK-BASED PRICING? EVIDENCE FROM CONSUMER LOAN DATA

Gabriela Kuvíková

CERGE EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

WORKING PAPER SERIES (ISSN 1211-3298)
Electronic Version

538

Working Paper Series 538
(ISSN 1211-3298)

Does Loan Maturity Matter in Risk-Based Pricing? Evidence from Consumer Loan Data

Gabriela Kuvíková

CERGE-EI
Prague, April 2015

ISBN 978-80-7343-343-7 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-335-1 (Akademie věd České republiky. Národohospodářský ústav)

Does Loan Maturity Matter in Risk-Based Pricing? Evidence from Consumer Loan Data

Gabriela Kuvíková*

Abstract

This paper investigates the role of loan contract terms in the performance of consumer credit. Taking advantage of a sample of accepted and rejected consumer loans from a Czech commercial bank, I estimate the elasticity of loan demand and find that borrowers with a high probability of default are more responsive to maturity than interest rate changes. I also argue that risk-based pricing may lead to an increase in loan maturity and loan default, rather than alleviating the adverse selection present on the lending market. Empirical evidence suggests that loan performance is time-dependent and default depends on the choice of loan duration.

Abstrakt

Tato práce zkoumá vliv podmínek v úvěrové smlouvě na výkon spotřebitelských úvěrů. S využitím vzorku přijatých a odmítnutých spotřebitelských úvěrů z české komerční banky odhaduji elasticitu poptávky po úvěrech a zjišťuji, že dlužníci s vysokou pravděpodobností nesplacení půjčky jsou více citliví na změny data splatnosti než na změny úrokových sazeb. Dále argumentuji, že tvorba cen na základě rizika může vést k prodloužení splatnosti úvěru a zvýšení pravděpodobnosti nesplacení úvěru spíše než k zmírnění nepříznivého výběru na trhu úvěrů. Empirické důkazy naznačují, že splacení úvěru je závislý na volbě doby trvání úvěru.

Keywords: credit scoring, consumer loans, asymmetric information

JEL Classification Numbers: D12, D14, D82, G21

---

* CERGE-EI is a joint workplace of Charles University in Prague and the Economics Institute of the Academy of Sciences of the Czech Republic.Address: CERGE-EI, P.O. Box 882, Politických vězňů 7, Prague 1, 111 21, Czech Republic.

I am grateful to Alena Bičáková, Randall K. Filer, Jan Hanousek, Štěpán Jurajda and Jan Kmenta for their valuable comments and guidance. This research was supported by a grant from the CERGE-EI Foundation under a program of the Global Development Network. All opinions expressed are those of the author and have not been endorsed by CERGE-EI or the GDN. All errors remaining in this text are the responsibility of the author.

1.Introduction

Over recent decades, substantial increases in the number of consumer loans[1] have been observed worldwide. Lending to individuals to finance the purchase of goods or services has become particularly popular in emerging markets. Despite the initial difficulties related to the availability of only minimal credit history on borrowers and pioneering methods used to evaluate the creditworthiness of borrowers, lending institutions instituted extensive provision of consumer loans. The quantitative importance of consumer loans in emerging markets can be illustrated using the example of the Czech Republic, where between 2000 and 2012 the total volume of consumer loans rose from CZK 31.1 bn to CZK 157.3 bn. [2]

The rapid growth of the consumer credit market has drawn increased attention to the asymmetric information present between lenders and borrowers. Stiglitz and Weiss's 1981 paper shows that lenders who are imperfectly informed about the default probability of borrowers (henceforth referred to as a borrower's 'riskiness') may suffer from adverse selection when deciding to grant a loan or not. Adverse selection occurs when, being aware of their own riskiness, "low-risk" borrowers with low probability of default will not be willing to pay increased prices for loans in the form of higher interest rates, while "high-risk" borrowers with a high probability of default will accept them. To minimize this, lenders may choose to deny loans rather than raise interest rates. As the price fails to regain equilibrium in the market, market imperfection appears. Stiglitz and Weiss (1981) define the solution of limiting the amount of credit as credit rationing

[1]The European Central Bank defines consumer loans in the following way: Credit for consumption (loans granted for mainly personal consumption of goods and services) includes loans to sole proprietors/unincorporated partnerships if the loan is predominantly used for personal consumption. Loans included in this category may or may not be collateralized by various forms of security or guarantee. Typical examples of loans in this category are loans granted for the financing of motor vehicles, furniture, domestic appliances and other consumer durables, holiday travel, etc. Loans to cover overdrafts and credit card loans also typically belong in this category. Lending for house purchase is excluded from this category.
Source: Manual on Monetary Financial Institution balance sheet statistics
https://www.ecb.europa.eu/pub/pdf/other/manualmfibalancesheetstatistics201204en.pdf?426543c0dbb56bb78f5afd978b44db17

[2]Source: Czech Statistical Office - Statistical Yearbook of the Czech Republic
http://www.czso.cz/csu/.nsf/engpubl/10n1-04-2004

equilibrium, a situation when certain borrowers are refused funds even if they are willing to pay higher interest rates, as lenders are already maximizing profit. According to Jaffee and Stiglitz (1990) lenders can also react to the adverse selection by offering multiple loan contract terms (i.e. loan packages with distinct loan amount, interest rate and maturity).

Differentiating interest rates based on the borrowers' riskiness (i.e. applying risk-based pricing of interest rates) is one such attempt to mitigate asymmetric information on the consumer loan market. A number of studies (Edelberg, 2006; Einav, Jenkins, and Levin, 2012) argue that borrowers are highly responsive to interest rate variations. Specifically, they provide evidence that risk-based pricing raises the borrowing costs of "high-risk" applicants'; and hence restricts the level of their debt.

Addressing excess loan demand under imperfect information becomes more important in a loan market where borrowers have liquidity constraints. An individual with liquidity constraints does not have sufficient funds to finance present consumption with income that will be accumulated in the future. Adams, Einav, and Levin (2009) show that this inability to reallocate funds over time can result in notable adverse selection (i.e. borrowers with high probability of default increase their debt amount). Supporting the results of the previous literature, Adams et al. (2009) highlight that risk-based pricing can effectively diminish the severity of the information problem (i.e. "high-risk" borrowers receive lower loan amounts). Nevertheless, in identifying loan demand and loan repayment the authors did not consider an important aspect for borrowers with liquidity constraints, the role of loan maturity.

Although practitioners and policymakers consider interest rates as a key driver of loan demand, the sensitivity of loan demand to maturity might be equally crucial. Estimating the demand elasticity with respect to both interest rate and maturity, Attanasio, Goldberg, and Kyriazidou (2008) and Karlan and Zinman (2008) show that borrowers with low income are more responsive to maturity changes than to interest rate changes. Their finding is consistent with binding liquidity constraints, a situation when borrowers with limited available cash choose longer loan maturity in order to reduce monthly payments, rather than decreased interest rates. The authors shed light on the role of maturity on purchasing behavior; however, limited and inconclusive empirical evidence exists about its implications for loan performance or pricing decisions.

The current paper attempts to fill this gap by estimating loan demand and loan performance jointly and highlighting the implications of maturity choice for screening out risky borrowers. First, I derive the econometric specifications for loan granting and repayment. I use these to estimate the elasticity of loan demand and probability of default with respect to both interest rate and loan maturity. Specifically, I test the null hypothesis that loan interest rate and maturity have no role in loan demand, whether borrowers are liquidity constrained or not. Second, I point out the role of a risk-based maturity setting in decreasing the information asymmetries on the loan market. In particular, I test the null hypothesis that maturity choice after risk-based pricing has no impact on loan default. Third, I show that the time of default is maturity-dependent and differs across borrowers in the different risk categories. The key contribution of this paper is that it shows that by reflecting the borrower's riskiness in the price of loan, both loan maturity terms and loan defaults increase. Specifically, liquidity constrained "high-risk" borrowers are offered high interest rates and most often choose long-term loans. This eventually increases their probability of default. Hence, a risk-based maturity setting does not necessarily improve the quality of consumer loans granted or alleviate the adverse selection present on the lending market.

This paper utilizes a unique dataset of rejected and accepted consumer loans from a Czech commercial bank (hereafter, the "Bank").[3] These include loans granted for the purchase of goods and services, loans granted for the modernization/reconstruction of housing and loans without a stated purpose. The unique dataset contains extensive information on borrower application characteristics, loan contract terms, and loan performance information of over 220 000 individuals who applied for a consumer loan between 2007 and 2013. From January 2012, the Bank has applied risk-based pricing, which is reviewed and developed periodically.

[3]The Bank does not wish to be explicitly identified. The anonymized data is available for replication.

2.The Lending Process

Altman (1980) defines the lending process as a sequence of activities involving two principal parties whose association spans from loan application to successful or unsuccessful loan repayment. Figure 1 illustrates the five key levels of the lending process.

Level 1

The individual enters the consumer loan market by submitting an application form for a loan.[4] The borrower discloses information about his/her socio-demographic characteristics such as age, marital status, education, etc. (application characteristics) and information related to the requested loan such as loan amount, loan maturity, etc. (loan term characteristics). The loan maturity is initially set by the applicant and is assumed to be driven by the long-term unemployment incidence of the region where the loan is requested.[5]

Level 2

The lender determines whether to grant the requested loan to the applicant. In order to assess the creditworthiness of their potential debtors, financial institutions use credit scoring techniques. The main purpose of these techniques is to estimate the probability that an applicant for credit will default by a given time in the future.[6] In its credit scoring model, the Bank estimates default probability using 3 types of credit scores: behavioral score (derived from the applicant's repayment history), application score (derived from the applicant's descriptive socio-demographic characteristics) and credit bureau score (derived from information about the applicant's existing and prior debt). Using these scores the bank assigns each applicant a risk band (four groups of "very

[4]On the consumer loan market, loan contract terms vary substantially across individual loan providers. Prior to loan application, the borrower has indicative information (for random loan amount and a minimum interest rate offer, each lender publishes a menu of maturities and annuity payments) about the lenders' offer from publicly available marketing materials. When entering the loan application process, the borrower uses this information to decide about his/her preferred loan maturity/amount given liquidity constraints – this requested loan amount and loan maturity can be considered the result of searching process.

[5]The change of loan maturity is subject to a new loan application.

[6]These are evaluated by analyzing a sample of customers who applied for loans in the past, where there is good information on subsequent loan performance history.

low-risk", "low-risk", "high-risk" and "very high-risk" borrowers). If the applicant's loan is pre-accepted (based on his/her aggregate credit score), the lender then assigns an interest rate for the requested loan maturity/amount. The interest rate is set primarily by the lender.[7] The lender offers loan contract terms that maximize its expected profit (taking into account the expected profit from an alternative investment of the loan amount). The interest rate is assumed to be driven by the applicant's risk margin, which is the price for the riskiness of the borrower and reflects the lender's risk aversion at the time of the loan request.

Level 3

Given the approved loan amount, interest rate and maturity the applicant has a chance to accept (open the account) or reject the loan contract conditions (no loan is originated). The borrower's decision is driven by its risk awareness and by the amount of monthly annuity payment (especially if the applicant is liquidity constrained). A loan is considered to be approved if it is approved by both the lender and the applicant. A loan is considered to be rejected if it is rejected by either the lender or the applicant.

Level 4

Given that the lender and the borrower agree on loan contract terms[8] and the borrower is granted the loan, the borrower starts repaying the principal and interest in the form of monthly annuity payments. The borrower can either follow the agreed repayment schedule, or renegotiate the loan contract terms (e.g. early repayment).[9]

---

[7]The assumption that loan maturity is primarily set by the borrower and the approved loan amount/interest rate is set primarily by the lender is made based on the Bank's best practice applied in the consumer loan market. It is also in line with the related literature. In Karlan and Zinman (2008) the lender identifies the loan price based on the borrower's pre-approved riskiness; and Attanasio et al. (2008) argue that credit-constrained borrowers' loan maturity is driven by their liquidity.

[8]The final loan contract terms are determined by the relative risk aversion across the borrower and the lender. The Bank's lending process is designed such that the lender reflects his/her risk aversion primarily through interest rate level and the borrower reflects its risk aversion primarily through maturity choice. In line with Adams et al. (2009) I assume that the competitive outcome is the contract that maximizes the borrowers' utility subject to lenders making nonnegative profits.

[9] Early repayment might be more likely for "high-risk" borrowers, since they can have then better credit after successful payments. However, early repayment is connected with additional borrowing costs in the form of prepayment penalty.

Figure 1.The lending process and data availability

Level 1

Borrower: Loan application (207 640 obs) l_i, x_i, u_i, t_i, n_i

Borrower: No loan application (no obs)

$$t_i = x_i\chi_1 + \chi_2 u_i + \chi_3 n_i + \chi_4 s_i + \varepsilon_{ti} \quad (3)$$

Level 2

Lender: Loan offer (114 944 obs) $l_i, x_i, t_i, u_i, m_i, b_i, r_i, n_i, s_i$

Lender: No loan offer (92 696 obs) $l_i, x_i, t_i, u_i, b_i, n_i, s_i$; $a_i = 0$

$$a_i = 1(x_i\delta_1 + \delta_2 m_i + \delta_3 u_i + \delta_4 b_i + \delta_5 s_i + \varepsilon_{bi} > 0) \quad (4)$$

$$r_i = x_i\beta_1 + \beta_2 m_i + \beta_3 n_i + \beta_4 s_i + \varepsilon_{ri} \quad (2)$$

Level 3

Borrower: Loan agreement (105 759 obs) $l_i, x_i, t_i, u_i, m_i, b_i, r_i, n_i, s_i$; $a_i = 1$

Borrower: No loan agreement (9 185 obs) $l_i, x_i, t_i, u_i, b_i, n_i, s_i$; $a_i = 0$

$$l_i = \alpha_1 x_i + \alpha_2 \hat{r}_i + \alpha_3 \hat{t}_i + \alpha_4 s_i + \alpha_5 \hat{\lambda}_i + \varepsilon_{gi} \quad (6)$$

Level 4

Renegotiation (2 090 obs) $l_i, x_i, t_i, u_i, m_i, b_i, r_i, n_i, s_i$; $a_i = 1$; $y_i = 1$

Regular payments (103 669 obs) $l_i, x_i, t_i, u_i, m_i, b_i, r_i, n_i, s_i$; $a_i = 1$; $y_i = 0$

Level 5

No Default (102 015 obs) $l_i, x_i, t_i, u_i, m_i, b_i, r_i, n_i, s_i$; $a_i = 1$; $d_i = 0$

Default (3 744 obs) $l_i, x_i, t_i, u_i, m_i, b_i, r_i, n_i, s_i$; $a_i = 1$; $d_i = 1$

$$\lambda(t_{di}, x_i) = \lambda_o(t_{di}).\exp(x_i\phi_1 + \phi_2 t_i s_i + \phi_3 t_i + \phi_4 s_i + \phi_5 r_i + \phi_6 n_i + \phi_7 \hat{\varepsilon}_{gi}) \quad (8)$$

Source: Author's illustration of lending process based on the description of the Bank.
Note: For loan request *i*the following information is available:
l_i - loan amount, r_i– loan interest rate, t_i- loan maturity, x_i - the borrower's application characteristics, n_i- the region in which is loan is requested,u_i- the region's long-term unemployment incidence, a_i– approved loan, b_i - number of debtors registered at the Czech Banking Credit Bureau at the time of loan request, m_i – risk margin, y_i- dummy for risk-based pricing, y_i – dummy for renegotiated loan, d_i – dummy for default, t_{di}– months till default. The individual equations of the econometric specification are described in Section 3 Methodology.

Level 5

The borrower either fully repays the loan or defaults. The borrower is considered to be in default if he/she is more than 90 days overdue with any payment connected with the loan.

3.Methodology

Overall, the main objective of this paper is to develop an econometric model that demonstrates the role of risk-based pricing and loan maturity on a consumer credit market with asymmetric information. I start by estimating the loan demand elasticity with respect to maturity and interest rate. Then I highlight the time dependency of default and examine maturity specific factors of loan performance.

The expected impact of selected variables and the predictions of the related literature are summarized in Table A1 in the Appendix.

3.1. Modeling Loan Demand

The loan demand estimation is complicated by the endogeneity of loan contract terms and sample selection (the nonrandom character) present in the consumer loan data. These can cause the parameter estimates to be biased. This section discusses how this paper deals with these two key issues in the loan demand estimation.

3.1.1. Loan Interest Rate

Interest rate endogeneity arises as lenders can change the loan price based on loan demand, and vice versa, the borrower can adjust his/her loan demand based on offered interest rates. In setting the price, the profit-maximizing lender aims to increase the interest rate, whereas the borrower aims to receive a loan at the lowest possible rate.

The literature deals with the endogeneity of interest rates in different ways. In Alessie, Weber, and Hochguertel (2005), the Italian usury law of 1997 (which limited interest rate charges) is used as an instrument for the identification of endogenous interest rate in loan demand estimation. The authors find evidence for the interest-rate elasticity of loan demand and argue that it is region specific. In Attanasio et al. (2008), the endogeneity of loan interest rate is addressed by exploiting data on the U.S. tax reform of 1986 (the change in interest deductibility affected the after-tax interest rate on the consumer loan market). Adams et al. (2009) identify loan demand on the car loan market by exploiting variation in list prices (i.e. catalogue car prices that different from negotiated prices) and variation in the level of down payments.

Similarly to Karlan and Zinman (2008), this paper captures the variation in the interest rate by information on the applicant's risk category. Applicants are classified into risk bands based on their estimated riskiness. These bands are then translated into risk margins taking into account the (conservative or aggressive) loan granting strategy of the lender. The higher the lender's risk aversion, the higher the risk margin and the final loan interest rate. I assume that the lender sets the final interest rate based on the loan‘s risk margin (the lender's willingness to accept the expected risk of the borrower). The interest margin has no effect on the loan amount, as the borrower is not aware of the lender's (frequently changing) loan granting strategy when setting its preferences aiming to smooth consumption.

3.1.2. Loan Maturity

Endogeneity of maturity is a further issue if the borrower cares primarily about monthly borrowing costs rather than the ultimate price of the loan. If the borrower is credit constrained and offered monthly payments (as result of maturity chosen by the borrower and interest rate set by the lender) that s/he cannot afford, s/he can either apply for a lower loan amount (which might decrease the interest rate) or prolong the maturity of the initially requested loan (accepting the initial interest rate). I assume that setting loan maturity is primarily the decision of the borrower, who aims to decrease the cost of lending by choosing shorter loans. S/he is willing to prolong the length of the loan only to such extent that the decreased monthly payments are acceptable for her expected future financial resources. The lender aims to prolong the loan maturity, as this is associated with higher interest income, while the higher riskiness of the borrower is implicitly reflected in the interest rate. It is questionable how successful the lender is in transferring the riskiness of borrower into the loan price or how significant is the adverse selection on the market. I discuss this issue in more detail in the next section.

The majority of studies neglect the effect of loan maturity on loan demand (Edelberg, 2006; Adams et al., 2009), and only limited empirical literature focuses on the role of loan maturity in borrowing behavior. In Attanasio et al. (2008) the endogeneity of loan maturity is addressed by using data on increased durability of cars (due to slower car depreciation, the maturity of loans is prolonged). Karlan and Zinman (2008) cooperate with the lender to generate exogenous variation in loan maturity. Specifically, randomly assigned “maturity suggestions” (loan offers for different maturities) are used to identify the elasticity of loan demand with respect to maturity. The randomized trial was conducted by a microfinance institution in South Africa.

To identify loan maturity in the loan demand equation, this paper utilizes data on the region’s unemployment duration. Specifically, I follow Jurajda and Munich (2002) and use the long-term unemployment incidence (hereafter, the “LTU incidence”) as a measure of unemployment duration. The LTU incidence is defined as the share of persons unemployed for 12 months or more in the total number of unemployed persons, expressed as a percentage.[11] There are two reasons to use LTU incidence as a measure of unemployment duration. First, as opposed to the LTU rate (the share of the number of long-term unemployed to the size of the labor force), the definition of LTU incidence is more transparent in transition countries where the concept of labor-force participation has been adopted gradually. Second, LTU incidence allows a researcher to capture the specifics of the business cycle (during recession it first declines driven by the increase in short-term unemployed workers, then it rises driven by the difficulty of the short-term unemployed to find employment) with the required regional granularity.

Several studies emphasize the role of unemployment in determining the duration of consumer loans. Navratil (1981) is the first to highlight that in periods of high unemployment rates, the short-term lending for auto loans is likely to increase, thus

[11]Source: Eurostat; http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tgs00053

decreasing loan maturity. A contrary finding is provided by the more recent paper by Chetty (2008), who shows that for the unemployed, the welfare gains of longer loans are much higher than the welfare gains of shorter loans. In particular, by prolonging the loan maturity, borrowers can decrease the monthly repayment amount and overcome financial difficulties during longer periods of unemployment. Attanasio et al. (2008) and Stephens (2008) argue that liquidity constraints determine the length of loans.

Motivated by the above studies, this paper utilizes the incidence of regional unemployment for the identification of loan maturity. Higher unemployment is expected to prolong consumer loans, as obtaining loans with longer maturity enables borrowers to take precautions against the risk of a long period of unemployment. On the other hand, the region's long-term unemployment incidence does not influence the number of loans requested, because the requested loan is primarily the result of the borrower's preferences about smoothing his/her consumption. If the borrower prefers to borrow some amount (rather than to save over a period of time for an expenditure), s/he is not discouraged from borrowing because s/he leaves in a region which has experienced an increase in its long-term unemployment incidence. What s/he primarily cares about in such a region are favorable loan contract terms.

3.1.3. Sample Selection

Sample selection arises for two reasons:

1) no information is available on those who did not wish to borrow;
2) information on rejected applicants is limited - loan contract terms are available only for those who were approved for a loan.

The related empirical literature acknowledges the difficulties in correcting for sample selection on the consumer loan market. Alessie et al. (2005) accept that the sample selection cannot be corrected, using Heckman's (1979) model, as the authors fail to find a variable that predicts loan approval but does not influence loan demand. They assume that a bank with a leading market position attracts applicants with good repayment behavior. Their solution is to estimate loan demand by controlling for the observable characteristics of the borrowers. Specifically, Alessie et al. (2005) correct for the composition effect connected to observable characteristics by p-score weighting the individual observations. Using data on auto loans, Attanasio et al. (2008) correct the sample selection in the loan demand equation through characteristics that have impact on buying a car, but do not necessarily affect loan amount (e.g. dummy for car ownership).

In line with the literature, this paper could not account for individuals who did not apply for a loan. I assume that the probability that an individual will apply for a loan has no endogenous effect on the probability of default. An individual can apply for a loan regardless of his/her expectation of the default probability it will be granted, as credit bureaus collect only information on borrowers who were eventually provided a loan. [12] If a potential borrower is rejected by the credit scoring evaluation, this is recorded in the credit bureau system for a maximum of 12 months. Thus, unless the customer has a bad loan repayment/default history connected with a previously provided loan, being rejected has no direct impact on the quality of his future loans after 12 months. In such cases, the probability of being accepted is equal in all institutions with no rejection history. The only cost implied by loan application is the time cost.

On the other hand, this paper does take into account the limited information on those who applied, but did not ultimately sign the loan contract. This includes both cases when the Bank rejects the applicant or when the applicant does not accept the loan contract terms offered by the Bank. To solve this problem of missing data on rejected loans, I follow Heckman (1979) and first estimate the selection equation on the whole sample of applicants. Similarly to Haas, Ferreira, and Taci (2010) and Bicakova, Prelcova, and Pasalicova (2010), the level of information-sharing about the borrowers' indebtedness is used to capture the variation in loan approval. Specifically, in this paper the exclusion restriction for the selection equation is the number of debtors monitored by the Czech Banking Credit Bureau. Over the past ten years, the credit bureaus have achieved substantial development both in the quality of information and the coverage of debt in the financial sector. This allows the use of information about a varying number of debtors to identify loan approval. The more positive information is available about

[12] The CBCB - Czech Banking Credit Bureau was established in 2002 for the purpose of operating the Client Information Bank Register (CIBR). It contains data on contractual (loan) relations between banks and their clients. http://www.cbcb.cz/

the debt level of a borrower, the more likely it is that the borrower is reliable and will maintain regular monthly loan repayments. At the same time, the borrower's decision about the requested loan amount is independent of developments in credit bureau information. His/her available credit history affects the decision of the prospective borrower to apply for a loan rather than the amount he/she applies for.

3.1.4. Model Specification

I specify the borrower's loan demand with respect to interest rate and maturity by the following econometric specification:

$$l_i = \log(L_i) = x_i\alpha_1 + \alpha_2 r_i + \alpha_3 t_i + \alpha_4 s_i + \varepsilon_{li}, \tag{1}$$

$$r_i = x_i\beta_1 + \beta_2 m_i + \beta_3 n_i + \beta_4 s_i + \varepsilon_{ri}, \tag{2}$$

$$t_i = x_i\chi_1 + \chi_2 u_i + \chi_3 n_i + \chi_4 s_i + \varepsilon_{ti}, \tag{3}$$

where for each loan application i = 1…N the following is known: L_i is the approved loan amount (takes logarithmic form as loans are nonnegative) , x_i is the vector of the information on application characteristics, behavioral and credit bureau score; r_i is the loan interest rate set primarily by the lender, t_i is the loan maturity set primarily by the borrower, m_i is the borrower's risk margin, u_i is the long-term unemployment incidence in the borrower's region, n_i is the region where the application *I* was submitted to the lender, s_i is a dummy for risk-based pricing introduced by the Bank in January 2012; and $\varepsilon_{ri}, \varepsilon_{ti}, \varepsilon_{li}$ are the unobserved error terms. Consequently, using loan repayment schedule with equal total payments, the lender charges the borrower a monthly annuity payment of $p_i(L_i) = (L_i * r_i)/(1-(1+r_i)^{-t_i})$.

To jointly account for both endogeneity and sample selection, I extend the sample selection model for endogeneous explanatory variables suggested by Wooldridge (2002) and estimate the structural equation of interest (1) together with the two equations describing the endogenous interest rate (2) and maturity (3), and the selection equation (4):

$$a_i = 1(x_i\delta_1 + \delta_2 m_i + \delta_3 u_i + \delta_4 b_i + \delta_5 s_i + \varepsilon_{bi} > 0) \tag{4}$$

where a_i is a binary variable indicating whether the loan is accepted $(a_i = 1)$ or rejected $(a_i = 0)$ either by the borrower or the lender, m_i is the borrower's risk margin, u_i is the long-term unemployment incidence in the borrower's region, b_i is the number of debtors registered at the Czech Banking Credit Bureau at the time of the loan request and ε_{bi} is the unobserved error term.

The following assumptions are made:

(a) $(x_i, n_i, m_i, u_i, b_i, s_i)$ is always observed, (l_i, r_i, t_i) is observed when $a_i = 1$;

(b) $(\varepsilon_{li}, \varepsilon_{bi})$ is independent of $(x_i, n_i, m_i, u_i, b_i, s_i)$;

(c) ε_b ~ Normal (0, 1);

(d) $\mathrm{E}(\varepsilon_{li} \mid \varepsilon_{bi}) = \gamma_4 \varepsilon_{bi}$;

(e) $\mathrm{E}(z_1' \varepsilon_{ri}) = 0$ (where $z_1\beta = x_i\beta_1 + \beta_2 m_i + \beta_3 n_i + \beta_4 s_i$) and $\beta_2 \neq 0$;

$\mathrm{E}(z_2' \varepsilon_{ti}) = 0$ (where $z_2\chi = x_i\chi_1 + \chi_2 u_i + \chi_3 n_i + \chi_4 s_i$) and $\chi_2 \neq 0$.

Assumption (a) emphasizes the nonrandom nature of the sample. The exogeneity of application characteristics x_i and the two exogenous variables m_i, u_i is formalized by assumption (b). Assumption (c) states that the error term of the selection equation follows standard normal distribution. Linearity in the regression of ε_{li} on ε_{bi} is required by assumption (d). Lastly, assumption (e) results from the endogeneity of loan contract terms in the loan demand equation (1). It states that (i) the error terms

$\varepsilon_{ri}, \varepsilon_{ti}$ have zero mean and are uncorrelated with the right-hand-side variables, and (ii), (β_2, χ_2) are nonzero, requiring that at least two exogenous variables (m_i, u_i) do not appear in the loan demand equation (the order condition). Under this assumption the parameters β_2 and χ_2 are identified.

The derived estimating equation has the following form:

$$l_i = \alpha_1 x_i + \alpha_2 r_i + \alpha_3 t_i + \alpha_4 s_i + g(x_i, n_i, m_i, u_i, b_i, s_i, a_i) + \varepsilon_{gi}, \tag{5}$$

where $g(x_i, m_i, n_i, u_i, b_i, s_i, a_i) \equiv \mathrm{E}(\varepsilon_{li} \mid x_i, n_i, m_i, u_i, b_i, s_i, a_i)$ and $\varepsilon_{gi} \equiv \varepsilon_{li} - \mathrm{E}(\varepsilon_{li} \mid x_i, n_i, m_i, u_i, b_i, s_i, a_i)$. By definition the error term is uncorrelated with the exogenous variables: $\mathrm{E}(\varepsilon_{gi} \mid x_i, n_i, m_i, u_i, b_i, s_i, a_i) = 0$. Equation (5) is estimated by 3SLS on the sample of accepted loan applications $(a_i = 1)$ using the exogenous variables and the estimated inverse Mills ratio, where
$\mathrm{E}(\varepsilon_{li} \mid x_i, m_i, u_i, b_i, s_i, a_i = 1) = \alpha_4 \lambda(x_i \delta_1 + \delta_2 m_i + \delta_3 u_i + \delta_4 b_i + \delta_5 s_i)$.

Specifically, the estimation is performed in two steps. First, using all observations the selection equation is estimated by probit and the estimated inverse Mills ratio $\hat{\lambda}_i$ is obtained. Second, using the subsample for which both (r_i, t_i) are observed, the equation

$$l_i = \alpha_1 x_i + \alpha_2 \hat{r}_i + \alpha_3 \hat{t}_i + \alpha_4 s_i + \alpha_5 \hat{\lambda}_i + \varepsilon_{gi} \tag{6}$$

is estimated by 3SLS, using the exogenous variables $(m_i, u_i, b_i, \hat{\lambda}_i)$. [13] In particular, I test the null hypothesis that interest rate and loan maturity have no effect on the approved loan amount: $(H_0 : \alpha_2 = 0)$ and $(H_0 : \alpha_3 = 0)$. The sensitivity of loan demand to loan contract terms is estimated both on the pooled sample (including all

[13] The sample selection correction is also present in the equations for interest rate (2) and maturity (3) as these are estimated on approved loans.

observations) and on the subsample of low-income borrowers (liquidity constrained borrowers[14] whose net monthly income at the time of loan application is below the sample's median net monthly income). Finally, the null hypothesis of no selection bias ($\mathrm{H}_0 : \alpha_5 = 0$) is tested by exploiting the 3SLS t statistic for $\hat{\alpha}_5$; and the null hypothesis of no endogeneity is tested by estimating the structural model (1) that includes the residuals from the two equations describing the endogenous interest rate (2) and maturity (3).

3.2. Modeling Default Probability

The goal of this section is to propose an econometric model that uses demand estimates for predicting default probability. The model should reflect how the different loan contract terms influencing consumer behavior affect the loan performance. Specifically, I focus on the time dependency of default (the length of time the borrower avoided default has impact on the probability of default) and test for the significance of asymmetric information hidden in the maturity choice.[15]

To do this, I take advantage of the semi-parametric proportional hazard model, which relates the individual covariates and the time of event (or failure, as I refer to default) occurrence in multiplicate form. If $\lambda(t_{di}, x_i)$ is the probability that an individual defaults at time t_{di} (conditional on making regular payments till default), x_i are application characteristics, the relationship between the distribution of failure times and the vector of application characteristics can be expressed by the semi-parametric proportional hazard model developed by Cox (1972) as

$$\lambda(t_{di}, x_i) = \lambda_o(t_{di}).\exp(x_i\phi_1 + \phi_2 s_i y_i + \phi_3 t_i + \phi_4 s_i + \phi_5 r_i + \phi_6 n_i + \phi_7 y_i), \qquad (7)$$

[14]Borrowers with liquidity constrains cannot be easily identified. This paper utilizes the approach of Attanasio et al. (2008) who assume that low-income borrowers are liquidity constrained borrowers.

[15]Flannery (1986), Diamond (1991) and Berger, Espinosa-Vega, Frame, and Miller (2005) were the first to suggest that the size of asymmetric information between lenders and borrowers can significantly affect the choice of loan maturity. They focused on commercial and industrial loans.

where s_i is a dummy variable taking the value 1 if the application was evaluated using risk-based pricing, and y_i is a dummy variable taking the value 1 if the application renegotiated ex post. The advantage of proportional hazard models is that whereas parametric models use information over the whole time horizon (distributional assumption for baseline hazard $\lambda_0(t_{di})$; estimation of the cumulative hazard), semi-parametric models use only the information at failure times (no distributional assumption for baseline hazard; estimation of the direct hazard).

The incomplete information on the occurrence of events during the observation period belongs among the specifics of duration time estimation. As the information about the loan performance after the end of the observation period is missing, I deal with right censored data. There are three possibilities of the event status: the event occurred by t_{di}^* (duration time), the event did not occur by the end of observation period or the event did not occur before loan completion (t_c). For each individual one observes t_{di}, where $t_{di} = \min(t_{di}^*, t_c)$.

Loan amount and default jointly are modelled jointly:

$$\lambda(t_{di}, x_i) = \lambda_o(t_{di}).\exp(x_i\phi_1 + \phi_2 t_i s_i + \phi_3 t_i + \phi_4 s_i + \phi_5 r_i + \phi_6 n_i + \phi_7 \hat{\varepsilon}_{gi}), \qquad (8)$$

I test the null hypothesis that loan maturity choice after risk-based pricing has no impact on the loan default; formally I test $H_0 : \phi_2 = 0$. Similarly to Adams et at. (2009), the identification is through the two-stage control function approach – to estimate the loan default I use the estimated residual $\hat{\varepsilon}_{gi}$ from loan demand estimation is used as control variable. The main goal is to identify the borrowers' private information at the time of loan application that affects both loan amount and loan default. The models for loan demand (6) and the default probability (8) are also estimated for short-, medium- and long- term loans[16] and across borrowers in the different risk categories.

[16] Glennon and Nigro (2005) argue that the determinants of default are maturity-specific.

4.Data

4.1. Data Description

The data sample consists of the consumer loan information of over 220 000 individuals. The dataset includes application characteristics (e.g. age, marital status, education, etc.), loan contract information (e.g. interest rate, loan maturity, loan amount, etc.) and performance indicators (e.g. date of default, monthly outstanding balance, past due, etc.). The consumers requested the loans between 2007 and 2013[17], where the last performance observation is from April 2013. Table A2 in the Appendix summarizes the list of available information on consumer loans. Table A3 in the Appendix, reporting the basic descriptive statistics, suggests that an average borrower is 40 years old, receives a net monthly income above 17 000 CZK and has been employed for more than 5 years.

In order to measure the performance of the loans, monthly data on repayment status is used. For each loan, one piece of the following information is available: the number of the months till default, the number of months till on-time repayment or the number of months till the end of the data observation interval (April 2013). That is, each loan has its survival time: either time to default or time to non-default (being repaid or censored data). This enables a more precise estimation of default, as the number of successful payments till default is also taken into account.

When monitored on the 30th of April 2013, 3.6 % of those who had obtained a loan had defaulted and the rest of the borrowers performed well. Although there are several different definitions of "defaulted" loans, the one of the Basel Committee on Banking Supervision (2004) is applied: a loan is in default if the borrower is more than 90 days overdue with any payment connected with the loan.

Rejected loans comprise 48.9 % of the total number of consumer loans. These include those applications that were either rejected by the lender (due to application characteristics or credit history) or the borrower (due to unfavorable loan terms offered by the lender). Figure 1 illustrates the number of rejected loans by lender (92 696 loans)

[17] The dataset differentiates between the date of loan request and loan opening. Year dummies are created based on the loan request date at which the Bank decided to accept or reject the applicant.

and by the borrower (9 185 loans). Rejection by the borrower is not identified separately, as 90% of the applicants are rejected based on the information gained from credit bureau.

In addition to information on interest rate, data on risk margin is also limited. Risk margin is observed only after the risk-based pricing is implemented (January 2012). I solve this issue by multiple imputation (similar to Adams et al. 2009). For each approved loan application prior to January 2012, the missing risk margin is replaced with predicted values from a regression analysis of the complete data. The development of risk margin over the observation period is summarized in Table A4 in the Appendix. The sample statistics indicate that there is a gradual increase in the risk margin and lenders requested the highest risk margin during 2012.

The consumer loan data utilized in this paper is application-specific – for one application I observe only one outcome (loan contract terms, loan performance) and the change of loan contract terms is subject to a new unidentified loan application. Renegotiated loans were first signed with initial loan contract terms, and then during the loan repayment period the loan contract terms were renegotiated. As the information on the renegotiated interest rates is not available, renegotiated loans cannot be used to study the incidence of change in loan maturity before and after the introduction of risk-based pricing.

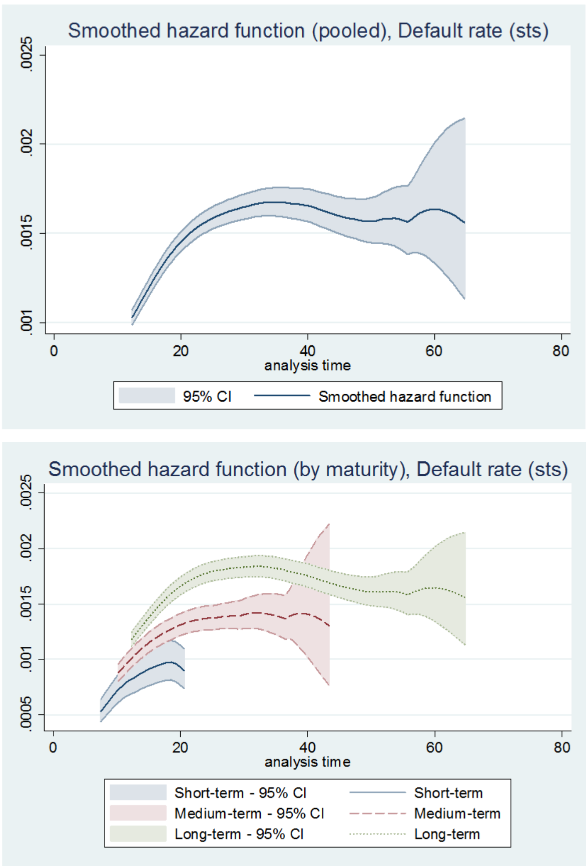
Figure 1 summarizes how the data availability differs over the individual levels of lending process.

4.2. Data Analysis

Although there are several estimation techniques of the survival functions, nonparametric methods are very useful for descriptive purposes in the first place. They illustrate the shape of the unconditional hazard and survival functions before introducing the covariates into the model. Specifically, the survivor and the hazard functions are easily interpretable and effective in describing the duration dependence.

Figure A1 in the Appendix depicts the cumulative hazard function (with 95% confidence intervals) estimated by the Nelson-Aalen method. It suggests that at the end of the consumer loan observation period, almost 90% of the sample remained without default. Figure 2 plots the estimated hazard rate (with 95% confidence intervals), which expresses the instantaneous probability of default conditional on making regular payments until a particular month during the time under analysis. According to the smoothed hazard function that treats all consumer loans equally and does not distinguish between maturity or risk bands (‘pooled’), defaults are most likely to occur around the 30th month from the date of loan provision. On the other hand, the smoothed hazard function by maturity suggests that the default is not only time-dependent, but also maturity dependent.

Figure 2. Smoothed hazard function pooled and by maturity



Source: Author’s computations, 2007-2013. *Note:* (1) The figure on the left depicts pooled data, i.e. treats all consumer loans equally and does not distinguish between maturity or risk bands. (2) The figure on the right depicts smoothed hazard functions for short term loans with maturity up to 2 years, medium term loans with maturity between 2 and 5 years and long term loans with maturity more than 5 years.

Table 1 presents the preliminary sample statistics of average maturity (Panel A) and average default rate (Panel B) before/after the introduction of risk-based pricing. Due to the limited observation period after the introduction of risk-based pricing (January 2012), the before/after periods are represented only by one year (2011/2012). After the introduction of risk-based pricing, borrowers in all risk bands increase their average loan duration, but the "very high-risk" group remains almost unchanged. The statistics from Table 1 (Panel A) are in line with Karlan and Zinman (2008), who show that by longer maturity the borrower can lower the amount of monthly payments and, hence, afford higher loan amount. Panel B summarizes the observed average default rate for risk bands and loans with different maturities. One year before the introduction of risk-based pricing, "very high-risk" borrowers with medium-term loans (2-year to 5-year) have the highest incidence of default. One year after the introduction of risk-based pricing, borrowers with long-term loans (more than 5-year) default the most frequently.

Hence, the main focus of this paper is whether banks applying risk-based pricing are able to decrease the adverse selection (i.e. borrowers with high probability of default increase their debt amount) for liquidity constrained borrowers who are more sensitive to maturity changes (relative to interest rate changes).

Table 1. Sample statistics on before/after risk-based pricing

Panel A - Average maturity

	Average loan maturity		Number of observations	
Risk band	**Before risk-based pricing**	**After risk-based pricing**	**Before risk-based pricing**	**After risk-based pricing**
Very low-risk	4,4	4,6	8 667	9 902
Low-risk	4,5	4,7	6 443	6 624
High-risk	4,1	4,2	1 450	1 580
Very high-risk	3,5	3,5	551	454
Total	**4,4**	**4,5**	**17 111**	**18 560**

(continued on next page)

Table 1. Sample statistics on before/after risk-based pricing

Panel B - Average default rate

Risk band	Before risk-based pricing	After risk-based pricing
Very low-risk	**0.6%**	**0.1%**
<2Y	0.1%	0.0%
2Y-5Y	0.4%	0.1%
>5Y	1.0%	0.2%
Low-risk	**1.8%**	**0.4%**
<2Y	1.0%	0.3%
2Y-5Y	1.8%	0.4%
>5Y	2.1%	0.4%
High-risk	**3.9%**	**1.3%**
<2Y	2.9%	1.0%
2Y-5Y	4.0%	1.3%
>5Y	4.2%	1.3%
Very high-risk	**5.3%**	**4.0%**
<2Y	3.1%	2.9%
2Y-5Y	6.2%	3.3%
>5Y	5.2%	6.4%
Total	**1.5%**	**0.4%**

Source: Author's calculations using the sample of Czech consumer loans. *Note*: (1) The Bank classifies borrowers into risk bands based on the estimated riskiness. (2) Before risk-based pricing is represented by year 2011, and after risk-based pricing is represented by year 2012.

5.Results

This section starts with the estimation of the loan demand model that accounts for both the presence of sample selection and the issue of endogeneity. Then I discuss the estimates of default probability derived from the Cox proportional hazard model and highlight the implications of risk-based pricing on the quality of granted loans, i.e. on the probability of default. Both loan demand and loan performance are examined with respect to loan contract terms and with respect to the borrower's application characteristics. Finally, I illustrate the maturity-dependent default probability for borrowers in the different risk categories.

5.1. The Elasticity of Loan Demand to Interest Rate and Maturity

First, I correct for the nonrandom feature of the data, by estimating the probability of loan approval based on selection equation (6). [18]The nonrandom issue of the sample arises as there is no information available on those individuals who do not apply for a loan and limited information on those who apply but do not sign the loan contract. Therefore, I estimate the Heckman (1979) selection model that corrects for this type of incomplete information. The number of individuals monitored in the Czech Banking Credit Bureau at the time of loan application is used as an exclusion restriction.

Second, using the estimated inverse Mills ratio from the Heckman (1979) model I estimate the loan demand equation (1) with the two equations describing the endogenous interest rate (2) and loan maturity (3). The three equations are estimated using 3SLS, where the two exclusion restrictions are the borrower's risk margin and the average long-term unemployment incidence in the borrower's region.

I reject the null hypothesis that loan interest rate and maturity have no role in loan demand (Table 2). Consistent with Alessie et al. (2005), the results suggest that increasing interest rates discourage individuals from borrowing (loan demand decreases), whereas with longer maturity the loan amount increases (similar to Attanasio et al.'s 2008 study).

The test results suggest that both the null hypothesis of no-sample-selection and the null hypothesis of no-endogeneity can be rejected at 1%. First, I use the t statistic on the inverse Mills ratio (variable INVMILLS) as a test for the presence of sample selection $\mathrm{H}_0 : \alpha_4 = 0$. The z-value of 15.6 is a strong evidence against the null hypothesis of no-sample-selection (Table 2, Column 2). Second, I test the endogeneity of interest rate and maturity jointly. Specifically, for both endogenous variables I obtain the reduced form residuals, and then I test the joint significance of these residuals in the

[18]I follow the variable (non)categorization of the Bank. In all models the variables are used in the same manner as they enter the Bank's credit scoring model. The individual estimates refer to indicated changes in the dependent variable due to a change in the particular application characteristic compared to its reference group.

structural equation using an F test. The F (2, 105723) being equal to 188.8 is well above the 1% critical value in the F distribution, so I reject the null hypothesis that interest rate and loan maturity have no effect on the approved loan amount. In addition, I reject the null hypothesis that risk margin has no effect on the loan interest rate (at 1% significance level) or that LTU incidence has no effect on the loan maturity (at 5% significance level). One percentage point increase in the risk margin leads to a 0.3 percentage point increase in the interest rate (similar to Karlan and Zinman 2008's findings); and one year increase in the region's long-term unemployment leads to a 0.4 year increase in the loan maturity rate (similar to Chetty 2008's findings).

In Table 2, I also compare the interest rate and maturity elasticity of loan demand for the pooled sample (Column 2) and for the subsample of low-income borrowers (Column 4). The results suggest that the loan amount of a low-income borrower increases with longer maturity (one month increase in the loan maturity results in a 2.1% increase in the loan amount), while the interest rate has statistically no significant effect for these borrowers. The increasing importance of loan maturity for low-income borrowers is in line with Karlan and Zinman's 2008 findings. However, this paper goes further and uses the maturity elastic demand estimates to see the probability of default they imply (see the details in the next section).

Table A5 in the Appendix summarizes how the borrower's application characteristics affect loan demand. The parameter estimates have the expected signs. If focusing on low-income borrowers, the results suggest that women, pensioners, students and borrowers who rent housing borrow less. Interestingly, married borrowers, with university education and employed in banking/insurance company have the higher loan demand. The results are qualitatively comparable to loan demand determinants derived by Attanasio et al. (2008) and Adams et al. (2009).

Table 2. Estimation results of loan demand and default probability

Dependent variable	Loan demand				Default probability			
	Pooled sample		Low-income subsample		Pooled sample		Low-income subsample	
	Coef.	*St.error*	Coef.	*St.error*	Haz. ratio	*St.error*	Haz. ratio	*St.error*
Interest rate	-0.035***	*0,004*	0.003	*0,006*	1.172***	*0.010*	1.084***	*0.017*
Approved maturity	0.015***	*0,001*	0.021***	*0,001*	1.004***	*0.001*	1.005***	*0.001*
Credit bureau score	0.001***	*0,000*	-0.001***	*0.001*	0.999***	*0.001*	0.999***	*0.000*
Behavioral score	0.001***	*0,000*	0.001***	*0.001*	0.998***	*0.001*	0.998***	*0.000*
Inverse Mills ratio	0.326***	*0,028*	0.013	*0,037*				
Risk-based pricing	0.060***	*0,005*	0.020**	*0,008*	0.442***	*0.127*	0.328**	*0.164*
Renegotiated loan					6.020***	*0.294*	6.671***	*0.415*
Approved maturity *Risk-based pricing					1.009**	*0.004*	1.013	*0.008*
Loan demand residual					0.819***	*0.024*	0.793***	*0.032*
R^2	0.5093		0.4639					
N	105 759		46 598		105 759		46 598	
Log likelihood					-38 221		-20 223	
Prob> chi2					0.000		0.000	
Loglikelihood ratio (LR) chi2					4 858		2 639	

Source: Author's computations, 2007-2013. *Note:* (1) For loan demand estimation the logarithmic form of approved loan amount is used. (2) INVMILLS denotes the Inverse Mills ratio calculated after estimating equation, (3), AAMOUNT_RES denotes the estimated residual from the loan demand equation. (4) Robust standard errors are used for statistical inferences. (5) Estimation results presented only for variables that were statistically significant at least in one model. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

5.2. The Impact of Risk-Based Pricing on Loan Performance

The borrower's probability of default is estimated on the loan contract term and the borrower's application characteristics using the Cox proportional hazard model. In addition to the loan and application characteristics, the estimated residual from the loan demand equation is included in the model as control variable. Table 2 summarizes the estimation results (hazard rates) for the pooled sample (Column 6) and for the subsample of low-income borrowers (Column 8). The Cox partial likelihood model provides a semi-parametric specification for the relationship between hazard rates and the application characteristics.[19] Column 6 and Column 8 in Table 2 quantify the hazard rate, $\exp(\beta)$, for the application characteristics as a percentage of the hazard rate for

[19]The reference group for the application factor variables is always the one with the lowest coding. For the coding of variables refer to Table A2 on the Appendix.

their reference groups. The results provide evidence of the effect of risk-based pricing (variable *RBPRICING*) introduced by the Bank over the observation time (in January 2012). As the elasticity of loan demand with respect to maturity has been shown to be statistically significant, I introduce an interaction term of risk-based pricing with approved maturity (*RBPRICING*AMATURITY*). The hazard ratio on this interaction term suggests that the null hypothesis that maturity choice after risk-based pricing has no impact on loan default can be rejected. Given risk-based pricing, prolonging loan maturity increases the probability of default for the pooled sample of borrowers by 1.3 % (derived from coefficients in Table 2 Column 6) and for the sub-sample of low-income borrowers by 1.2%.[20] The time-dependence in default described below suggests that the negative impact of long-term loans is likely to increase as the observation period is extended (loan performance after introducing risk-based pricing is examined only over the fourteen month period between January 2012 and April 2013). In other words, differentiating between borrowers solely through different interest rates causes borrowers to choose either to reduce the loan amount or to prolong maturity to compensate the lender for their riskiness. The latter then leads to higher default probability for both the liquidity constrained and liquidity unconstrained borrowers. Thus, banks seeking to mitigate adverse selection by developing risk-based pricing should also test for the increasing riskiness of the borrower pool with respect to loan duration. These results complement the findings of Adams et al. (2009), who quantify the positive impact of risk-based pricing on loan performance without controlling for the endogeneity of loan maturity.

The effect of individual application characteristics on default probability presented in Table A5 in the Appendix is in line with the expectations. For instance, consistent with Kocenda and Vojtek (2009), the hazard ratio for low-income borrowers with university education is only 56% of the hazard rate for those who have secondary technical education. The longer survival time without default increases with longer period of employment as in Bicakova (2007). Borrowers who own property are

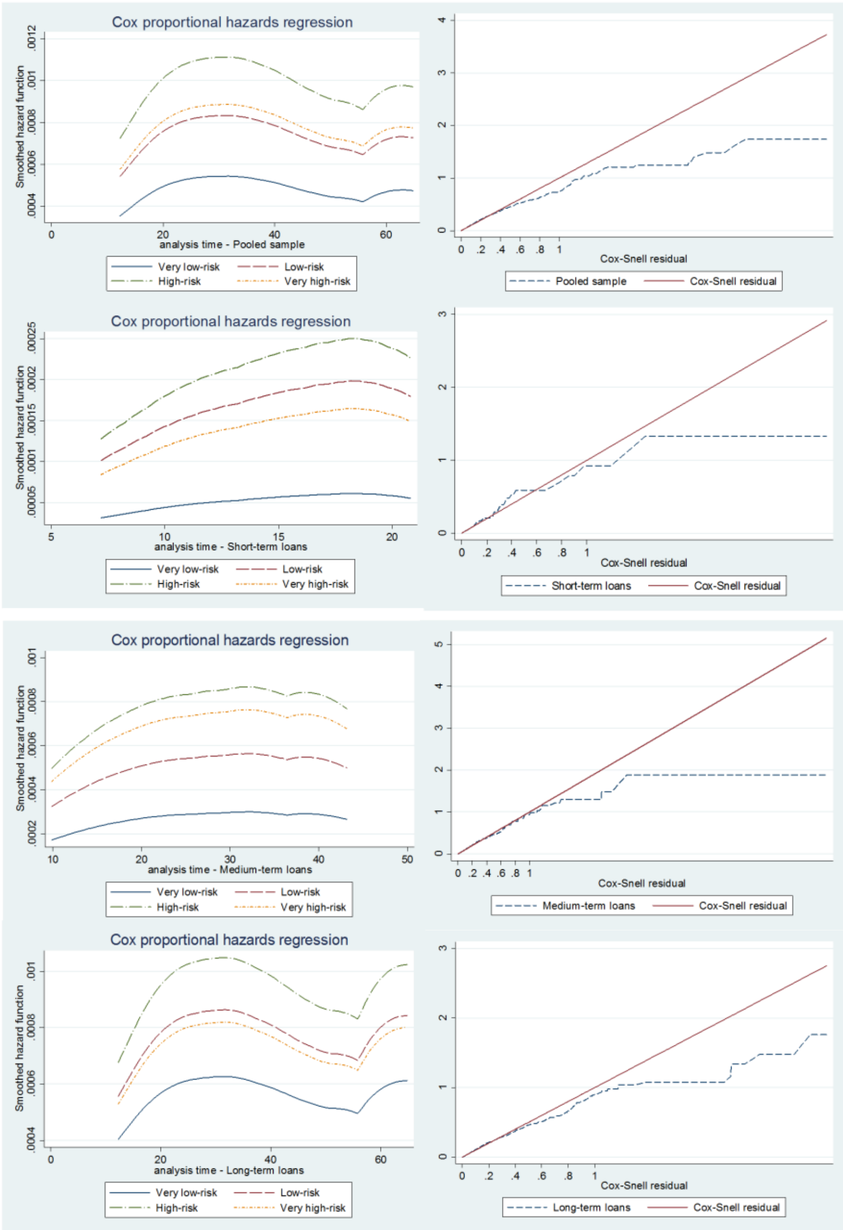
[20]As a robustness check the simple probit of loan default was performed on all observations (with default occurring within 24 months after loan origination). This alternative specification yields similar conclusions as those derived from the Cox proportional hazard model.

associated with a 43% lower risk of default than those who do not own property. These results are in line with the predictions of Einav et al. (2012).

Figure A2 in the Appendix plots the fitted Cox proportional hazards regression by loan maturity. It depicts the estimated default probability for the pooled sample and for the subsamples with different maturity: borrowers of short-term loans (maturity up to two years) are the most likely to default after the 18th month of granting; medium term loans (maturity between two and five years) are the most likely to default at the 30th month, and long term loans (more than five years maturity) default most frequently around the 34th month. Comparing the pooled proportional hazards and the proportional hazards by maturity, all achieve their peak before the end of the third year. These results suggest that the timing of default is maturity-specific. While Glennon and Nigro (2005) find that between 1983 and 1998 the default most frequently occurs before the end of the second year after loan origination, Figure A2 shows that between 2007 and 2013 the default occurrence peaks around the third year. This can be explained by the overall prolongation of consumer loans.

To see the how significant the time-dependent default is across borrowers in the different risk categories, I also plot the proportional hazard by maturity and by risk band (Figure 3). The overall model fit of the individual hazard regressions is assessed by computing the Cox-Snell residuals. If the model is correct, the real cumulative hazard function based on the covariate vector has an exponential distribution and a hazard rate of one. The default variation plotted in Figure 3 is the most significant for long-term loans. Comparing the dashed line with Cox-Snell residuals in Figure 3, it can be concluded that the maturity-specific models fit the data equally as well as the model for the pooled sample. The results suggest that in addition to risk-based consumer loan pricing, maturity-based credit scoring is also inevitable.

Figure 3. Cox proportional hazards regression pooled and by maturity/by risk bands



Source: Author's computations, 2007-2013. *Note:* The model fit is evaluated by the comparison of the Cox cumulative hazard to the Cox Snell residual.

6.Conclusion

Driven by the sharp increase in consumer loan demand, the role of credit scoring methods in assessing a borrowers' creditworthiness is becoming more and more important. Thanks to the wide range of credit history collected by credit bureaus, lenders can screen out risky borrowers in their credit scoring models, not only based on application characteristics, but on behavioral and credit history information. However, the ultimate effect of different loan contact terms on loan demand and loan performance has not yet been examined in the process of loan provision.

The aim of this paper is to present empirical evidence about whether a risk-based maturity setting improves the quality of granted consumer loans and alleviates the adverse selection present on the lending market. Taking advantage of a sample of both accepted and rejected consumer loans from a Czech commercial bank, this paper is the first to point out the importance of maturity in loan demand and loan performance.

This study contributes to the existing literature on consumer loan markets in several ways. First, it shows that low-income borrowers can be credit constrained and thus have limited access to credit at market interest rates. Empirical evidence suggests that loan demand for low-income borrowers is more sensitive to available cash and loan maturity changes than to interest rate changes. This is consistent with the assumption that borrowers with liquidity constraints are likely to prolong the maturity of their loans in order to borrow the desired loan amount. Second, by reflecting the borrower's riskiness in the interest rate, lenders discourage risky borrowers from obtaining short-term loans. This then leads to higher default probability for both liquidity constrained and liquidity unconstrained borrowers. The finding is consistent with the theoretical prediction that reduced asymmetric information encourages "high-risk" borrowers to either demand lower loan amounts or to prolong their loan maturity to compensate the lender for their riskiness. Therefore, banks seeking to mitigate adverse selection by developing risk-based pricing should also test the increasing riskiness of borrower pool due to the sensitivity to loan duration. Finally, this paper provides evidence that the time of default is maturity-dependent and differs across borrowers in the different risk categories. Hazard models that differentiate between loan maturities and risk bands have an equally good model fit as one that treats all consumer loans as pooled and does not distinguish between these two factors. These results further advocate the necessity of maturity-based credit scoring.

References

Adams, W., Einav, L., and Levin, J. (2009). Liquidity Constraints and Imperfect Information in Subprime Lending. *American Economic Review, 99*, 49-84.

Alessie, R., Weber, G., and Hochguertel, S. (2005). Consumer Credit: Evidence from Italian Micro Data. *Journal of the European Economic Association, 1*(3), 144-178.

Altman, E.I. (1980). Commercial Bank Lending: Process, Credit Scoring, and Costs of Errors in Lending. *Journal of Financial and Quantitative Analysis, 15(*4), 813-832.

Attanasio, O.P, Goldberg, P.K., and Kyriazidou, E. (2008). Credit Constraints in the Market for Consumer Durables: Evidence from Micro Data on Car Loans. *International Economic Review, 49*(2), 401-436.

Basel Committee on Banking Supervision (2004). International Convergence of Capital Measurement and Capital Standards. *Revised Framework, Bank for International Settlements.*

Berger, A.N., Espinosa-Vega, M.A., Frame, W.S., and Miller, N. H. (2005). Debt Maturity, Risk, and Asymmetric Information. *Journal of Finance*, *60*(6), 2895-2923.

Bicakova, A. (2007). Does the Good Matter? Evidence on Moral Hazard and Adverse Selection from Consumer Credit Market. *European University Institute Working Papers.ECO No. 2007/2.*

Bicakova, A. Prelcova, Z. and Pasalicova, R. (2010). Who Borrows and Who May Not Repay. *Czech National Bank, Working Papers 10*.

Chetty, R. (2008). Moral Hazard versus Liquidity and Optimal Unemployment Insurance. *Journal of Political Economy, 116* (2), 173-234.

Cox, D.R. (1972). Regression Models and Life-Tables (with discussion). *Journal of Royal Statistical Society, 34*(2), 248-275.

Diamond, D. W. (1991). Debt Maturity Structure and Liquidity Risk. *Quarterly Journal of Economics, 106*, 709-738.

Edelberg, W. (2006). Risk-Based Pricing of Interest Rates for Consumer Loans. *Journal of Monatery Economics*, *53*, 2283-2298.

Einav, L., Jenkins, M. and Levin, J. (2012). Contract Pricing in Consumer Credit Markets. *Econometrica, 80*(4), 1387-1432.

Flannery, M. J. (1986). Asymmetric Information and Risky Debt Maturity Choice. *Journal of Finance, 41*, 19-37.

Glennon, D. and Nigro, P. (2005). An Analysis of SBA Loan Defaults by Maturity Structure. *Journal of Financial Services Research, 28*, 77-111.

Haas, R. D., Ferreira, D. and Taci, A. (2010). What Determines the Composition of Banks' Loan Portfolios? Evidence from transition countries. *Journal of Banking & Finance, 34*, 388–398.

Heckman, J.J. (1979). Sample Selection Bias as Specification Error. *Econometrica, 47*, 153 – 161.

Jaffee, D. and Stiglitz, J.E. (1990). Credit Rationing' In Handbook of Monetary Economics, Volume 2, ed. B. M. Friedman and F. H. Hahn, 837–88. Amsterdam: Elsevier.

Jurajda, S. and Munich, D. (2002). Understanding Czech Long-Term Unemployment. *William Davidson Working Paper 498.*

Karlan, D., and Zinman, J. (2008). Credit Elasticities in Less Developed Countries: Implications for Microfinance. *American Economic Review, 98*, 1040–1068.

Kocenda, E. and Vojtek, M. (2009). Default Predictors and Credit Scoring Models for Retail Banking. *CESifo Working Paper Series*2862, CESifo Group Munich.

Navratil, F.J. (1981). An Aggregate Model of the Credit Union Industry. *The Journal of Finance, 36* (2), 539-549.

Stephens, M. (2008). The Consumption Response to Predictable Changes in Discretionary Income: Evidence from the Repayment of Vehicle Loans. *The Review of Economics and Statistics*, *90*(2), 241–252.

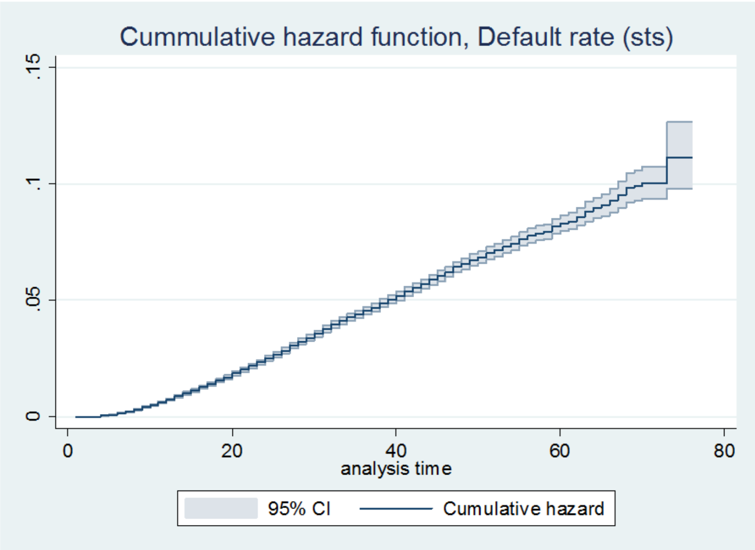
Stiglitz, J.E. and Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information. *American Economic Review, 71*, 393-409.

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data.The MIT Press.*Cambridge, MA*.

Appendix

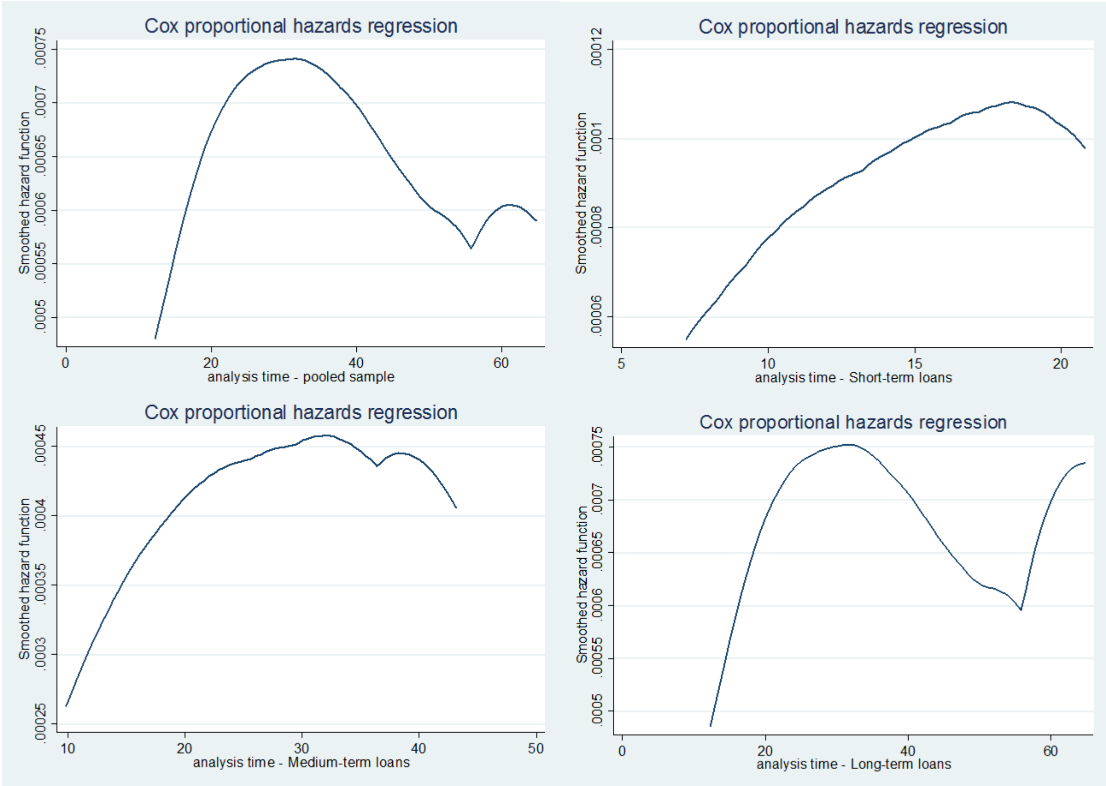
Figures

Figure A1. Nelson-Aalen estimator of the cumulative hazard function



Source: Author's computations, 2007-2013.

Figure A2. Cox proportional hazards regression pooled and by maturity



Source: Author's computations, 2007-2013. The figure on the upper left corner depicts Cox proportional hazards for pooled data, i.e. treats all consumer loans equally and does not distinguish between maturity. The other three figures depict the Cox proportional hazards for short term loans with up to 2 years, medium term loans with maturity between 2 and 5 years and long term loans with maturity more than 5 years.

Tables

Table A1. Expected relationship between selected dependent and independent variables

Dependent variable	Independent variable	Expected relationship	Literature
Loan approval	Car ownership	+	Alessie et al. (2005)
	Use of credit bureau information	+	Bicakova et al. (2010)
Loan demand	Interest rate	-	Alessie et al. (2005)
	Maturity	+	Attanasio et al. (2008)
Loan interest rate	Risk category	+	Karlan and Zinman (2008)
	Tax reform on phase out of interest deductibility	+	Attanasio et al. (2008)
	Usury law on max interest rate level	-	Alessie et al. (2005)
Loan maturity	Unemployment rate	+	Chetty et al. (2008)
	Durability of cars	+	Attanasio et al. (2008)
Default probability	Interest rate	+	Adams et al. (2009)
	Maturity	+	Adams et al. (2009)

Source: Author’s literature review.

Table A2. The list of personal loan information (Panel A)

Variable description	Variable name in dataset	Encoding
Application characteristics		
Age (in months)	***AGE***	***continuous***
Female	***FEMALE***	***dummy***
Marital status	***MARITS***	
Unspecified		*1*
Divorced		*2*
Married		*3*
Partner		*4*
Single		*5*
Widow/er		*6*
Education	***EDU***	
Secondary (technical)		*1*
Secondary (general)		*2*
Post-secondary (technical)		*3*
Secondary (vocational)		*4*
Post-secondary (vocational)		*5*
University		*6*
Housing status	***HOUSE***	
Unspecified		*1*
Living with parents		*2*
Sharing property		*3*
Personal property		*4*
Renting		*5*
Student dormitory		*6*
Employment status	***EMPLOYS***	
Employed		*1*
House wife		*2*
Pensioner		*3*
Student		*4*
Employment duration (in months)	***EMPLOYY***	***continuous***
Employment type	***EMPLOYT***	
Unspecified		*1*
Bank/insurance company		*2*
Enterpreneur		*3*
Foreign company		*4*
Private company		*5*
Public organization		*6*
Net monthly income (in CZK)	***INCOME***	***continuous***
Region (NUTS2)	***REGION***	***dummy***
Credit bureau score	***CBSCORE***	***continuous***
Application score	***APPSCORE***	***continuous***
Behavioral score	***BEHAVSCORE***	***continuous***

Source: Random sample of consumer loans from the Bank, data from 2007-2013.

Table A2. The list of personal loan information (Panel B)

Variable description	Variable name in dataset	Encoding
Loan term characteristics		
Requested amount (in CZK)	***RAMOUNT***	***continuous***
Year of loan request	***RYEAR***	***dummy***
Loan approval indicator	***APPROVED***	***dummy***
Approved amount (in CZK)	***AAMOUNT***	***continuous***
Interest rate (in %)	***IR***	***continuous***
Risk margin (in %)	***RM***	***continuous***
Approved loan maturity (in months)	***AMATURITY***	***continuous***
Risk band	***NRISK***	
Very low-risk		*1*
Low-risk		*2*
High-risk		*3*
Very high-risk		*4*
Credit bureau information	***CBINFO***	***dummy***
Loan with specified purpose	***PURPOSE***	***dummy***
Number of individuals monitored in the CBCB (in mil.)	***CBIND***	***continuous***
Long-term unemployment rate (in %)	***UNDUR***	***continuous***
Risk-based pricing	***RBPRICING***	***dummy***
Default indicator	***DEF***	***dummy***
Renegotiated loan	***RENEG***	***dummy***
Number of months to default	***DEFAULT***	***continuous***

Source: Random sample of consumer loans from the Bank, data from 2007-2013.

Table A3. Descriptive statistics (Panel A)

Variable name	Mean	Std. Dev.	Min	Max
Application characteristics	*Accepted and rejected loans (N=207 640)*			
Age (in months)	485	155	216	1 159
Female	0,479	0,500	0	1
Marital status				
Divorced	0,184	0,387	0	1
Married	0,418	0,493	0	1
Partner	0,012	0,107	0	1
Single	0,335	0,472	0	1
Widow/er	0,010	0,100	0	1
Education				
Secondary (general)	0,103	0,303	0	1
Post-secondary (technical)	0,015	0,120	0	1
Secondary (vocational)	0,400	0,490	0	1
Post-secondary (vocational)	0,387	0,487	0	1
University	0,084	0,278	0	1
Housing status				
Living with parents	0,170	0,375	0	1
Sharing property	0,033	0,180	0	1
Personal property	0,541	0,498	0	1
Renting	0,220	0,414	0	1
Student dormitory	0,000	0,009	0	1
Employment status				
House wife	0,030	0,172	0	1
Pensioner	0,142	0,349	0	1
Student	0,001	0,029	0	1
Employment duration (in months)	71	90	0	579
Employment type				
Bank/insurance company	0,017	0,129	0	1
Enterpreneur	0,027	0,161	0	1
Foreign company	0,032	0,176	0	1
Private company	0,261	0,439	0	1
Public organization	0,178	0,383	0	1
Net monthly income (in CZK)	17 451	11 861	1	500 000
Loan with specified purpose	0,102	0,303	0	1
Existence of credit bureau information	0,756	0,429	0	1
Risk band				
Low-risk	0,362	0,480	0	1
High-risk	0,122	0,327	0	1
Very high-risk	0,029	0,167	0	1
Loan approval indicator	0,510	0,500	0	1

Source: Author's (2014) computations, data from 2007-2013. *Note:* Loan characteristics are available only for approved loans.

Table A3. Descriptive statistics (Panel B)

Variable name	Mean	Std. Dev.	Min	Max
Loan term characteristics	*Accepted loans (N=105 759)*			
Approved amount (in CZK)	93 653	82 100	4 000	1 000 000
Approved loan maturity (in months)	54,0	26,5	1,0	134
Interest rate (in %)	13,4	2,8	3,7	25,9
Long-term unemployment rate (in %)	2,8	1,2	0,7	6,1
Risk margin (in %)	1,8	1,4	-5,2	10,6
Number of individuals monitored in the CBCB (in mil.)	4,9	0,3	4,2	5,3
Default indicator	0,04	0,19	0	1
Credit bureau score	318	269	-40	1 120
Application score	178	222	-4	998
Behavioral score	454	192	0	1 012

Source: Author's (2014) computations, data from 2007-2013. *Note:* Loan characteristics are available only for approved loans.

Table A4. Summary statistics of risk margin by year

Year of loan request	N	Mean	Standard deviation
2007	12 167	1.40	1.44
2008	16 567	1.52	1.38
2009	18 378	1.79	1.39
2010	17 784	1.88	1.39
2011	17 122	1.99	1.44
2012	36 866	2.65	2.32
2013	10 523	2.47	2.18
Total	**129 407**	**2.06**	**1.84**

Source: Author's calculations. *Note:* Prior January 2012 the missing risk margin data is derived based on predicted value from a regression analysis of the complete data.

Table A5. Estimation results of loan demand and default probability

Dependent variable	Loan demand				Default probability			
	Pooled sample		Low-income sample		Pooled sample		Low-income sample	
	Coef.	***St.error***	**Coef.**	***St.error***	**Haz.ratio**	***St.error***	**Haz.ratio**	***St.error***
Age	-0.001***	*0,001*	-0.001	*0,001*	1.000**	*0.001*	0.999	*0.001*
Female	-0.125***	*0,004*	-0.053***	*0,006*	0.742***	*0.028*	0.710***	*0.035*
Education								
Secondary (general)	-0.125***	*0.178*	-0.080***	*0.023*	1.762***	*0.216*	1.592**	*0.253*
Post-secondary (techn.)	0.065***	*0,021*	0.016	*0,031*	0.607**	*0.127*	0.449**	*0.158*
Secondary (voc.)	0.034***	*0,016*	0.001	*0,022*	0.751**	*0.088*	0.674**	*0.105*
Post-secondary (voc.)	-0.054**	*0,017*	-0.048**	*0,022*	1.067	*0.124*	1.017	*0.156*
University	0.130***	*0.173*	0.076***	*0.026*	0.396***	*0.059*	0.559**	*0.134*
Employment status								
House wife	-0.129***	*0,015*	0.063***	*0,017*	1.050	*0.121*	0.956	*0.125*
Pensioner	-0.160***	*0,010*	-0.023**	*0,011*	0.529***	*0.040*	0.573***	*0.052*
Student	-0.257***	*0,056*	-0.119*	*0,064*	1.597	*0.717*	1.409	*0.712*
Employment duration	-0.001*	*0.001*	-0.001***	*0.000*	0.996***	*0.001*	0.997***	*0.001*
Employment type								
Bank/insurance company	-0.038**	*0,018*	0.157***	*0,040*	0.410**	*0.111*	0.635	*0.243*
Enterpreneur	-0.013	*0,012*	0.021	*0,014*	1.180*	*0.102*	1.148	*0.124*
Foreign company	0.060***	*0,010*	0.017	*0.016*	0.986	*0.062*	1.152	*0.102*
Private company	0.010*	*0.005*	-0.017**	*0.009*	0.880**	*0.046*	1.129	*0.085*
Public organization	-0.071***	*0.006*	-0.039***	*0.010*	0.691***	*0.042*	0.786**	*0.066*
Net monthly income	0.001***	*0,001*	0.001***	*0.001*	1.001	*0.001*	0.999***	*0.001*
Marital status								
Divorced	0.022	*0.019*	-0.063***	*0.023*	1.013	*0.144*	0.984	*0.173*
Married	0.116***	*0.018*	0.080***	*0.022*	0.793*	*0.111*	0.840	*0.144*
Partner	0.091***	*0,026*	0.078**	*0.033*	0.929	*0.188*	0.928	*0.243*
Single	0.098***	*0.019*	0.026	*0.023*	0.979	*0.138*	1.001	*0.173*
Widow/er	0.072***	*0,021*	-0.012	*0.024*	0.908	*0.159*	1.001	*0.213*

(continued on next page)

Table A5. Estimation results of loan demand and default probability

Dependent variable	Loan demand				Default probability			
	Pooled sample		Low-income sample		Pooled sample		Low-income sample	
	Coef.	***St.error***	**Coef.**	***St.error***	**Haz.ratio**	***St.error***	**Haz.ratio**	***St.error***
Housing status								
Living with parents	0.104***	*0,012*	0.093***	*0.015*	0.564***	*0.048*	0.564***	*0.061*
Sharing property	0.030**	*0.014*	-0.031	*0.019*	0.616***	*0.066*	0.670**	*0.094*
Personal property	0.035***	*0.011*	0.005	*0.014*	0.562***	*0.046*	0.574***	*0.060*
Renting	0.030***	*0,011*	-0.026*	*0.015*	1.000	*0.079*	1.015	*0.104*
Student dormitory	0.089***	*0,199*	0.067	*0.234*	1.812	*1.294*	1.685	*1.213*
Region. Loan Purpose					*yes*			
$\mathbf{R^2}$	0,5093		0,4639					
N	105 759		46 598		105 759		46 598	
Log likelihood					-38 221		-20 223	
Prob> chi2					0.000		0.000	
Loglikelihood ratio (LR) chi2					4 858		2 639	

Source: Author's computations, 2007-2013. *Note:* (1) For loan demand estimation the logarithmic form of approved loan amount is used. (2) Robust standard errors are used for statistical inferences. (3). ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Working Paper Series
ISSN 1211-3298
Registration No. (Ministry of Culture): E 19443

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the ASCR, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: http://www.cerge-ei.cz

Editor: Marek Kapička

The paper is available online at http://www.cerge-ei.cz/publications/working_papers/.

ISBN 978-80-7343-343-7 (Univerzita Karlova. Centrum pro ekonomický výzkum
a doktorské studium)
ISBN 978-80-7344-335-1 (Akademie věd České republiky. Národohospodářský ústav)

CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
http://www.cerge-ei.cz