Does Loan Maturity Matter in Risk-based Pricing?

Evidence from Consumer Loan Data

Gabriela Pásztorová[1]

Abstract

This paper investigates the role of loan contract terms in the performance of consumer credit. Taking advantage of a sample of both 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 provide evidence that loan performance is time- dependent and default depends on the choice of loan duration. I argue that a risk-based maturity setting improves the quality of granted consumer loans and alleviates the adverse selection present on the lending market.

Keywords: credit scoring, consumer loans, asymmetric information
JEL Codes: D12, D14, D82, G21

---

[1] CERGE-EI, Prague (a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic), Politickych veznu 7, Praha 1, Czech Republic. 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.

1 Introduction

Over recent decades, substantial changes in the volume of consumer loans have been observed worldwide. Particularly in emerging markets, despite the initial difficulties of the availability of only minimal credit history on borrowers and pioneering methods used to evaluate the creditworthiness of borrowers, lending institutions instituted prudent but extensive provision of consumer loans. The quantitative importance of consumer loans in the emerging markets can be illustrated using the example of the Czech Republic, where between 2000 and 2011 the total volume of consumer loans rose from CZK 31.1 bn to CZK 159.4 bn [2].

The rapid growth of the consumer credit market brought 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 the borrowers (henceforth referred to as a borrower's 'riskiness') may suffer from adverse selection when deciding to grant a loan or not. Adverse selection appears when, being aware of their own riskiness, "low-risk" borrowers with low probability of default will not be willing to pay an increased price, but "high-risk" borrowers with a high probability of default will accept a higher interest rate. In other words, "high-risk" borrowers demand higher loan amounts and default with highest probability. To eliminate this type of excess demand, lenders might choose to deny loans instead of raising 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 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.

More recently, a number of studies find evidence of credit rationing on the loan market where borrowers have liquidity constraints. If these are binding and borrowers do not have sufficient funds to finance their desired consumption with resources that they will accumulate in the future, addressing access loan demand under imperfect information becomes more important. Alessie et al. (2005), Edelberg (2006) and Adams et al. (2009) examine how

[2] Source: Czech Statistical Office; http://www.czso.cz/csu/2004edicniplan.nsf/engpubl/10n1-04-2004

lenders cope with information asymmetry and address loan demand through differentiating interest rates based on the borrowers' riskiness. They find evidence that risk-based pricing of interest rates substantially mitigates the adverse selection present on the loan market.

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 et al. (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 while still acquiring the desired loan amount. The key assumption valid for borrowers with liquidity constraints is that they prefer to have reduced monthly cash flow rather than decreased interest rates. The authors shed light on the role of maturity on purchasing behavior (the demand side of the consumer loan market); however, limited and inconclusive empirical evidence exists about its implications for loan performance or pricing decisions (supply side of the consumer loan market).

The current paper attempts to fill in 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 loan granting and repayment equation, then estimate the elasticity of loan demand and probability of default with respect to interest rate and loan maturity. Second, I use the demand estimates to point out the role of a risk-based maturity setting in decreasing the information asymmetries on the loan market. Third, I show that the time of default is maturity-dependent and differs across borrowers in the different risk categories. The key contribution of the paper is that it shows that by reflecting the borrower's riskiness in the interest rate, lenders discourage risky borrowers from short-term loans and by prolonging maturity decrease their probability of default. Hence, a risk-based maturity setting improves the quality of granted consumer loans and alleviates the adverse selection present on the lending market.

The paper utilizes a unique dataset of both rejected and accepted consumer loans from a Czech commercial bank (hereafter, the "Bank")[3]. This 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.

2 Consumer Loan Market

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 four key steps of the lending process.

As the first step, the individual enters the consumer loan market by submitting an application form for a loan. The borrower discloses information about his/her socio-demographic characteristics (e.g. age, marital status, education, etc.) and information related to the requested loan (e.g. loan amount, purpose of loan, etc.).

In the second step, the lender determines whether to grant the requested loan to the borrower. 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. Lenders try to predict default and make a decision to grant a loan (or not) based on the loan application characteristics and the available credit history[4] of the customer. 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. Applicants are then given a score by summing up the points based on application characteristics and verified credit history. Depending on this score and the corresponding probability of default, applicants are

[3] The Bank does not wish to be explicitly identified.

[4] In the case of new clients the Bank gains credit history about the client from the CBCB - Czech Banking Credit Bureau and SOLUS, which collect positive and negative information on a client's credit history credibility and payment behavior In the case of existing clients the Bank utilizes credit history information also from its own records.

categorized in a band with customers with similar characteristics. Applicants are accepted or rejected for the loan based on the band's loan approval cut-off threshold. Applicants are classified based on the bank's assessment of probability of default into the following bands: "very low-risk", "low-risk", "high-risk" and "very high-risk" [5]. Based on their riskiness applicants interested in obtaining a loan amount l with loan maturity t are offered by the lender an interest rate r for borrowing. Consequently, using loan repayment schedule with equal total payments, the lender charges the borrower a monthly annuity payment of $a(l) = (l * r)/(1-(1+r)^{-t})$.

In the third step, the borrower decides whether to accept the loan contract terms offered by the lender. Assuming that the interest rate is derived based on the borrower's application characteristics and requested loan amount, the borrower can either decide not to accept the loan or accept the loan with offered loan contract terms. The borrower will accept the contract offered by the lender if his/her utility from accepting contract c given application characteristics x is higher than his/her from not accepting contract c given application characteristics.

As a last step, given that the lender and the borrower agreed on loan contract terms and the borrower is granted the loan, the borrower starts repaying the principal and interest in the form of monthly annuity payments. During the period of loan repayment, the borrower can choose early repayment or regular payments. In the case of early repayment, the borrower is penalized to cover the interest loss of the lender.[6] In the case of regular payments, the

[5] From January 2012 the Bank applies risk-based pricing, which is reviewed and developed periodically.

[6] While the costs granting a loan vary slightly over time, the returns on loans are spread over time and short term loans can cause losses for financial institutions in the case of early repayment. That is, despite the fact that these loans are never predicted to default, the discounted net loss in case of early repayment might exceed the profits that they are presumed to generate.

borrower either fully repays the loan or defaults[7]. The lender always chooses to offer loan contract terms that maximize its profit. [8]

3 Methodology

Overall, the main objective of this paper is to develop an empirical method that demonstrates the role of risk-based pricing and loan maturity on the 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.

3.1 Modeling Loan Demand

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

$$l = x\alpha_1 + \alpha_2 r + \alpha_3 t + \varepsilon_l \,, \tag{1}$$

where l is the granted loan amount, x is the vector of the information on application characteristics, risk bank, credit history ; r is the loan interest rate, t is the loan maturity, and ε_l is the unobserved error term.

Endogeneity

The loan demand estimation is complicated by the endogeneity of interest rate and maturity. The endogeneity of loan contract terms can cause the parameter estimates to be

[7] If the borrower does not meet its monthly payment obligations through three subsequent months, s/he is considered to be in default.

[8] During loan repayment the borrower can decide to renegotiate the originally signed contract terms – s/he can make an extraordinary payment, restructure the loan or consolidate several loans. The model described in this paper does not allow for such renegotiation of loan contract terms.

biased. Interest rate endogeneity arises as lenders can change the interest rate 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. Endogeneity of maturity is a further issue if the borrower cares primarily about monthly cash flow rather than the price that is paid for 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 the 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 loan only to that extent that the decreased monthly payments are acceptable for her expected future cash flow. The lender aims to prolong the loan maturity as it is associated with higher interest income, while this higher riskiness of borrower is implicitly reflected in the higher interest rate. It is questionable how successful the lender is in transferring the riskiness of borrower into the loan price or how significant the adverse selection on the market is. I discuss this issue in more detail in the next section.

To tackle the endogeneity problem and obtain unbiased estimates, I take advantage of selected local labor market conditions to create instruments for loan interest rate and maturity.

The exogenous variation in the interest rate is captured by information on the average monthly income of the borrower's region. Specifically, I calculate the rate of a borrower's monthly income compared to the average disposable income observed in his/her region at the time of loan application. Borrowers with a monthly income lower than the region's average signal low probability of repayment for the lender. The lender's response is a higher interest rate (lower interest rates are offered only on smaller loans) to capture the expected riskiness of the borrower. Being aware of his/her own riskiness the „low-risk" borrower refuses to pay the increased loan price, while the „high-risk" borrower will accept it, expecting lower probability of repayment. I assume that the monthly income serves as a proxy for the riskiness of the borrower and that the lender is the one who primarily sets the final interest rate on the market. At the same time, the variable has no effect on the loan amount, as independent of the region's average disposable income, both higher-income and lower-income borrowers can have different needs or preferences for smoothing their consumption. Bicakova et al. (2011) support this assumption by providing evidence that the correlation between borrowers' indebtedness and the average monthly income in the Czech regions is not statistically significant[9].

The change in unemployment duration in the region serves as an instrument for the borrower's choice of loan maturity. Specifically, I follow Jurajda and Munich (2002) and use the long-term unemployment rate (LTU, hereafter) as a measure of unemployment duration. The LTU is defined as the number of unemployed looking for a job over one year divided by the total number of unemployed workers. The borrower's maturity decision entering the credit loan market reflects the local labor market conditions in the form of months required to find a job in the region. In a region with a long average duration of unemployment, maturity is likely to be shortened, as the borrower does not want have debt burden in the case of being unemployed for a longer period. I assume that the borrower is the one who primarily decides about the length of the loan. On the other hand, we assume that the change in a region's long-term unemployment rate does not influence the individual's decision about the amount of a loan. 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 save over a period of time for an expenditure), s/he is not discouraged from borrowing just because s/he leaves in a region which experienced an increase in long-term unemployment rates. What s/he primarily cares about in such a region are the favorable loan contract terms.

Specifically, I estimate interest rate and loan maturity by the following equations:

$$r = x\beta_1 + w\beta_2 + \varepsilon_r \ , \qquad (2)$$

$$t = x\chi_1 + u\chi_2 + \varepsilon_t \ , \qquad (3)$$

[9] According to Bicakova et al. (2011) this labor market condition affects mainly the marginal borrowers who are at the edge of their repayment ability.

where w is rate of a borrower's monthly income compared to the average disposable income observed in his/her region, u is the change in the unemployment duration in the borrower's region, and $\varepsilon_r, \varepsilon_t$ are the unobserved error terms.

Sample selection

Before estimating the model, I have to also deal with the nonrandom character of the consumer loan data. 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.

This paper does not account for those individuals who did not apply for a loan. We 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 probability it will be granted, as credit bureaus collect only information on borrowers who were eventually provided a loan[10]. If the borrower only tries the credit scoring evaluation and is rejected, it is not recorded in any credit bureau system. Thus, unless the customer has a bad loan repayment or default history connected with a previously provided loan, being rejected has no direct impact on the quality of his future loans. As loan application does not imply cost to the customers, there is no reason why an individual should not try the bank's credit scoring process.

On the other hand, the paper takes into account the limited information on those who applied, but eventually did not sign the loan contract. This appears either because the Bank rejects the applicant or because the applicant does not accept the loan contract terms offered by the Bank. Therefore, I follow Heckman (1979) and first estimate the selection equation on the whole sample of applicants that applied for a loan. The exclusion restriction for the selection equation is the Bank's behavioral score derived based on the individual's credit

[10] The CBCB - Czech Banking Credit Bureau was established in 2000 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/

history. This information is gained either from the Bank's own records (if the individual is already a client of the bank) or from the databases of credit bureaus. This behavioral score is assumed to be the key factor that decides whether the Bank approves the applicant's loan request. The more positive information is available about the credit history of the borrower from the Bank's records, the more likely it is that the borrower is reliable and will have no difficulties to maintain the regular monthly cash flow for loan repayment. At the same time, the borrower's decision about the requested loan amount is independent of credit history in the Bank. The available credit history affects the decision of the prospective borrower to apply for a loan rather than the amount he/she applies for.

To jointly account for both endogeneity and sample selection, I extend the sample selection model for endogeneous explanatory variables suggested by Wooldridge (2010) 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):

$$b = 1(x\delta_1 + w\delta_2 + u\delta_3 + h\delta_4 + \varepsilon_b > 0) \tag{4}$$

where w is rate of a borrower's monthly income compared to the average disposable income observed in his/her region, u is the change in the unemployment duration in the borrower's region, h is the behavioral score of the individual and $\varepsilon_r, \varepsilon_t, \varepsilon_b$ are the unobserved error terms.

The following assumptions are made:

(i) (x, w, u, h, b) is always observed, (l, r, t) is observed when $b = 1$;

(ii) $(\varepsilon_l, \varepsilon_b)$ is independent of (x, w, u);

(iii) $\varepsilon_b \sim$ Normal (0, 1);

(iv) $\mathrm{E}(\varepsilon_l \mid \varepsilon_b) = \alpha_4 \varepsilon_b$;

(v) $\mathrm{E}(z_1{}'\varepsilon_r) = 0$, where $z_1\beta = x\beta_1 + u\beta_2$ and $\beta_2 \neq 0$;

$\mathrm{E}(z_2'\varepsilon_t) = 0$, where $z_2\chi = x\chi_1 + u\chi_2$ and $\chi_2 \neq 0$.

I estimate the loan demand equation by 3SLS, where I add the inverse Mills ratio to the explanatory variables:

$$l = x\alpha_1 + \alpha_2 r + \alpha_3 t + g(x,w,u,h,b) + \varepsilon_g \,, \tag{5}$$

where $g(x,w,u,h,b) \equiv \mathrm{E}(\varepsilon_l \mid x,w,u,h,b)$ and $\varepsilon_g \equiv \varepsilon_l - \mathrm{E}(\varepsilon_l \mid x,w,u,h,b)$ implies $\mathrm{E}(\varepsilon_g \mid x,w,u,h,b) = 0$.

It also holds that $\mathrm{E}(\varepsilon_l \mid x,w,u,h,b=1) = \alpha_4\lambda(x\delta_1 + w\delta_2 + u\delta_3 + h\delta_4)$.

To sum up, the estimation is performed in two steps. First, $\delta_1, \delta_2, \delta_3, \delta_4$ are consistently estimated by probit from the selection equation (4) using all observations and the estimated inverse Mills ratio $\hat{\lambda}_i = \lambda(x_i\hat{\delta}_1 + w_i\hat{\delta}_2 + u_i\hat{\delta}_3 + h_i\hat{\delta}_4)$ is obtained. Second, on the subsample where r and t are observed, I estimate by 3SLS the following equation:

$$l_i = x_i\alpha_1 + \alpha_2 r_i + \alpha_3 t_i + \alpha_4\hat{\lambda}_i + \varepsilon_i \,. \tag{6}$$

I test for the null hypothesis of no selection bias ($\mathrm{H}_0 : \alpha_4 = 0$) by exploiting the 3SLS t statistic for $\hat{\alpha}_4$; and test the null hypothesis of no endogeneity by estimating the structural model (1) that includes the residuals from the 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 a model that uses the 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 and test for the significance of asymmetric information hidden in the maturity choice.[11]

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 talk about default) occurrence in multiplicate form. If $\lambda(t_d, x)$ is the probability that an individual defaults at time t_d (conditional on paying regular payments till default), x 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

$$\lambda(t_d, x) = \lambda_o(t_d).\exp(x\phi_1 + l\phi_2 + r\phi_3 + m\phi_4) \qquad (7)$$

The advantage of proportional hazard models is that whereas parametric models use information over the whole time horizon (distributional assumption for baseline hazard $\lambda_0(t_d)$ and estimation of the cumulative hazard), semi-parametric models use only the information at failure times (no distributional assumption for baseline hazard and estimation of the direct hazard).

The incomplete information on the occurrence of event during observation period belongs among the specifics of duration time estimation. I deal with censored data, a situation in which I stop following the individuals in the sample.[12] There are two possibilities of the event status: the event occurred by t_d^* (duration time) or the event did not occur by the end of observation period t_c (censoring time). For each individual one observes t_d, where $t_d = \min(t_d^*, t_c)$.

[11] Flannery (1986), Diamond (1991) and Berger et al. (2005) are 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.

[12] As the information about the loan performance after the end of the observation period is missing, the data is right censored.

I model loan size and default jointly. The final default probability model includes the estimated residuals ξ from loan demand as a control variable:

$$\lambda(t_d, x) = \lambda_o(t_d).\exp(x\phi_1 + \phi_2 r + \phi_3 m + \phi_4 s + \phi_5 \xi) \qquad (8)$$

The identification is through the quarterly change of saving rate s as a time-varying macroeconomic shock during the period of loan repayment. This factor affects the probability of repayment, but has no implications for the requested loan amount that has been already agreed. The models for loan demand (6) and the default probability (8) are estimated for short-, medium- and long- term loans[13] and across borrowers in the different risk categories.

4 Data

4.1 Consumer Loan Data

The data sample consists of the consumer loan information of over 220,000 individuals. It includes installment loans (consumer loans, cash loans). The dataset includes application characteristics (e.g. age, marital status, education, etc.), loan contract information (e.g. interest rate, loan maturity, loan size, etc.) and performance indicators (e.g. date of default, monthly outstanding balance, past due, etc.). The consumers requested the loans between 2007 and 2013[14], where the last performance observation is from April 2013. Table 1 summarizes the list of information on the available consumer loan. Table 2 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

[13] Glennon and Nigro (2005) show the determinants of default are maturity-specific.

[14] 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.

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, the 3.6 % of those who had obtained a loan had defaulted and the rest of the individuals performed well. Although there are several different definitions of "defaulted" loans, the one of the Basel Committee on Banking Supervision (2004) is used: a loan is defaulted 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 low score) or the borrower (due to unfavorable loan terms offered by the lender). Although information on application characteristics are available for the whole sample of consumer loans, information on interest rates for rejected borrowers are not observed. As discussed in the previous section, I deal with this problem by employing a sample selection model for loan demand estimation that accounts for endogeneity as well.

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. As opposed to the density, the survivor and the hazard functions are easily interpretable and effective in describing the duration dependence.

In Figure 1 the cumulative (integrated) hazard function with 95% confidence intervals is plotted estimated by Nelson-Aalen method. It suggests that at the end of the consumer loan observation period, more than 95% 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 paying regular payments until a particular month during analysis time. According to the smoothed hazard function that treats all consumer loans equally and does not distinguish between maturity or risk bands (henceforth referred to as 'pooled'), defaults are most likely to occur around the 20th 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.

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 I highlight the implications of risk-based pricing on the quality of granted loans, i.e. on the probability of default. Finally, I illustrate the maturity-dependent default probability for borrowers in the different risk categories.

5.1 Loan Demand

As the first step in the estimation of loan demand, I correct for the nonrandom feature of the data and estimate the probability of loan approval based on selection equation (6)[15]. 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 borrower's behavioral score (credit history) is used as an exclusion restriction.

As a second step, using the estimated inverse Mills ratio 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

[15] 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.

restrictions are the borrower's monthly income relative to the disposable income observed in her/his region[16] and the increase in the average long-term unemployment rate in the borrower's region[17]. The F-tests confirmed that the instruments are valid. Overall, the elasticity of loan demand is statistically significant with respect to both loan term conditions (Table 3, Panel A). With increasing interest rate, individuals are discouraged from borrowing, whereas with longer maturity the loan amount increases. In Table 3 I compare the interest rate and maturity elasticity of loan demand for the pooled sample (Panel A, Column 2) and for the subsample of low-income borrowers[18] (Panel A, Column 4). The results suggest that the loan demand of a low-income borrower increases with longer maturity, while the interest rate has no statistically 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.

5.2 Probability of Default

The default probability estimation based on consumer loan application characteristics is conducted using the Cox proportional hazard model. In addition to the application and loan characteristics, the estimated residual from the loan demand equation and the quarterly change of saving rate[19] are included into the model as control variables. Table 3 summarizes the estimation results 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

[16] Source: Czech Statistical Office;
http://apl.czso.cz/pll/rocenka/rocenkavyber.volba?titul=Ukazatele%20v%20region%E1ln%EDm%20%E8len%ECn%ED&mypriznak=RC&typ=2&proc=rocenka.presmsocas&mylang=CZ&jak=4

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

[18] The sub-sample of low-income borrowers represents those borrowers who have their net monthly income at the time of application below the sample's median net income.

[19] Source: Eurostat;
http://epp.eurostat.ec.europa.eu/portal/page/portal/sector_accounts/data/quarterly_data

specification for the relationship between hazard rates and the application characteristics.[20] Column 6 and Column 8 in Table 3 quantify the hazard rate, $\exp(\beta)$, for the application characteristics as a percentage of the hazard rate for their reference groups. The effect of individual application characteristics on default probability is in line with the expectations. The longer survival time without default increases with higher education and being employed for longer period. For instance, the hazard ratio for borrowers with university education is only 44% of the hazard rate for those who have upper secondary technical education. Similarly, borrowers with own property are associated with a 42% lower risk of default at any time from loan provision than those not indicating housing status with the same observed characteristics.

More importantly, the results also provide evidence of the effect of risk-based pricing (variable *RBPRICING*) introduced in 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*AMATURITYC*). The hazard ratio on this interaction term suggests that given risk-based pricing, an increase in loan maturity decreases the probability of default by 12% (derived from coefficients corresponding to Table 3, Panel A, Column 6). In other words, with the decrease of asymmetric information between lenders and borrowers, "high-risk" borrowers choose either to reduce the loan amount or to prolong maturity to compensate the lender for their riskiness. The effect of risk-based pricing for the subsample of low-income borrowers is not statistically significant (Table 3, Panel A, Column 8) due to the reduced sample size (low default occurrence) in the observation period after January 2012.

Figure 3 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: short-term loans (maturity up to two years) are the most likely to default after the 18^{th} month of granting; medium term loans (maturity between two and five years) are the most likely to default at the 20^{th} month, and long term loans (more than five years maturity)

[20] The reference group for the application factor variables is always the one with the lowest coding. For the coding of variables refer to Table 1.

default most frequently after the 24th month. Comparing the pooled proportional hazards and the proportional hazards by maturity, all achieve their peak just before the end of second year.

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. The default variation plotted in Figure 4 is the most significant for medium-term loans. 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. Comparing the dashed line with Cox-Snell residuals in Figure 4, it can be concluded that the maturity-specific models fit the data equally good as the model for the pooled sample.

6 Conclusion

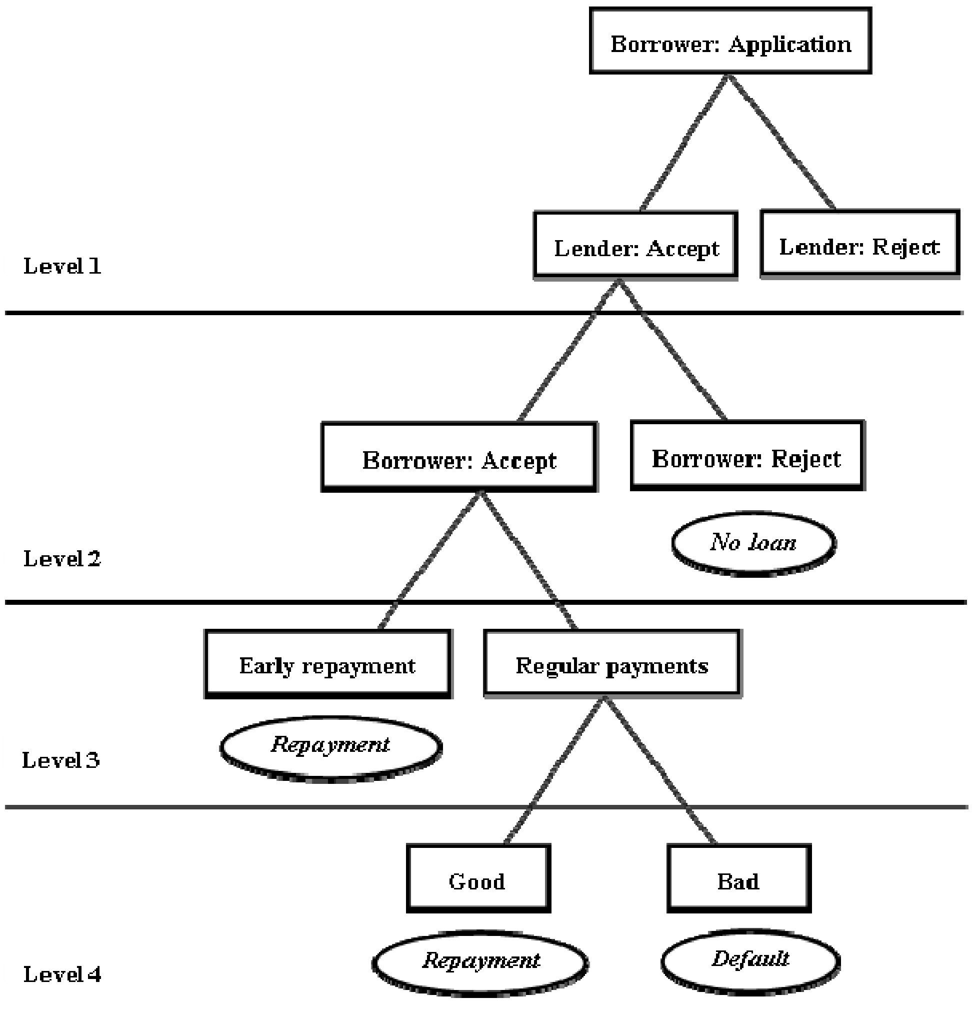
Driven by the sharp increase in consumer loan demand, the role of credit scoring methods in assessing a borrowers' creditworthiness is becoming more and more important. Thanks to the wide range of credit history collected by credit bureaus, lenders can screen out risky borrowers in their credit scoring 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 quantified.

This paper presents empirical evidence that 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, the paper contributes to the growing literature on credit scoring models by pointing 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, I show that low-income borrowers in this sample are 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 and by prolonging maturity decrease their probability of default. This is consistent with the theoretical predictions 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. Finally, I provide 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 the one that treats all consumer loans as pooled and does not distinguish between these two factors.

Appendix

Figure 1. The lending process and data availability



Source: Author's illustration of lending process based on the description of the Bank's representatives.

Note: Level 1 – Based on the borrowers application characteristics the borrowers decides to accept (then offers interest rate) or reject the borrower for a loan (no loan is originated). An initial maturity is requested by the borrower, but the lender can propose its change. Accept – available both application and loan contract characteristics, Reject – available only application characteristics.

Level 2 – Based on the lender's interest rate offer, the borrower has a chance to accept the loan contract conditions (open account) or reject (no loan is originated). Accept - available both application and loan contract characteristics, Reject – available only application characteristics. This paper treats as rejected both loans that were rejected by the lender or by the borrower.

Level 3 – The borrower either repays the loan in regular payments or makes an early repayment. Early repayment - information available, but the simplified model of this paper this is not taken into account. Regular payments – available full information on the time of repayment.

Level 4 – Good – the time of full repayment is observed, Bad – the time of default is observed.

Table 1. The list of personal loan information (Panel A)

Application characteristics	Name and encoding
Age	***AGE***
Female	***FEMALE***
Marital status	***MARITS***
Unspecified	*1*
Divorced	*2*
Married	*3*
Partner	*4*
Single	*5*
Widow/er	
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*
With parents	*2*
Sharing property	*3*
Owner of property	*4*
Rent	*5*
Student dormitory	*6*
Employment status	***EMPLOYS***
Employed	*1*
Housewife	*2*
Pensioner	*3*
Student	*4*
Years of being employed	***EMPLOYY***
Employment type	***EMPLOYT***
Unspecified	*1*
Financial	*2*
Enterpreneur	*3*
Foreign company	*4*
Private company	*5*
Public organization	*6*
Net monthly income	***INCOME***
Region	***NUTS 2***

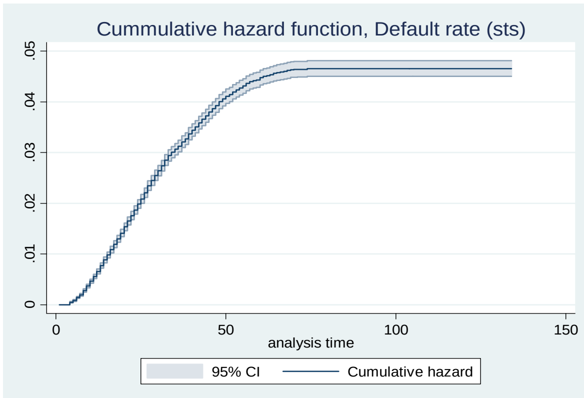
Source: Random sample of consumer loans from the Bank. *Note:* Dummies are created for the following variables: FEMALE (1/0). Continuous variables include: AGE, EMPLOYY, INCOME.

Table 1. The list of personal loan information (Panel B)

Loan term characteristics	Name and encoding
Requested amount	***RAMOUNT***
Year of loan request	***RYEAR***
Loan approval indicator	***APPROVED***
Approved amount	***AAMOUNT***
Interest rate	***IR***
Approved maturity	***AMATURITY***
Risk band	***RISK***
Very low-risk	*1*
Low-risk	*2*
High-risk	*3*
Very high-risk	*4*
Availability of credit bureau information	***CBINFO***
Loan purpose	***LOANPURP***
Unspecified	*1*
Purchase of a flat/house	*2*
Reconstruction of a flat/house	*3*
Construction of a flat/house	*4*
Share in a housing cooperation	*5*
Co-purchasing a flat	*6*
Purchase of a piece of land/garage	*7*
Purchase of a recreational facility	*8*
Reconstruction of a recreational facility	*9*
Electronic equipment	*10*
Settlement of inheritence	*11*
Purchase of a new car	*12*
Purchase of a used car	*13*
Youth housing	*14*
Education purpose	*15*
Behavioral score	***BEHAVSCORE***
Borrower's income relative to the region's disposable income	***ISHARE***
Change in long-term unemployment rate	***UNDURCH***
Change in saving rate	***SRATECH***
Risk-based pricing applied	***RBPRICING***
Default indicator	***DEFAULT***

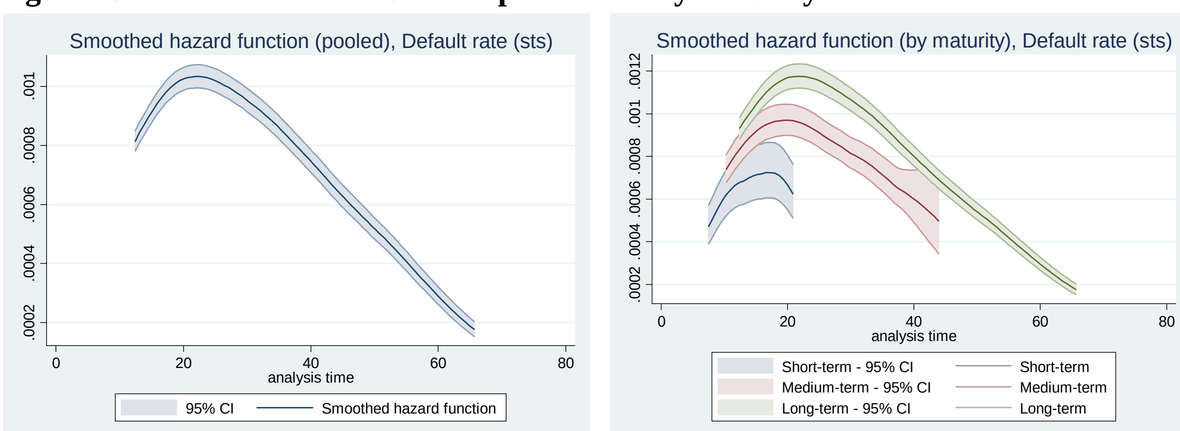
Source: Random sample of consumer loans from the Bank. *Note:* The requested loan amount (RAMOUNT) and the approved loan demand (AAMOUNT) are categorized into ten quantile categories. Dummies are created for the following variables: APPROVED (1/0), CBINFO (1/0), DEFAULT (1/0), RBPRICING (1/0) and RYEAR (year dummy). Continuous variables include IR and AMATURITY.

Figure 1. Nelson-Aalen estimator of the cumulative hazard function



Source: Author's computations, 2007-2013.

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, e.i. 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 2. Descriptive statistics

Variable name	N	Mean	Std. Dev.	Min	Max
Application characteristics					
AGE	207 640	485	155	216	1 159
FEMALE	207 640	0,479	0,500	0	1
MARITS					
Divorced	207 640	0,184	0,387	0	1
Married	207 640	0,418	0,493	0	1
Partner	207 640	0,012	0,107	0	1
Single	207 640	0,335	0,472	0	1
Widow/er	207 640	0,010	0,100	0	1
EDU					
Secondary (general)	207 640	0,103	0,303	0	1
Post-secondary (technical)	207 640	0,015	0,120	0	1
Secondary (vocational)	207 640	0,400	0,490	0	1
Post-secondary (vocational)	207 640	0,387	0,487	0	1
University	207 640	0,084	0,278	0	1
HOUSE					
With parents	207 640	0,170	0,375	0	1
Sharing property	207 640	0,033	0,180	0	1
Owner of property	207 640	0,541	0,498	0	1
Rent	207 640	0,220	0,414	0	1
Student dormitory	207 640	0,000	0,009	0	1
EMPLOYS					
Housewife	207 640	0,030	0,172	0	1
Pensioner	207 640	0,142	0,349	0	1
Student	207 640	0,001	0,029	0	1
EMPLOYY	207 640	63	85	0	1 325
EMPLOYT					
Financial company	207 640	0,017	0,129	0	1
Enterpreneur	207 640	0,027	0,161	0	1
Foreign company	207 640	0,032	0,176	0	1
Private company	207 640	0,261	0,439	0	1
Public organization	207 640	0,178	0,383	0	1
INCOME	207 640	17 451	11 861	1	500 000
CBINFO	207 640	0,756	0,429	0	1
RISK					
Low-risk	207 640	0,372	0,483	0	1
High-risk	207 640	0,136	0,343	0	1
Very high-risk	207 640	0,098	0,297	0	1
APPROVED	207 640	0,511	0,500	0	1
Loan characteristics					
AAMOUNT	106 100	93 710	82 255	4 000	1 000 000
AMATURITY	106 100	54	27	1	134
IR	106 100	17	19	2	73
DEFAULT	106 100	0,036	0,185	0	1

Source: Author's computations, 2007-2013. *Note:* (1) Variables AGE and EMPLOYY are reported in months. (2) Loan characteristics are available only for approved loans.

Table 3. Estimation results of loan demand and default probability (Panel A)

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***
IR	-0,031**	*0,016*	-0,001	*0,024*	0,975***	*0,003*	0,968***	*0,005*
AMATURITY/C	0,054***	*0,013*	0,058***	*0,011*	1,373***	*0,041*	1,274***	*0,048*
RBPRICING	0,484**	*0,209*	-0,012	*0,216*	0,340**	*0,145*	0,239**	*0,158*
AMATURITYC*RBPRICING					0,637**	*0,107*	0,704	*0,187*
INVMILLS	-0,430**	*0,194*	-0,317***	*0,082*				
AAMOUNT_RES					0,951***	*0,009*	0,943***	*0,013*
R-squared	0,398		0,464					
N	106 100		46 753		106 100		46 753	
Log likelihood					-40 105		-21 262	
LR chi-square test					5 371		2 542	

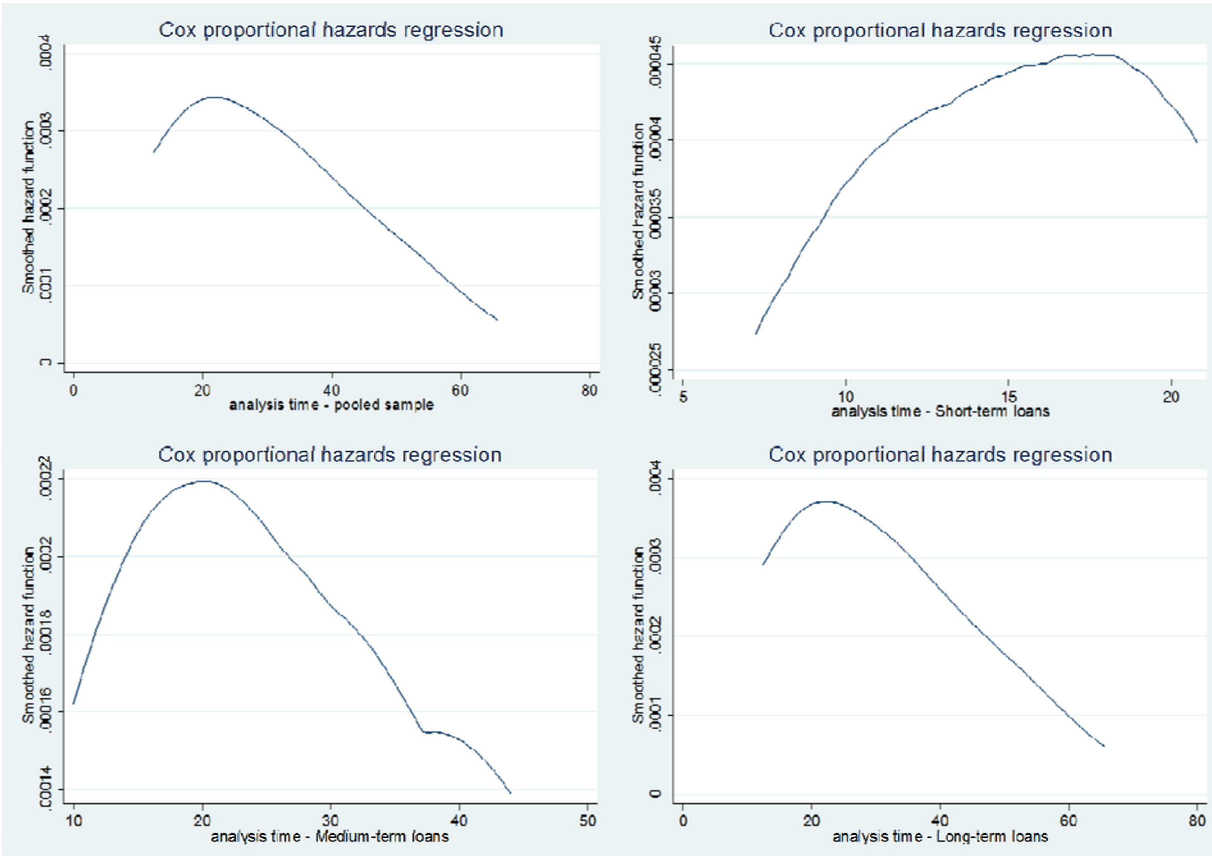
Source: Author's computations, 2007-2013. *Note:* (1) Loan demand expresses the approved loan amount. (2) INVMILLS denotes the Inverse Mills ratio calculated after estimating equation (4), AAMOUNT_RES denotes the estimated residual from the loan demand equation and AMATURITYC in default probability estimation denotes maturity categorized into short-term (up to 2years), medium-term (between 2 and 5years) and long-term (more than 5years). (3) Robust standard errors are used for statistical inferences. (4) Estimation results presented only for variables that were statistically significant at least in one model. * represents statistically significant at 10%, ** statistically significant at 5%, and *** statistically significant at 1%.

Table 3. Estimation results of loan demand and default probability (Panel B)

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*
AGE	0,001	*0,001*	0,001	*0,000*	1,001	*0,001*	0,999*	*0,001*
FEMALE	-0,333***	*0,014*	-0,176***	*0,023*	0,786***	*0,029*	0,724***	*0,035*
MARITS								
Divorced	-0,166***	*0,047*	-0,162***	*0,049*	1,056	*0,115*	0,998	*0,131*
Married	0,076**	*0,034*	0,220***	*0,041*	0,875	*0,094*	0,883	*0,114*
Partner	0,071	*0,070*	0,255***	*0,089*	0,970	*0,178*	0,898	*0,214*
Widow/er	-0,163**	*0,070*	0,028	*0,078*	1,171	*0,205*	1,221	*0,257*
EDU								
Secondary (general)	-0,315***	*0,059*	-0,205***	*0,077*	1,732***	*0,211*	1,525**	*0,240*
Post-secondary (technical)	0,106	*0,069*	-0,020	*0,095*	0,628**	*0,026*	0,433**	*0,152*
Post-secondary (vocational)	-0,199***	*0,055*	-0,151**	*0,070*	1,070	*0,123*	0,981	*0,148*
University	0,309***	*0,056*	0,139*	*0,078*	0,446***	*0,065*	0,562**	*0,132*
HOUSE								
At parents	0,262***	*0,041*	0,274***	*0,049*	0,603***	*0,051*	0,591***	*0,064*
Sharing property	-0,010	*0,051*	-0,082	*0,057*	0,756**	*0,081*	0,770*	*0,107*
Personal property	0,028	*0,046*	0,001	*0,044*	0,585***	*0,048*	0,615***	*0,065*
EMPLOYS								
House wife	-0,091*	*0,053*	0,246***	*0,047*	0,967	*0,111*	0,825	*0,106*
Pensioner	-0,342***	*0,099*	-0,037	*0,059*	0,526***	*0,039*	0,509***	*0,045*
Student	-0,853***	*0,200*	-0,398	*0,244*	1,175	*0,552*	0,929	*0,469*
EMPLOYT								
Financial company	1,140**	*0,481*	0,397	*0,748*	0,513**	*0,148*	1,113	*0,434*
Foreign company	0,105**	*0,039*	0,035	*0,046*	1,146	*0,098*	1,159*	*0,102*
Public organization	-0,146***	*0,034*	-0,103*	*0,054*	0,686***	*0,042*	0,771**	*0,065*
EMPLOYY	-0,001	*0,001*	-0,001***	*0,000*	0,996***	*0,001*	0,997***	*0,001*
INCOME	0,001***	*0,001*	0,001***	*0,000*	0,999	*0,001*	0,999***	*0,000*
RISK								
Low	-0,308***	*0,022*	-0,208***	*0,026*	2,247***	*0,104*	2,215***	*0,142*
High	-0,470***	*0,065*	-0,348***	*0,054*	3,339***	*0,172*	3,241***	*0,228*
Very-high	-0,281	*0,173*	-0,079	*0,094*	4,221***	*0,286*	4,082***	*0,369*
CBINFO	-0,389***	*0,124*	-0,435***	*0,087*	0,618***	*0,024*	0,724***	*0,037*
SRATECH					0,997	*0,003*	1,003	*0,003*
Constant	4,406***	*0,577*	2,668***	*0,696*				
Year dummy					***yes***			
Loan purpose categories					***yes***			
R-squared	0,398		0,464					
N	106 100		46 753		106 100		46 753	
Log likelihood					-40 105		-21 262	
LR chi-square test					5 371		2 542	

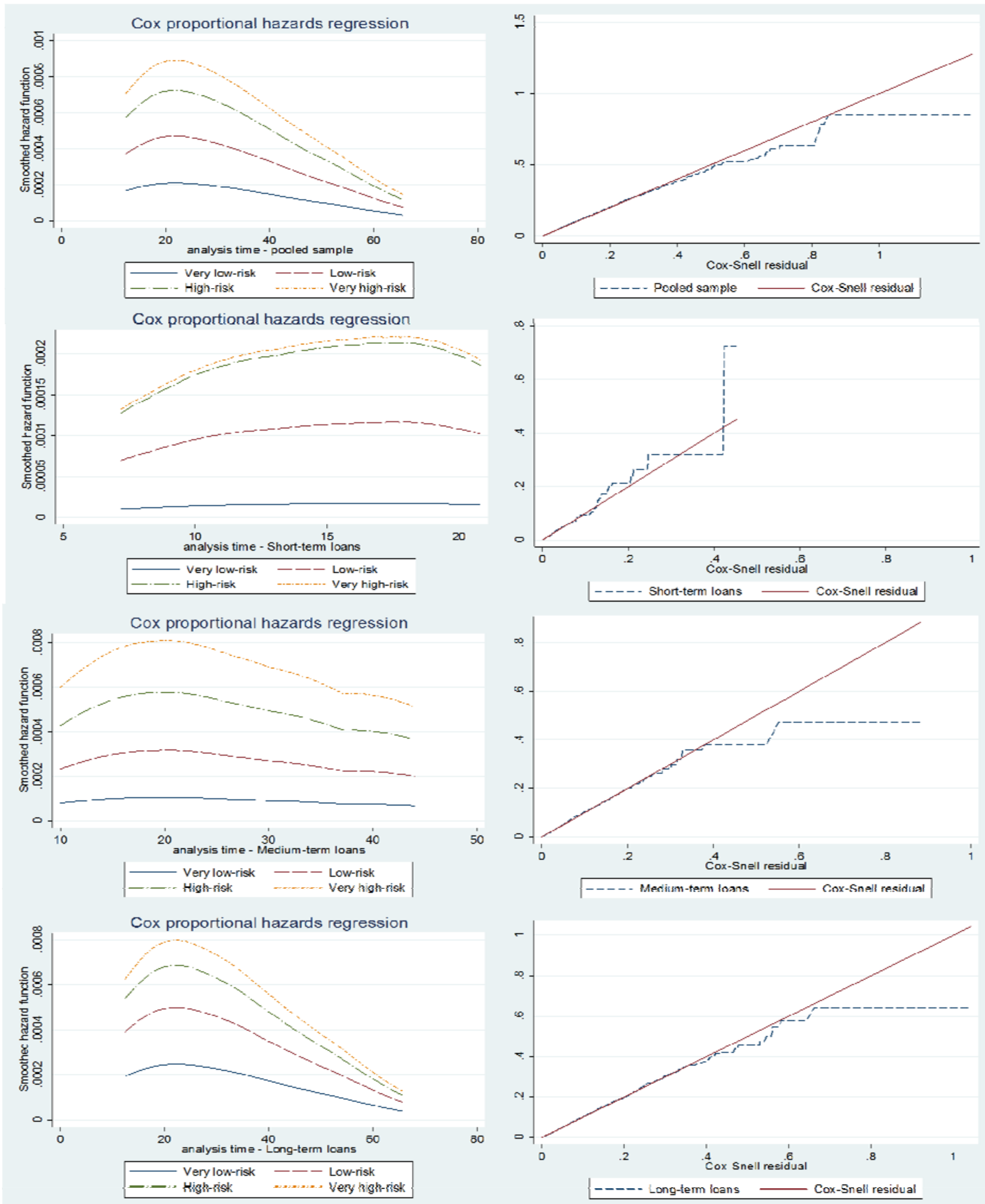
Source: Author's computations, 2007-2013. *Note:* Robust standard errors are used for statistical inferences.
* represents statistically significant at 10%, ** statistically significant at 5%, and *** statistically significant at 1%.

Figure 4. 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, e.i. 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.

Figure 5. 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.

References

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

Alessie, R., Weber, G., 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., 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.*

Bičáková, A., Prelcová, Z., and Pašaličová, R. (2011). 'Who Borrows and Who May Not Repay?' *CERGE-EI Working Paper Series* 443.

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

Cox, D.R. (1972). 'Regression models and life-tables (with discussion).' *Journal of Royal Statistical Society* B(74): 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.

Égert, B., Crespo-Cuaresma, J. and Reininger, T. (2007). 'Interest Rate Pass-Through in Central and Eastern Europe: Reborn from Ashes Merely to Pass Away?' *Journal of Policy Modeling* 29: 209-225.

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.

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

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

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

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

Wooldridge, J. M. (2010). 'Econometric Analysis of Cross Section and Panel Data. The MIT Press.' *Cambridge, MA, 2nd edition*.