LOANS FOR BETTER LIVING: THE ROLE OF INFORMAL COLLATERAL

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

541

Working Paper Series 541
(ISSN 1211-3298)

Loans for Better Living: The Role of Informal Collateral

Gabriela Kuvíková

CERGE-EI
Prague, May 2015

ISBN 978-80-7343-346-8 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-338-2 (Akademie věd České republiky. Národohospodářský ústav)

Loans for Better Living:

The Role of Informal Collateral

Gabriela Kuvíková*

Abstract

Consumers with insufficient resources can finance purchases by applying for specific purpose loans or unspecified purpose loans. I examine the default gap of these two types of loans by using a unique dataset of consumer loans from a Czech commercial bank. In line with theoretical models that perceive collateral as a screening device mitigating adverse selection, the paper confirms a negative relationship between the default rate and the presence of informal collateral. More importantly, it is not the purpose for the loan, but mainly the unobserved characteristics of the borrower that drive the default rate. The paper also provides empirical evidence that the interest rate differential between specific purpose loans and unspecified purpose loans is systematically higher than their default rate differential.

Abstrakt

Spotřebitelé s nedostatkem vlastních zdrojů mohou financovat svojí spotřebu účelovými nebo neúčelovými úvěry. Předmětem mého výzkumu je zmapování rozdílnosti v defaultu u těchto dvou typů úvěrů s využitím unikátní databáze spotřebitelských úvěrů získaných z jedné české komerční banky. V souladu s teoretickými modely, které vnímají záruku jako prostředek sloužící k zmírnění nežádoucího výběru na trhu úvěrů, tento článek potvrzuje negativní vztah mezi mírou defaultu a přítomností neformální záruky. Co je však důležité, není to účel úvěru, ale hlavně nepozorované vlastnosti dlužníka, které řídí míru defaultu. Tento článek také poskytuje empirické důkazy, že rozdíl v úrokové míře mezi účelovými a neúčelovými úvěry je systematicky vyšší než jejich rozdíl v míře defaultu.

Keywords: Consumer loans, Asymmetric information, Collateral, Default

JEL Classification Numbers: D12, G14, 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 Randall K. Filer, Jan Hanousek, Jan Kmenta and Evžen Kočenda for their valuable comments and guidance. All errors remaining in this text are the responsibility of the author.

1.Introduction

Since the early 2000s, the ways consumers may finance their expenditures have become diversified to a large extent. The range of loan products is particularly wide for financing housing-related expenditures. In addition to mortgage loans and building savings schemes, individuals can apply for housing loans granted for financing investments related to a property (e.g. home purchase, home renovation, home equipment). The key distinction between mortgage and housing loans is that the repayment of the latter is not secured by a lien on the property. Hence, housing loans are notably more attractive to those who are not willing or able to secure their loan with property. Alternatively, if the loan is intended to finance expenditures that are not housing-related, the borrower can apply for consumer credit. The key distinction between consumer credit and housing loans is that housing loans are granted conditional on the ownership of the real estate they finance, even though it does not serve as collateral. In this paper, housing loans and consumer credit with a designated purpose are jointly referred to as specific purpose loans ('purpose-loans'), while consumer loans without a designated purpose are referred to as unspecified purpose loans ('non-purpose loans'). The latter are viewed as bearing the highest risk, as no information is available on the expenditure they are intended to finance.

The cost of the loan products varies by their perceived riskiness. Mortgage loans are secured (the financed property serves as collateral and can be claimed by the lender in case of borrower bankruptcy) - their interest rate and probability of default (henceforth referred to as 'default rate') is relatively low compared to other types of loans. At the end of 2013 in the Czech Republic, the interest rate on new mortgage loans was 3.4 percent, while the share of non-performing loans to total mortgage loans was 3.0 percent.[1] By contrast, housing loans and consumer credit are unsecured loans (there is only a general claim on the borrower's assets in the case of default), and their interest rates and default rates are substantially higher than for mortgage loans. As of the end of 2013 in the Czech Republic, the interest rates on new consumer loans were 14.5 percent, while the share of non-performing loans to total consumer loans was 12.2 percent.[2]

Although previous literature has long emphasized the role of collateral in mitigating the asymmetric information between lenders and borrowers at the time of loan granting, their conclusions are contradictory. The theoretical predictions of Boot, Thakor and Udell (1991), Manove and Padilla (2001) and Inderst and Mueller (2007) suggest that with higher collateral the probability of default rises. The authors support their findings with several main arguments: (1) when they require increased collateral, financial institutions often weaken their screening mechanisms, (2) to achieve financing, the borrowers are likely to provide all the required collateral irrespective of their probability of default. A contrary view from Jimenez, Salas and Saurina (2006) supports the private information hypothesis. It says that collateral sorts loan applicants such that low risk borrowers prefer to pledge their loans (due to their low probability of default)

---

[1]Source: Czech National Bank –ARAD database – Monetary and financial statistics
http://www.cnb.cz/cnb/STAT.ARADY_PKG.STROM_DRILL?p_strid=A&p_lang=EN
[2]Source: Czech National Bank –Financial Market Supervision Report
http://www.cnb.cz/en/supervision_financial_market/aggregate_information_financial_sector/financial_market_supervision_reports/index.html

and have lower interest rates, while high risk borrowers prefer not to pledge their loans (given their higher probability of default) and have higher interest rates.

Despite the broad debate on collateral and its impact on loan performance, limited research has focused on the role of informal collateral in the housing loan market. Housing loans finance home equity (similar to mortgage loans), but are granted without collateral (similar to standard consumer loans). Instead, their loan contract terms are conditional on informal collateral, which exists whenever the lender has evidence of the good the loan is intended to finance. For a housing loan, homeownership and an invoice verifying the purpose of the loan serves as evidence of collateral. These help individuals applying for a housing loan signal their better creditworthiness. Because the existence of informal collateral makes the borrower eligible for favorable loan contract terms without a lien on the property, the information asymmetry between the lender and the borrower might be more severe. This paper addresses this issue and tests the effectiveness of informal collateral in alleviating adverse selection on the consumer loan market. It contributes to the findings of Kocenda and Vojtek (2011), who were the first to study the default probability of Czech consumer loans with different purposes.

This empirical paper focuses on three questions. First, I test whether the existence of informal collateral influences the likelihood of successful loan repayment, by applying a probit model to measure the effect of different loan types on the borrower's default rate. Second, I examine whether the lower default rate on purpose-loans is driven by the type of product they are intended to finance. This is tested by including loan purpose and applicant type dummies into the probit model. The latter is derived from information on multiple loan contracts per applicant and accounts for the fact that applicants with different default probability select different loan purposes. Third, I test whether applicants with the same application characteristics and loan contract terms have the same default rate and interest rate differential, regardless of whether they apply for loans with specified or unspecified purpose. I tackle the issue of self-selection by using propensity score matching.

The paper exploits a unique dataset of over 207 000 rejected and accepted consumer loans from a Czech commercial bank.[3] The dataset covers consumer loans granted from 2007 till 2013. It comprises three different types of consumer loans: housing loans and consumer loans with specified and unspecified purposes.

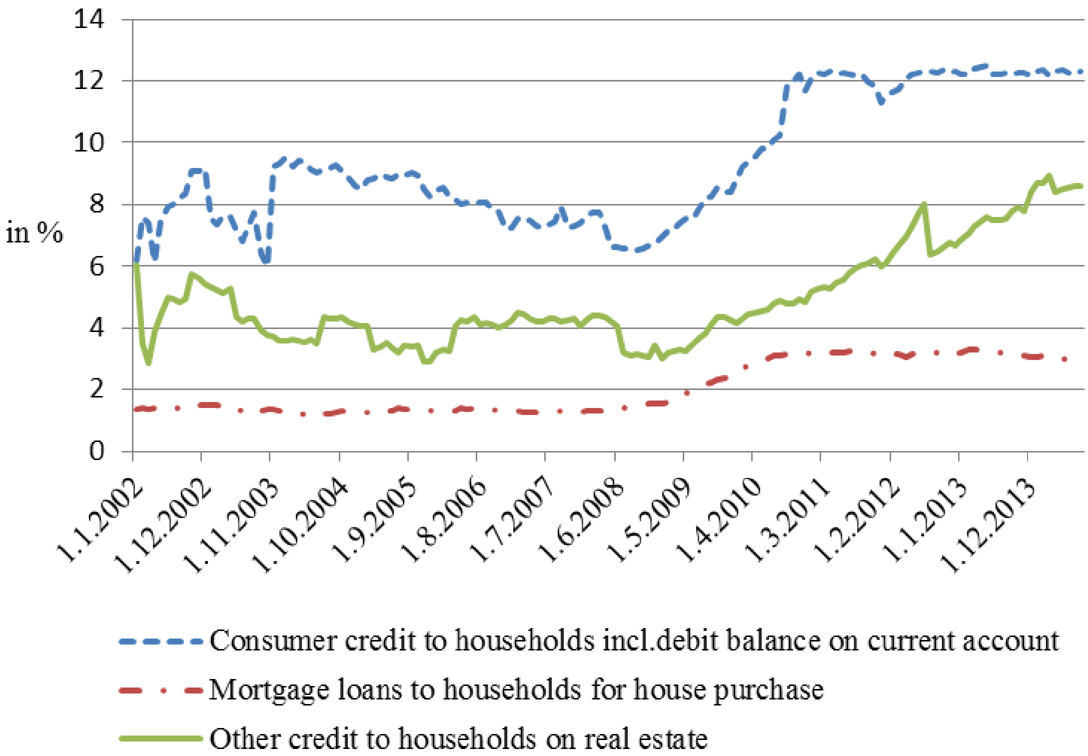
2.Why the Type of Consumer Loan Matters

2.1. Description of Consumer Loan Types

The share of non-performing loans[4] of total loans (hereafter, the "NPL ratio") varies substantially among the consumer loan types. Its significance is illustrated in Figure 1, which depicts the share of NPL in consumer credit[5], mortgage loans and housing loans in the Czech Republic. Whereas mortgage loans maintained a solid performance between 2002 and 2013, the share of problem loans in the case of consumer credit and housing loans sharply increased. Although neither consumer credit nor housing loans are backed by collateral, there is a 3.7 percentage point difference in their NPL ratio (based on the most recent results from August 2014).

[3]The Bank does not wish to be explicitly identified.
[4]According to CNB Regulation No. 123/2007, § 196, § 197 non-performing loans are receivables with default classified as substandard, doubtful or loss loans.
[5]The statistics cover consumer credit with both specified and unspecified loan purpose.

Figure 1. Share of non-performing loans on total loans (by consumer loan type)



Source: Czech National Bank (CNB) – ARAD database – Monetary and financial statistics. *Note:* (1) The statistic covers consumer loan provided in the Czech Republic. (2) Non-performing loans include substandard, doubtful or loss loans.

The loan-application process for consumer credit (with specified and unspecified purpose) and housing loans begins identically. In order to assess the creditworthiness of their potential debtors and to decide whether to grant a loan, financial institutions use automated credit scoring techniques. Their main purpose is to estimate the probability that an applicant will default by a given time in the future. Lenders make loan-granting decisions based on the loan application information provided by their customers and the probability of default. Application information is evaluated by analyzing a sample of past customers who applied for a loan, whose records provide good information on subsequent loan performance history. Credit scoring divides loan applicants into ‘good’ and 'bad' and assists independent lending institutions in their loan-granting decisions. Kocenda and Vojtek (2011) provide an extensive survey of literature on existing credit scoring techniques (e.g. logistic regression, classification and regression trees) and compare their efficiency and discriminatory power.

2.2. The Role of Loan Contract Terms in Alleviating Adverse Selection

Loan contract terms for the individual consumer loan types differ. After mortgage loans (secured by a lien on a property), housing loans offer the second lowest interest rates. To be eligible for a favorable interest rate in the case of housing loans, the applicant must document both loan purpose and proof of homeownership.[6] The loan must finance a property-related investment, and the real estate should be in the name of the applicant. Though housing loans can be used to finance home renovation, home purchase (to some upper limit), or home equipment, they are not secured by a lien on the property.[7] Another advantage is that housing loans are also subject to favorable tax treatment. Upon fulfilling certain conditions, a borrower can deduct the interest expenses of housing loans when tax returns are filed. The interest rates of loans for auto purchase or other purpose are less favorable[8], as the loan is not backed by homeownership. Borrowers are obliged to submit an invoice verifying that the loan was used for the specified purpose, and are then entitled to the lower interest rates. If the

[6]The terms and conditions of housing loans usually include also requirements on the share of total costs (e.g. 20 or 30 percent) to be financed from the borrowers' own resources.

[7]After exceeding some upper loan limit, the bank may insist on securing a loan by collateral. Nevertheless, if the applicant decides to back his loan with property, then the loan application is changed to a mortgage loan request, due to the even lower interest rate it offers. In some cases the bank might require the property to be insured (the cost of insurance is paid by the borrower).

[8]Comparing loans for auto purchase or other purpose, the former offer more favorable loan contract terms. This is because cars are easier for the bank to repossess in case the borrower defaults.

borrower does not deliver this evidence, the price of the loan is raised to the interest rate level of loans without a specified purpose. Loans for unspecified purposes bear the highest interest rate. This is because individuals who cannot or who are not willing to specify the purpose of financing are perceived as risky.

Housing loans thus benefit from the presence of informal collateral.[9] Individuals providing evidence of loan purpose and homeownership can signal their creditworthiness and gain favorable loan contract terms. This can prevent market inefficiencies that arise on consumer loan markets when the borrower has private information related to loan repayment. To mitigate this asymmetric information between lenders and borrowers, the bank can design such loan contract terms (most importantly, set interest rates) that aim to reveal the borrower's risk type. This paper tests the effectiveness of informal collateral to alleviate adverse selection on the housing loan market, a field that has not previously been studied.

3.Methodology

This section outlines the identification strategy applied to measure the impact of consumer loan type on the borrower's default rate. To estimate the impact of loan type on a borrower's default rate, first the simple probit is applied. Compared to the linear probability model, the probit model offers a better modeling of dichotomous outcome estimation. Second, the propensity score matching is used to see how the results change after the potential selection bias on consumer loan market is accounted for. This paper

[9]Pavan (2008) is the first to define the role of durable goods as informal collateral in the loan performance of unsecured debts.

does not model the process of loan approval and the setting of loan contract terms (loan amount, interest rate, maturity).[10] These are assumed to be the result of equilibrium outcome.

3.1. Probit Models

Hypothesis 1. The purpose of the loan has no impact on the probability of loan repayment.

The default rate is a function of information available about the borrower at the time of loan application. In Model 1 the probability of default is estimated by the following probit model:

$$Y_i^* = \beta_0 + X_i^{'}\beta_1 + \beta_2 PURPOSE_i + \varphi_i \tag{1}$$

where Y_i denotes default for borrower $i, X_i^{'}$ is the vector of application characteristics and the loan contract terms of application $i, PURPOSE_i$ is a categorical variable indicating the purpose of a loan (Table A1 in the Appendix summarizes the individual variables and their coding) and φ_i are unobserved factors assumed to have a standard normal distribution with zero mean and variance equal to one. Although the latent variable Y_i^* is not observed, Y_i takes the value of 0 if the borrower does not default (

[10]Kuvikova (2015) estimates loan demand and loan performance jointly, while accounting for the number of successful payments till default using the endogeneity of loan contract terms, the potential sample selection on the consumer loan market. The paper also offers an alternative model for default estimation by utilizing a duration model that takes into account the number of successful payments till default.

$Y_i^* < 0$) and Y_i takes the value of 1 if the borrower defaults ($Y_i^* > 0$). Assuming the standard normal cumulative distribution $\Phi(.)$ the probability of default can then be derived as follows:

$$\Pr(Y_i^* = 1 \mid PURPOSE_i, X_i) = \Phi(\beta_{0i} + X_i^{'}\beta_1 + \beta_2 PURPOSE + \varphi_i) \tag{2}$$

Although the coefficients of the probit model $(\frac{\partial \mathrm{E}(Y_i^*)}{\partial X_k} = \beta_k)$ express the direction of the impact of the explanatory variables on the binary outcome, unlike in the linear probability model they do not express the marginal effects and hence need to be calculated explicitly. To quantify the magnitude of the effect $(\frac{\partial \Pr(Y_i=1|X_i)}{\partial X_k})$ the average marginal effect is used. It expresses the impact of a one-unit change in the explanatory variable on the average change in the probability of the outcome variable.

Specifically, the null hypothesis tested in Model 1 is that $H_0 : \beta_2 = 0$.

Hypothesis 2. The type of applicant choosing loans with different riskiness does not affect the loan default.

Applicants with different default probability might select loans with certain loan purpose. To differentiate between the effect of loan type and the type of individuals that apply for certain loans, Model 2 is defined as

$$Y_i^* = \beta_0 + X_i^{'}\beta_1 + \beta_2 PURPOSE_i + \beta_3 APPTYPE_i + \varphi_i \tag{3}$$

where the categorical variable $APPTYPE_i$ indicates the applicant's type with respect to loan purpose. The variable is created using information on multiple loan contracts per applicant (both accepted and rejected loans), in which to each loan purpose *j* an applicant type dummy is assigned. This dummy takes the value of one if the loan application *i* is submitted by an individual who has already applied for a loan purpose *j*. This specification enables one to account for the unobserved individual heterogeneity connected to the good (certain applicants are more prone to buy more riskier goods) and quantify whether the default is driven by the riskiness of the applicant or the riskiness of the good the loan is intended to finance (Bicakova, 2007).

Specifically, the null hypothesis tested in Model 2 is that $H_0 : \beta_3 = 0$.

3.2. Propensity Score Matching

Hypothesis 3. The average effect of loan purpose on the loan default is not significantly different from zero when similar applicants are compared.

In estimating the effect of loan type on the default rate of borrowers, self-selection becomes an issue. Specifically, borrowers applying for a purpose-loan may differ significantly from those applying for a non-purpose loan. To account for self-selection and check the robustness of results based on probit regression, the matching approach is utilized. The method is used for estimating causal effects, and aims to resemble a randomized experiment by comparing treated and control groups with similar distribution of covariates.[11] Contrary to a standard regression approach that might suffer from selection on unobservable characteristics, matching is a non-experimental method that focuses on controlling for observables. As the method is non-parametric, it does not impose a functional form and requires fewer assumptions than the regression approach.[12]

In order to see whether the default rate of purpose-loans differ from the default rate of non-purpose loans, I take advantage of the non-experimental matching method suggested by Rosenbaum and Rubin (1983). The method allows us to quantify the impact of treatment programs that differ across individuals. In particular, it describes what would have happened in the absence of treatment. The method assumes that the selection of individuals into control and treatment groups is based on a sufficient number of observables, where the unobservables are assumed to be unimportant.

Two potential outcomes of probability of default are compared: y_{1i} is the probability of default for purpose-loans and y_{0i} is the probability of default for non-purpose loans. I assume that a population of borrowers exists in which everyone is equally eligible to choose between the two types of loans. I observe y_{1i} only if $D_i = 1$ (the borrower applied for a purpose-loan) and observe y_{0i} only if $D_i = 0$ (the borrower applied for a non-purpose loan).

[11] Stuart (2010) offers a detailed review of matching techniques.

[12] Angrist (1998) argues that the primary difference between the estimates of the approaches lies in the weights corresponding to the explanatory variables. Whereas in the regression model the weights are larger when the variance of treatment is larger, in the matching approach the weights are larger when the probability of treatment is larger.

Assuming the borrower has a choice between loan types, the aim is to measure whether the purpose makes a difference in the default rate of borrowers. The average effect of treatment on treated (ATT, hereafter) is chosen to quantify the average effect of loan type on the probability of default:

$$E(y_{1i} - y_{0i} \mid D_i = 1) = E(y_{1i} \mid D_i = 1) - E(y_{0i} \mid D_i = 1).$$

If the choice of loan type was completely random, i.e. $E(y_{0i} \mid D_i = 1) = E(y_{1i} \mid D_i = 0)$, we could simply compare the treatment group (the borrower applied for a loan with specified purpose) and control group (the borrower applied for a loan with unspecified purpose) as in a randomized experiment. However, as we deal with a non-randomized observational dataset on application characteristics, the treatment and control groups are not comparable before the treatment. Thus, a non-parametric matching method[13] is required to estimate the average effect of loan type. This reduces the bias caused by confounding factors in observational datasets where the assignment of customers to the treatment and control groups is not random. Controlling for confounding factors, the matching method corrects for the selection bias by balancing the distribution of covariates in the treated and control groups.

As I deal with a large number of application characteristics when testing the null hypothesis, I take advantage of the propensity score matching.[14] This approach groups

[13]The matching estimators can identify and give consistent estimates of the choice of loan type on default rates under the following two assumptions: (1) D_i is independent of (y_{1i}, y_{0i}) conditional on $X = x$. (2) $c < P(D_i = 1 \mid X = x) < 1 - c$, for some $c > 0$. The first assumption (the unconfoundedness assumption) ensures that, conditional on the application characteristics of the borrower, the loan type is independent of the default rate of the borrower. The second assumption (the identification assumption) allows for borrowers of the two loan types to have similar application characteristics and thus they can be compared.

[14]Abadie and Imbens (2002) suggest that a bias of simple matching estimators exists, and the simple method might be not suitable in cases where there is a wide range of covariates.

the pre-treatment characteristics of each individual into a single scalar and the matching is realized solely on this propensity score.[15] The propensity score matching is done by pairing each treated individual with one or more individuals from the control group based on their propensity scores. Motivated by Heckman, Ichimura and Todd (1997), who compare different matching methods depending on sample size, I use the "nearest neighbor" method for ATT estimation.

Specifically, the null hypothesis tested in Model 3 is that $H_0 : E(y_{1i} - y_{0i} \mid D_i = 1) = 0$.

4.Data

In order to analyze the default pattern of the Czech consumer loan market, a dataset of over 207 000 consumer loans covering the entire Czech Republic has been obtained. It consists of application information on those individuals who were granted/rejected a consumer loan between 2007 and 2013, together with their monthly repayment status. The data observation period lasts till 2013. The sample of consumer loans includes only CZK-denominated loans. It is a representative sample, as the vast majority of loans in the Czech Republic are CZK-denominated - the share of loans to households denominated in foreign currency is below 1%.[16] Table A1 in the Appendix lists the available information on consumer loans.

[15] Rosenbaum and Rubin (1983) define the propensity score as the propensity towards exposure to treatment 1 given the observed pre-treatment covariates. In other words, the propensity score is the probability of being granted a purpose loan, conditional on the borrower's application characteristics and the loan contract terms.
[16] Source: Czech National Bank – ARAD databasehttp://www.cnb.cz/cnb/STAT.ARADY_PKG.STROM_SESTAVY?p_strid=AABBAA&p_sestuid=&p_lang=EN

The selection of variables predicting default is driven by the information the Bank includes on their loan application form (the borrower's application characteristics and the loan contract terms). Nevertheless, following Kocenda and Vojtek (2011), I also conduct a single factor analysis to check the discriminatory power of the variables applied in the Bank's credit scoring. The overall information value of the application characteristics is calculated as the sum of information values for each category of application characteristics, defined for loan application i as

$$\text{Information Value}_i = \ln(\text{Odds}_i)\left(\frac{Default_i}{Default} - \frac{NoDefault_i}{NoDefault}\right) \tag{4}$$

$$\text{Odds}_i = \left(\frac{Default_i}{Default}\right)\left(\frac{NoDefault}{NoDefault_i}\right) \tag{5}$$

where *Default* represents the total number of defaulted loans and *NoDefault* represents the total number of loans that were repaid. The information value of application variables summarized in Table 4 confirms that the majority of application characteristics have information value between 0.1 and 0.2. The higher the information value, the higher the discriminatory power of the variable with the given categorization.

4.1. The Expected Impact of Loan Contract and Application Characteristics on Default

Table A1 in the Appendix summarizes the expected impact of loan term characteristics and application characteristics on the probability of loan default based on the related literature. The first set of variables include loan contract terms (Table A1, Panel A), which describe the loan the borrower and lender agreed on. Several application and loan term characteristics might signal a borrower's low probability of default. Recent literature findings suggest that lower default is likely on loans of high amounts (Dobbie and Skiba, 2013), on loans with a specific purpose (Kocenda and Vojtek, 2011), and for loans that were evaluated by applying risk-based pricing[17] (Adams, Einav, and Levin, 2009). A high credit bureau score expresses the applicant's low indebtedness (the score is highest if the borrower has no other debt) and a high behavioral score expresses the applicant's good repayment history (the score is the highest if the borrower has had no problems in previous debt repayment).

The second set of variables contains individual application characteristics (Table A1, Panel B), which represent the socio-demographic characteristics of the potential borrower at the time of loan application. From the application characteristics, the likelihood of bankruptcy is expected to diminish for older (Dobbie and Skiba, 2013), female (Chandler and Ewert, 1976), married and university-graduated applicants (Kocenda and Vojtek, 2011). In addition, previous literature suggests that employment with stable income (Gross and Souleles, 2002), home ownership (Adams et al., 2009) and long employment duration (Kocenda and Vojtek, 2011) should also have a positive impact on loan repayment. Certain application characteristics might be omitted from credit scoring models. Chandler and Ewert (1976) show that if gender is allowed, men have a significantly smaller chance of being granted a loan. This can be because other variables, like low income and part-time employment, signal good repayment behavior in the case of females, but bad repayment behavior in the whole population. In order to

[17]The Bank has been applying risk-based pricing (i.e. pricing based on the borrower's expected riskiness) from January 2012.

estimate the probability of default, this paper uses the list application information (including gender) that the Bank applies in its credit scoring model.

4.2. Descriptive Statistics

The descriptive statistics of the application characteristics and loan contract terms are presented in Table A3 in the Appendix. The mean values of personal loan information suggest that an average borrower has been employed for more than 5 years and has an average net income above 17 000 CZK monthly. On average the applicants were approved for a loan amount of 100 000 CZK with four and half year maturity at an interest rate of 14 percent.

Although there are several different definitions of 'defaulted' loans, similar to the literature on installment loans (Gross and Souleles, 2002; Barron, Chong, and Staten, 2008), I measure loan performance using delinquency rate as a proxy for expected default rate. I consider a loan to be in default if the borrower is more than 30 days overdue on any payment connected with the loan. Later, for the purposes of the sensitivity analysis, I use the definition set by the Basel Committee on Banking Supervision (2004): a loan is considered to be in default if the borrower is more than 90 days overdue on any payment connected with the loan. Table 1 summarizes the default rate of loans by loan type.

Table 1. Default rate by loan type

Loan type	No default	Default	Accepted loans	Acceptance rate
Unspecified purpose	94,5%	5,5%	91 305	50,9%
Specified purpose	98,6%	1,4%	14 454	51,4%
Total	**100 508**	**5 219**	**105 759**	**50,9%**

Source: Author's (2014) computations, data from 2007-2013.

Purpose-loans include loans for home purchase, home renovation, purchase of home equipment, purchase of a new/used car and loans for other purposes (e.g. mobile phone, computers, etc) and represent 14% of the total dataset (accepted and rejected loans). Table 2 presents the default rate and interest rate differentials of accepted loans. Consistent with national statistics, housing loans (loans for home purchase, home renovation, and purchase of home equipment) have lower default rate around a 3% than consumer loans with unspecified purpose. The interest rates reflect how easy it would be for the bank to repossess assets from the borrower in case of default: the cheapest are housing loans (connected to the ownership of property), then loans for auto purchase (connected to ownership of a car), and the least favorable interest rate is for loans with other or unspecified purposes.

Table 2. Default rates and interest rates per loan purpose

Loan purpose	**Default rate**	**Interest rate**	**Accepted loans**	**Acceptance rate**
Unspecified purpose	5,5%	14,0%	91 305	50,9%
Home purchase	2,3%	8,1%	3 171	51,2%
Home renovation	1,2%	8,2%	6 818	51,4%
Home equipment	1,3%	7,8%	477	38,9%
New/used car purchase	1,6%	11,6%	251	48,6%
Other purpose	0,8%	13,4%	3 737	54,2%
Total	**4,9%**	**13,4%**	**105 759**	**50,9%**

Source: Author’s (2014) computations, data from 2007-2013.

5.Results

The results of estimating the effect of application characteristics and loan contract terms on borrower’s probability of default conforms to expectations. The estimation results from the probit model and propensity score matching suggest that the impact of loan purpose on the probability of default rate is significant. Interestingly, the default rate differential between purpose-loans and non-purpose loans is much smaller than the interest rate differential.

5.1. The Effect of Informal Collateral on Loan Default

In order to interpret the effect of the individual loan determinants on the probability of default while keeping all the other covariates constant, I follow Greene (2003) and calculate the marginal effects from the estimation results. Table 3 displays the calculated average marginal effects of the probit model with corresponding standard errors for Model 1 and Model 2. [18]

Table 3 (Panel A) presents the probit estimation results with respect to loan term characteristics. The results from Model 1 indicate that the hypothesis that the purpose of the loan has no impact on the probability of loan repayment can be rejected. In particular, the probability of default decreases with an indicated loan purpose. Applicants with clear intentions and carefully planned objectives default less. Specifically, as a result of financing a home purchase, the borrower's probability of default decreases on average by 3.6 percentage points (compared to a non-purpose loan). The effect of financing home renovation, purchase of home equipment or used car is analogous.

[18] The reference group for the application factor variables is always the one with the lowest coding (summarized in Table A1 in the Appendix) and 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. For example, according to the positive sign of education level, relative to primary education being the reference group, the higher a customer's level of education, the lower the predicted default is expected to be.

Table 3.Probit estimation results (Panel A – Loan term characteristics)

Dependent variable: Default		**Model 1**	**Model 2**
		dy/dx (Delta method - standard error)	
Loan term characteristics	**Risk-based pricing**	-0.043***	-0.043***
		(0.001)	(0.001)
	Approved amount	-0.003***	-0.003***
		(0.001)	(0.001)
	Loan maturity	0.026***	0.026***
		(0.001)	(0.001)
	Loan purpose		
	Home purchase	-0.036***	-0.028***
		(0.002)	(0.005)
	Home renovation	-0.034***	-0.010
		(0.002)	(0.007)
	Home equipment	-0.037***	-0.039***
		(0.006)	(0.006)
	New/used car purchase	-0.036***	0.835
		(0.008)	(17.367)
	Other purpose	-0.028***	0.031**
		(0.004)	(0.012)
	Applicant type		
	Home purchase		-0.014
			(0.009)
	Home renovation		-0.037***
			(0.008)
	Home equipment		0.011
			(0.014)
	New/used car purchase		-0.352
			(9.828)
	Other purpose		-0.062***
			(0.006)
N		105 759	105 759
$\mathbf{R^2}$		0.2079	0.2132
Prob> chi2		0.000	0.000
Loglikelihood ratio (LR) chi2		8 647.4	8 864.8

Source: Author's (2014) computations, data from 2007-2013. *Note:* (1) The estimates denote the calculated average marginal effects for factor levels (dy/dx) expressing the discrete change from the base level. (2) The reference groups for the categorical variables are the following: Loan purpose - Non-purpose loans; Application type – Applicants only requesting non-purpose loans. (3) Only statistically significant results (***, **, and * denote significance at the 1%, 5%, and 10% levels) are presented.

Applicants funding other purposes (e.g. mobile phones, computers, etc.) are also less likely to have repayment difficulties, though the default declines only by 2.8 percentage points on average (compared to a non-purpose loan). This is natural as these applicants most likely finance one-time expenditures that have a relatively short lifespan (unlike investments in real estate). These findings complement the results of Kocenda and Vojtek (2011), who also utilize data from a Czech commercial bank and find that compared to loans for house building, loans with other purposes (e.g. renovation, purchase of apartment, land or house) have higher estimated probability of default. Nevertheless, this paper goes further and aims to compare the default rate and pricing differential of purpose-loans and non-purpose loans after accounting for potential selection bias.

When controlling for unobserved individual heterogeneity, the negative relationship between loan purpose and default probability is altered. The results from Model 2 suggest that the hypothesis that the type of applicant choosing loans with different riskiness does not affect the loan default can be rejected. In particular, after accounting for the applicant's type (*j* dummies created for borrowers who applied for the loan purpose *j* at least once), the effect of the loan purpose diminishes and it is the different type of borrowers with unobserved riskiness that drives the default rate. Compared to non-purpose loans, home renovations default less by 3.7 percentage points solely due to the fact that these borrowers have higher repayment incentives than loans without specific purpose. In the case of applicants financing a home purchase, the effect of loan purpose overweighs the effect of applicant type in explaining the lower default. The lower default rate of home equipment loans is also driven by the loan type. The applicant type has the most extreme impact on loans for other purposes (e.g. loans for mobile phones, computers, etc.): although borrowers of these durable goods default more, it is the applicant's lower riskiness that drives the better loan repayment. These findings are in line with Bicakova (2007), who presents qualitatively similar results on a sample of Italian consumer loans.

The remaining loan contract terms have similar influence on default probability for Model 1 and Model 2. In line with the findings of Dobbie and Skiba (2013), default declines with approved loan amount. This result is surprising given the asymmetric information between lenders and borrowers that stimulates the prominence of moral hazard (i.e. default is more likely on larger loans, while borrowers do not pay for the increased default costs) on the consumer loan markets (Adams et al., 2009). On the other hand, the default increases with longer loan maturity similar to Adams et al. (2009). This is predictable as default is more probable over a longer time period. Interestingly, interest rates turn out to be statistically insignificant. Credit bureau score (indicating the applicant's indebtedness) can also successfully reveal the borrower's riskiness. Both Gross and Souleles (2002) and Barron et al. (2008) confirm that the higher the credit bureau score, the less likely the borrower will default. The behavioral score (indicating the applicant's repayment history) encompasses information about whether the borrower historically accepted/rejected and repaid/defaulted on loans. The higher the score, the better the applicant's credit history and the better his/her future loan repayment behavior. These results follow the findings of Marshall, Tang and Milne (2010) who argue that longer lending relationship improves the quality of loan portfolios.

Application characteristics explain loan performance well and conform to expectations. Panel B of Table 3 indicates that the results are stable across the models.

From the set of variables, monthly income is perceived as a key indicator of a borrower's creditworthiness. With respect to its relationship to loan repayment, it is expected that the higher the applicant's monthly income, the lower the probability s/he will go bankrupt. Similarly to Gross and Souleles (2002), this paper provides empirical evidence that after accounting for other application characteristics, the impact of monthly income on default probability is very low in magnitude. This is also in line with Kocenda and Vojtek (2011), who find that including income in the credit scoring specification improves discrimination between 'good' and 'bad' applicants only marginally. Though Marshall et al. (2010) highlight that students are less likely to default, Model 1 and Model 2 cannot support this finding with statistically significant results. Instead, pensioners have on average (by 2.5 percentage points) lower default rate than employed applicants. The level of education is also a key characteristic that indicates how reliable the borrower will be in repaying the loan. Applicants with only primary school education have the highest probability of default. In line with the results of Kocenda and Vojtek (2011), with every additional level of education the likelihood of loan default declines. Similarly, I also find that lower probability of default is expected for married applicants (due to the assumption that they have an additional source of income in the case of job loss), and borrowers employed for a longer period or employed by a public organization (due to the assumption that they are more risk-averse). The results suggest that borrowers who own real estate are also less likely to default (similar to Adams et al., 2009). Application and loan term characteristics not presented in Table 3 yield statistically insignificant estimation results.

Table 3.Probit estimation results (Panel B – Application characteristics)

Dependent variable: Default		**Model 1**	**Model 2**
		dy/dx (Delta method - standard error)	
Application characteristics	**Behavioral score**	-0.001***	-0.001***
		(0.001)	(0.001)
	Credit bureau score	-0.001***	-0.001***
		(0.001)	(0.001)
	Female	-0.008***	-0.008***
		(0.001)	(0.001)
	Education		
	Secondary (general)	0.029***	0.029***
		(0.006)	(0.006)
	Post-secondary (technical)	-0.020***	-0.020***
		(0.007)	(0.007)
	Secondary (vocational)	-0.011**	-0.010**
		(0.005)	(0.005)
	University	-0.026***	-0.025***
		(0.005)	(0.005)
	Employment status		
	Pensioner	-0.025***	-0.025***
		(0.002)	(0.002)
	Employment duration	-0.001***	-0.001***
		(0.001)	(0.001)
	Employment type		
	Bank/insurance company	-0.041***	-0.040***
		(0.003)	(0.003)
	Private company	-0.014***	-0.013***
		(0.002)	(0.002)
	Public organization	-0.018***	-0.018***
		(0.002)	(0.002)
	Net monthly income	0.001**	0.001**
		(0.001)	(0.001)
	Marital status		
	Married	-0.012**	-0.012**
		(0.006)	(0.006)
	Housing status		
	Living with parents	-0.025***	-0.024***
		(0.004)	(0.004)
	Sharing property	-0.021***	-0.020***
		(0.005)	(0.005)
	Personal property	-0.026***	-0.025***
		(0.004)	(0.004)
	Region	yes	yes
N		105 759	105 759
$\mathbf{R^2}$		0.2079	0.2132
Prob> chi2		0.000	0.000
Loglikelihood ratio (LR) chi2		8 647.4	8 864.8

Source: Author's (2014) computations, data from 2007-2013. *Note:* (1) The estimates denote the calculated average marginal effects for factor levels (dy/dx) expressing the discrete change from the base level. (2) The reference groups for the categorical variables are the following: Education – Secondary (technical); Employment status – Employed; Employment type, Marital status, Housing status – Unspecified by the applicant. (2) Only statistically significant results (***, **, and * denote significance at the 1%, 5%, and 10% levels) are presented.

5.2. Default and Interest Rate Differential between Purpose-Loans and Non-Purpose Loans

To predict the probability that an applicant for a consumer loan will default, lenders need a credit scoring model that captures the behavior of an average applicant. The information most frequently used is the repayment behavior of applicants who were granted a loan; the characteristics of those applicants who were denied a loan is not recorded. Yet estimating the probability of default only on a sample of accepted applicants and then applying it to the sample of all applicants leads to biased estimates of the parameters. This exclusion of rejected applicants then results in an underestimation of the predictive power of the credit scoring model.

In order to ensure that borrowers with the same application characteristics are compared when quantifying the impact of loan purpose on the probability of default, propensity score matching is applied. The ATT is estimated in the following steps:

First, on the sample of consumer loan application data where all individuals have a unique observation, I estimate the propensity score on the individual characteristics by fitting a logistic regression:

$$PURPOSEL_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \tag{6}$$

where $PURPOSEL_i$ is the binary variable taking the value of one for purpose-loans and taking the value of zero for non-purpose loans, X_i is the set of application characteristics and ε_i is the error term. This gives the predicted probability of loan type based on the set of application characteristics as a composite score.

As a second step, I test whether the above specification is applicable. That is, after the propensity score is created, I test for the balancing hypothesis. It says that observations with the same propensity score must have the same distribution of application characteristics independent of loan type. The results of the balancing hypothesis summarized in Table A3 in the Appendix suggest that a significant part of the covariates is well-balanced.

Finally, once the propensity score satisfies the balancing hypothesis, I examine the effect of loan type on default by using propensity score matching. Specifically, I group applicants with similar application characteristics and loan contract terms to show that the variation in default rate remains even after controlling for observable borrower risk. The ATT estimation results using the "nearest neighbor" matching method with bootstrapped standard errors are summarized in Tables 4. The results suggest that the hypothesis that the average effect of loan purpose on the loan default is not significantly different from zero when similar applicants are compared can be rejected. Purpose-loans have a 0.7 percentage point lower default rate when compared to non-purpose loans. The statistically significant result at 1% level is achieved by matching over 14 000 purpose-loans with over 90 000 non-purpose loans (Table A5 in the Appendix). When compared to the unmatched sample results, for the matched sample, the default rate differential between purpose-loans and non-purpose loans decreased by 3.4 percentage points.

Table 4. Default rate differential - ATT estimation results

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Default rate	**Unmatched**	0.0137	0.0550	-0.0413	0.0019	-21.34
	Sample	**Treated**	**Controls**	**Difference**	**Bootstrap Std. Err.**	**z**
	ATT	0.0138	0.0206	-0.0067	0.0023	-2.88

Source: Author's (2014) computations, data from 2007-2013. *Note:* A loan is in default if the borrower is more than 30 days overdue on any payment connected with the loan.

To see the interest rate differential between the two loan types, propensity score matching is conducted on the same observable characteristics and loan contract terms. The results summarized in Table 5 suggest that after controlling for observable characteristics, purpose-loans have 3.6 percentage point higher interest rates than non-purpose loans. The test of the balancing hypothesis (summarized in Table A6 in the Appendix) is favorable and only two observations are off common support (summarized in Table A7 in the Appendix) during the propensity score matching.

Table 5. Interest rate differential - ATT estimation results

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Interest rate	**Unmatched**	9.5712	13.9741	-4.4030	0.0214	-206.12
	Sample	**Treated**	**Controls**	**Difference**	**Bootstrap Std. Err.**	**z**
	ATT	9.5714	13.1382	-3.5668	0.0344	-103.80

Source: Author's (2014) computations, data from 2007-2013.

The high interest rate differential for loans with similar default probability is another evidence of the heterogeneity in pricing policy for different loan types. On the example of the Czech Republic, Horváth and Podpiera (2012) show that the interest rate for consumer loans does not follow the market interest rate as closely as those of other types of loans. Alternatively, the authors suggest that the high interest rate for consumer loans is linked to the high risk margin that financial institutions impose on these loans. This paper goes further and points out that high risk margin can be the result of mispricing or the conservative loan granting strategy of the financial institution. Therefore, the pricing policy of financial institutions should be closely monitored in order to limit subsequent difficulties in consumer loan repayment.

6.Sensitivity Analysis

To test the validity of the identification strategy, the propensity score matching is performed applying an alternative definition of default. In particular, I use the definition set by the Basel Committee on Banking Supervision (2004) and consider a loan to be in default if the borrower is more than 90 days overdue on any payment connected with the loan. Table A8 in the Appendix summarizes the default rate of loans by loan type under the original definition (default occurs after 30 days overdue in payments) and the alternative definition (default occurs after 90 days overdue in payments). By relaxing the definition of default, the sample of loans in default is significantly reduced (from 5 219 to 3 744 observations). More importantly, after the definition change there is a substantial drop in the default rate differential between purpose-loans and non-purpose loans (from 4.1pp to 3.1pp).

Table 6. Sensitivity analysis - ATT estimation results

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Default rate	**Unmatched**	0.0082	0.0397	-0.0315	0.0017	-19.05
	Sample	**Treated**	**Controls**	**Difference**	**Bootstrap Std. Err.**	**z**
	ATT	0.0083	0.0144	-0.0061	0.0019	-3.26

Source: Author's (2014) computations, data from 2007-2013. *Note:* A loan is in default if the borrower is more than 90 days overdue on any payment connected with the loan.

The sensitivity analysis indicates that controlling for observable characteristics, the small difference between the default rate of the two loan types remains. The estimation results with the alternative definition of default are presented in Table 6 - the ATT is equal to 0.6 percentage points and is statistically significant at 1% level.[19]That is, when comparing applicants with same characteristics and loan contract terms, purpose-loans have a default rate only 0.6 percentage points higher than non-purpose loans. Hence, the alternative definition of default confirms the validity of the identification strategy and the robustness of the results.

7.Conclusion

Loans to households constitute the largest part of the loan portfolios of most banks. This paper addresses a primary problem of lending institutions, that is, how to evaluate customers' probability of default prior to granting loans. Utilizing data from a large set of consumer loans from the Czech Republic, the default rates of purpose-loans and non-purpose loans are analyzed and compared.

[19] The detailed results of the propensity score matching are available upon request.

The paper offers several contributions to the current literature on consumer loan market. First, results provide evidence that housing loans default less often. The existence of informal collateral (i.e. evidence of homeownership and invoice about loan purpose) signals better loan repayment. This is in line with theories that consider collateral as a tool to alleviate adverse selection on the consumer loan market. Second, the default rate differentials between consumer loan types are in several cases not driven by the purpose they intend to finance, but the type of borrower. This effect is most significant in the case of loans for home renovation. Third, controlling for observable application characteristics and loan contract terms, the default rate differential between purpose-loans and non-purpose loans decreases, though the interest rates differential between these two types of loans remains substantial. Specifically, while purpose-loans have, on average, only a 0.7pp higher default rate, their interest rate is 3.6pp higher than for non-purpose loans.

These findings provide evidence of the asymmetric information present on the consumer loan market. Borrowers applying for purpose-loans and non-purpose loans have very similar default probability, but are charged substantially different interest rates. This is in line with the empirical literature, according to which, financial institutions are prudent in consumer loan pricing and charge high mark-up when compared to mortgage or corporate loans. Nevertheless, this raises a question whether the interest rate for housing loans (being subject to tax-deductibility) should not be re-evaluated due to their lack of collateralization and higher average amount.

References

Abadie, A., and Imbens, G. (2002). Simple and Bias-Corrected Matching Estimators for Average Treatment Effects. *National Bureau of Economic Research Technical Paper, 283.*

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

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

Angrist, J.D. (1998). Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants, *Econometrica, 66*(2), 249-88.

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

Barron, J.M., Chong, B.-U., and Staten, M.E. (2008). Emergence of Captive Finance Companies and Risk Segmentation in Loan Markets: Theory and Evidence. *Journal of Money, Credit and Banking, 40*(1), 173-192.

Bicakova, A. (2007). Does the Good Matter? Evidence on Moral Hazard and Adverse Selection from Consumer Credit Market. *Giornale degli Economisti e Annali di Economia, 66* (1), 29-66.

Boot, A.W.A., Thakor, A.V., and Udell, G.F. (1991). Secured Lending and Default Risk: Equilibrium Analysis, Policy Implications and Empirical Results. *Economic Journal*, *101,* 458–472.

Dobbie, W., and Skiba, P.M. (2013). Information Asymmetries in Consumer Credit Markets: Evidence from Payday Lending. *American Economic Journal: Applied Economics*, *5*(4), 256–282.

Greene, W.H. (2003). Econometric Analysis, Upper Saddle River, NJ: Prentice Hall.

Gross, D.B., and Souleles, N.S. (2002). An Empirical Analysis of Personal Bankruptcy and Deliquency. *Review of Financial Studies*, *15,* 319-347.

Chandler, G.G., and Ewert, D.C. (1976). Discrimination on Basis of Sex and the Equal Credit Opportunity Act. *Working Paper* 4, Credit Research Center, Purdue University.

Heckman, J.J., H. Ichimura and P. Todd (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*: 64(4), Special Issue: Evaluation of Training and Other Social Programmes, 605-654.

Horváth, R., and Podpiera, A. (2012). Heterogeneity in Bank Pricing Policies: The Czech Evidence. *Economic Systems, 36,* 87-108.

Inderst, R., and Mueller, H.M. (2007). A Lender-Based Theory of Collateral. *Journal of Financial Economics, 84,* 826–859.

Jimenez, G., Salas, V., and Saurina, J. (2006). Determinants of Collateral. *Journal of Financial Economics, 81*, 255–281.

Kocenda, E., and Vojtek, M. (2011). Default Predictors and Credit Scoring Models for Retail Banking. *Emerging Markets Finance and Trade*, *47*(6), 80-98.

Kuvikova, G. (2015). Does Loan Maturity Matter in Risk-Based Pricing? Evidence from Consumer Loan Data. *CERGE-EI Working Paper 538.*

Manove, M., and Padilla, A.J. (2001). Collateral Versus Project Screening: A Model of Lazy Banks. *RAND Journal of Economics, 32*(4), 726–744.

Marshall, A., Tang, L., and Milne, A. (2010). Variable Reduction, Sample Selection and Bank Retail Credit Scoring. *Journal of Empirical Finance*, *17,* 501-512.

Rosenbaum, R., and D.B. Rubin (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70*(1), 41-55.

Rubin, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, *66,* 688–701.

Stuart, E.A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science*, *25*(1), 1–21.

Appendix

Table A1. The list of personal loan information (Panel A – Loan term characteristics)

Variable description	Variable name in dataset	Encoding	Expected effect on default	Recent literature
Loan term characteristics				
Loan approval indicator	***APPROVED***	***dummy***		
Approved amount (in CZK)	***AAMOUNT***	***Continuous***	-	**Dobbie and Skiba (2013)**
Interest rate (in %)	***IR***	***continuous***	+	**Adams et al. (2009)**
Approved loan maturity (in months)	***AMATURITY***	***continuous***	+	**Adams et al. (2009)**
Risk band	***NRISK***			
Very low-risk		*1*		
Low-risk		*2*		
High-risk		*3*		
Very high-risk		*4*		
Credit bureau information	***CBINFO***	***dummy***		
Purpose-loan	***PURPOSEL***	***dummy***	-	**Kocenda and Vojtek (2011)**
Loan purpose	***PURPOSE***			
Non-purpose loan		***1***		
Home purchase		***2***		
Home renovation		***3***		
Home equipment		***4***		
New/used car purchase		***5***		
Other purpose		***6***		
Risk-based pricing	***RBPRICING***	***dummy***	-	**Adams et al. (2009)**

Source: Random sample of consumer loans from the Bank. Author's (2014) computations, data from 2007-2013.

Table A1. The list of personal loan information (Panel B – Application characteristics)

Variable description	Variable name in dataset	Encoding	Expected effect on default	Recent literature
Application characteristics				
Age (in months)	***AGE***	***continuous***	**-**	**Dobbie and Skiba (2013)**
Female	***FEMALE***	***dummy***	**-**	**Chandler and Ewert (1976)**
Marital status	***MARITS***			
Unspecified		*1*		
Divorced		*2*	+	**Barron et al. (2008)**
Married		*3*	-	**Kocenda and Vojtek (2011)**
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*	-	**Kocenda and Vojtek (2011)**
Housing status	***HOUSE***			
Unspecified		*1*		
Living with parents		*2*		
Sharing property		*3*		
Personal property		*4*	-	**Adams et al. (2009)**
Renting		*5*		
Student dormitory		*6*		
Employment status	***EMPLOYS***			
Employed		*1*		
House wife		*2*		
Pensioner		*3*		
Student		*4*	-	**Marshall et al. (2010)**
Employment duration (in months)	***EMPLOYY***	***continuous***	-	**Kocenda and Vojtek (2011)**
Employment type	***EMPLOYT***			
Unspecified		*1*		
Bank/insurance company		*2*		
Enterpreneur		*3*	+	**Marshall et al. (2010)**
Foreign company		*4*		
Private company		*5*		
Public organization		*6*	-	**Kocenda and Vojtek (2011)**
Net monthly income (in CZK)	***INCOME***	***continuous***	-	**Gross and Souleles (2002)**
Region (NUTS2)	***REGION***	***dummy***		
Credit bureau score	***CBSCORE***	***continuous***	-	**Barron et al. (2008)**
Application score	***APPSCORE***	***continuous***		
Behavioral score	***BEHAVSCORE***	***continuous***	-	**Marshall et al. (2010)**

Source: Random sample of consumer loans from the Bank. Author's (2014) computations, data from 2007-2013.

Table A2. Information value of application characteristics

Variable	No default	Default	Total	Odds	Information value
Education					**0.2**
Secondary (technical)	1 108	108	1 216	2	
Secondary (general)	6 775	690	7 465	2	
Post-secondary (technical)	1 713	48	1 761	1	
Secondary (vocational)	43 951	1 813	45 764	1	
Post-secondary (vocational)	36 512	2 371	38 883	1	
University	10 480	190	10 670	0	
Employment type					**0.3**
Unspecified	45 331	3 222	48 553	1	
Bank/insurance company	2 142	20	2 162	0	
Enterpreneur	2 246	201	2 447	2	
Foreign company	3 029	380	3 409	2	
Private company	28 023	912	28 935	1	
Public organization	19 768	485	20 253	0	
Marital status					**0.1**
Unspecified	896	76	972	2	
Divorced	17 638	943	18 581	1	
Married	45 385	1 672	47 057	1	
Partner	905	63	968	1	
Single	32 304	2 342	34 646	1	
Widow/er	3 411	124	3 535	1	
Gender					**0.0**
Male	53 205	3 269	56 474	1	
Female	47 334	1 951	49 285	1	
Housing status					**0.2**
Unspecified	2 685	245	2 930	2	
Living with parents	15 922	1 121	17 043	1	
Sharing property	3 333	248	3 581	1	
Personal property	59 617	1 892	61 509	1	
Renting	18 976	1 712	20 688	2	
Student dormitory	6	2	8	6	
Employment status					**0.0**
Employed	86 999	4 654	91 653	1	
House wife	1 747	122	1 869	1	
Pensioner	11 698	437	12 135	1	
Student	95	7	102	1	
Loan purpose					**0.2**
Non-purpose loan	86 283	5 022	91 305	1	
Home purchase	3 098	73	3 171	0	
Home renovation	6 734	84	6 818	0	
Home equipment	471	6	477	0	
New/used car purchase	247	4	251	0	
Other purpose	3 706	31	3 737	0	

Source: Author's calculations.

Table A3. Descriptive statistics (Panel A – Loan term characteristics)

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
Default indicator	0,04	0,19	0	1
Purpose-loan	0,102	0,303	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 A3. Descriptive statistics (Panel B – Application characteristics)

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
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 A4. Balancing hypothesis – Default rate estimation

Application and loan term characteristics	Mean			t-test	
	Treated	Control	%bias	t	p>|t|
Risk-based pricing	0.342	0.569	-51.4	-39.45	0.000
Behavioral score	476.66	516.14	-20.0	-17.66	0.000
Credit bureau score	385.14	513.48	-47.1	-39.92	0.000
Interest rate	9.642	10.344	-26.8	-23.78	0.000
Loan maturity	2.736	2.739	-0.5	-0.50	0.615
Approved amount (in CZK)	2.713	2.731	-2.0	-2.24	0.025
Age	485.14	507.5	-16.1	-14.74	0.000
Female	0.435	0.462	-5.5	-4.59	0.000
Secondary (general)	0.037	0.036	0.2	0.19	0.849
Post-secondary (technical)	0.020	0.016	3.5	2.96	0.003
Secondary (vocational)	0.471	0.471	-0.1	-0.08	0.934
Post-secondary (vocational)	0.293	0.300	-1.4	-1.24	0.216
University	0.170	0.173	-0.8	-0.60	0.549
House wife	0.018	0.016	1.5	1.29	0.196
Pensioner	0.063	0.081	-6.4	-6.03	0.000
Student	0.001	0.001	0.8	1.00	0.317
Employment duration (in months)	82.11	88.965	-7.6	-6.12	0.000
Bank/insurance company	0.029	0.020	5.5	4.58	0.000
Enterpreneur	0.023	0.016	4.3	3.97	0.000
Foreign company	0.022	0.011	6.5	7.06	0.000
Private company	0.331	0.487	-34.2	-27.04	0.000
Public organization	0.215	0.182	8.3	6.99	0.000
Net monthly income	22585	24360	-12.4	-9.26	0.000
Divorced	0.171	0.201	-8.0	-6.57	0.000
Married	0.522	0.537	-2.9	-2.44	0.015
Partner	0.011	0.011	0.1	0.11	0.910
Single	0.266	0.225	9.0	8.01	0.000
Widow/er	0.020	0.023	-2.1	-1.96	0.050
Living with parents	0.114	0.101	3.9	3.66	0.000
Sharing property	0.028	0.009	10.6	11.53	0.000
Personal property	0.690	0.747	-11.8	-10.63	0.000
Renting	0.148	0.125	6.0	5.61	0.000

Summary of the distribution of |bias|					
	Pseudo R2	LR chi2	p>chi2	MeanB	MedB
	0.089	3488.96	0.000	8.7	5.5

Source: Author's (2014) computations, data from 2007-2013. *Note*: "Treated" and "Control" stands for purpose-loans and non-purpose loans, respectively.

Table A5. Common support – Default rate estimation

Treatment assigment	Common support		Total
	Off support	**On support**	
Untreated	0	91,297	91,297
Treated	294	14,160	14,454
Total	**294**	**105,457**	**105,751**

Source: Author's (2014) computations, data from 2007-2013. *Note*: "Treated" and "Control" stands for purpose-loans and non-purpose loans, respectively.

Table A6. Balancing hypothesis – Interest rate estimation

Application and loan term characteristics	Mean			t-test	
	Treated	**Control**	**%bias**	**t**	**p>|t|**
Risk-based pricing applied	0.342	0.336	1.2	0.98	0.326
Behavioral score	474.34	478.92	-2.3	-2.00	0.045
Credit bureau score	384.51	381.06	1.3	1.07	0.286
Loan maturity	2.740	2.735	0.8	0.87	0.387
Approved amount (in CZK)	2.717	2.725	-0.9	-1.09	0.276
Age	486.52	488.2	-1.2	-1.11	0.269
Female	0.436	0.444	-1.5	-1.32	0.188
Secondary (general)	0.037	0.037	0.1	0.06	0.950
Post-secondary (technical)	0.020	0.020	0.2	0.17	0.866
Secondary (vocational)	0.470	0.466	0.8	0.71	0.479
Post-secondary (vocational)	0.292	0.301	-1.8	-1.58	0.113
University	0.171	0.167	1.4	1.05	0.293
House wife	0.018	0.017	0.6	0.54	0.589
Pensioner	0.064	0.062	0.7	0.75	0.453
Student	0.001	0.001	0.0	-0.00	1.000
Employment duration (in months)	82.267	84.032	-2.0	-1.62	0.106
Bank/insurance company	0.028	0.028	0.0	0.04	0.972
Enterpreneur	0.023	0.021	1.3	1.13	0.259
Foreign company	0.022	0.021	0.5	0.49	0.624
Private company	0.330	0.333	-0.5	-0.41	0.680
Public organization	0.216	0.219	-0.8	-0.66	0.512
Net monthly income	22467	22253	1.5	1.15	0.250
Divorced	0.171	0.166	1.2	1.05	0.293
Married	0.525	0.529	-0.8	-0.71	0.480
Partner	0.011	0.012	-0.5	-0.38	0.702
Single	0.263	0.258	1.0	0.88	0.377
Widow/er	0.020	0.022	-1.3	-1.26	0.206
Living with parents	0.113	0.112	0.2	0.22	0.823
Sharing property	0.028	0.027	0.5	0.47	0.638
Personal property	0.693	0.690	0.5	0.45	0.656
Renting	0.147	0.149	-0.7	-0.66	0.508

Summary of the distribution of |bias|	Pseudo R2	LR chi2	p>chi2	MeanB	MedB
	0.001	39.89	0.475	0.9	0.8

Source: Author’s (2014) computations, data from 2007-2013. *Note*: “Treated” and “Control” stands for purpose-loans and non-purpose loans, respectively.

Table A7. Common support – Interest rate estimation

Treatment assigment	Common support		Total
	Off support	On support	
Untreated	0	91,297	91,297
Treated	2	14,452	14,454
Total	**2**	**105,749**	**105,751**

Source: Author's (2014) computations, data from 2007-2013. *Note*: "Treated" and "Control" stands for purpose-loans and non-purpose loans, respectively.

Table A8. Default rate by loan type

	Default = 30 days overdue		Default = 90 days overdue	
Loan type	**No default**	**Default**	**No default**	**Default**
Unspecified purpose	94.5%	5.5%	96.0%	4.0%
Specified purpose	98.6%	1.4%	99.2%	0.8%
Total	**100 508**	**5 219**	**101 983**	**3 744**

Source: Author's (2014) computations, data from 2007-2013. *Note*: "Treated" and "Control" stands for purpose-loans and non-purpose loans, 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 978-80-7343-346-8 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-338-2 (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