DOCUMENTOS DE TRABAJO

Adverse selection, loan access and default in the Chilean consumer debt market

Carlos Madeira

N° 838 Septiembre 2019 BANCO CENTRAL DE CHILE







CENTRAL BANK OF CHILE

La serie Documentos de Trabajo es una publicación del Banco Central de Chile que divulga los trabajos de investigación económica realizados por profesionales de esta institución o encargados por ella a terceros. El objetivo de la serie es aportar al debate temas relevantes y presentar nuevos enfoques en el análisis de los mismos. La difusión de los Documentos de Trabajo sólo intenta facilitar el intercambio de ideas y dar a conocer investigaciones, con carácter preliminar, para su discusión y comentarios.

La publicación de los Documentos de Trabajo no está sujeta a la aprobación previa de los miembros del Consejo del Banco Central de Chile. Tanto el contenido de los Documentos de Trabajo como también los análisis y conclusiones que de ellos se deriven, son de exclusiva responsabilidad de su o sus autores y no reflejan necesariamente la opinión del Banco Central de Chile o de sus Consejeros.

The Working Papers series of the Central Bank of Chile disseminates economic research conducted by Central Bank staff or third parties under the sponsorship of the Bank. The purpose of the series is to contribute to the discussion of relevant issues and develop new analytical or empirical approaches in their analyses. The only aim of the Working Papers is to disseminate preliminary research for its discussion and comments.

Publication of Working Papers is not subject to previous approval by the members of the Board of the Central Bank. The views and conclusions presented in the papers are exclusively those of the author(s) and do not necessarily reflect the position of the Central Bank of Chile or of the Board members.

Documentos de Trabajo del Banco Central de Chile Working Papers of the Central Bank of Chile Agustinas 1180, Santiago, Chile Teléfono: (56-2) 3882475; Fax: (56-2) 3882231 Documento de Trabajo Nº 838 Working Paper N° 838

Adverse selection, loan access and default in the Chilean consumer debt market*

Carlos Madeira Central Bank of Chile

Abstract

Using household survey data I estimate a model of consumer loan access, lender choice and default behavior. Households are sorted into different lenders, with higher income and education being associated with bank lending. Families with no demand for debt have low unemployment and wage risk. Also, debt amounts are quadratic in age and increase with unemployment risk, wage volatility and household size, which may be motivated by consumption smoothing of life-cycle and income shocks. Default behavior decreases with income and increases with higher indebtedness, health shocks, unemployment and wage risk, confirming there is imperfect screening of individual risk.

Resumen

Utilizando datos de la Encuesta Financiera de Hogares (EFH) estimo un modelo de acceso a créditos de consumo, selección de acreedor y comportamiento de impago. La evidencia muestra que los hogares se agrupan entre distintos tipos de acreedores, con mayor educación e ingreso estando asociados a crédito bancario. Familias sin demanda por deuda presentan bajo riesgo de desempleo y de volatilidad del ingreso laboral. Además, el monto de deuda es cuadrático en edad y aumenta con riesgo de desempleo, volatilidad del ingreso laboral y tamaño del hogar, lo que puede ser motivado por suavización de consumo debido al ciclo de vida y choques de ingreso. El comportamiento de impago disminuye con mayor ingreso y aumenta con mayor endeudamiento, choques de salud, riesgo de desempleo y volatilidad del ingreso laboral, confirmando que el mecanismo de monitoreo de los riesgos de los deudores por parte de los acreedores es imperfecto.

^{*} madeira@bcentral.cl. I would like to thank seminar participants at the Central Bank of Chile, FRB of New York, FRB of Philadelphia, Deustche Bundesbank, U. Carlos III, U. of York, Bank of Spain and Bank of England. All errors are my own.

1 Introduction

Household debt increased consistently in the last decades, both in emerging economies (IMF, 2006) and developed countries (Girouard, Kennedy, André, 2007). This evolution in the quantity of household credit coincided with a period of strong financial innovation, with a great range of loan products being available to consumers. Many consumers are able to access credit from a variety of sources, such as credit cards, auto loans, education loans, and for motives as diverse as health, vacations, purchase of durable goods, or a renegotiation of previous debts. Also, the technological evolution has allowed lenders to process larger and better databases on the characteristics of debtors, allowing for an increased use of credit scoring and an heterogeneity of loan terms for each applicant (Roszbach, 2004, Edelberg, 2006, Einav, Jenkins and Levin, 2012). Yet despite an increasing availability of consumer credit, several families are still unable to access credit markets or obtain lower loan amounts than desired (Adams, Einav and Levin, 2009). The factors behind consumers' loan choice and their credit constraints have been documented in recent studies for the United States (see Dynan and Kohn, 2007, Attanasio, Goldberg and Kyriazidou, 2008, Adams, Einav and Levin, 2009). However, loan choice in developing countries is still understudied.

This paper studies the access to consumer credit, choice of lender and repayment behavior of families in Chile. Consumer loans are particularly relevant in Chile, since over 60% of the households have some consumer debt. Using data from the Chilean Household Finance Survey (EFH), I estimate an econometric model in which families choose among a variety of lender types according to their earnings, labor risk, demographics, and unobserved factors. I find that families are sorted into different lenders according to their observable labor market risk. Furthermore, households' debt levels, income, and labor market risk have a significant impact on default behavior. According to the theory of contract pricing and adverse selection (Adams, Einav and Levin, 2009), lenders should offer higher loan amounts to households that have observables that represent a safer risk. However, in the presence of unobservable risks known to borrowers but not to lenders, then we should expect that riskier borrowers will borrow more and be more indebted. Therefore evidence that higher indebtedness is correlated with default implies that Chilean lenders are unable to discriminate for all the risk characteristics of the borrowers (Einav, Jenkins and Levin, 2012).

In Chile the market for consumer loans has several different providers and their credit offers

represent imperfect substitutes for consumers. These loan providers access different customer lists and information, besides being subject to different legal regulations, which affects their loan terms and the ability to target specific markets (see Marinovic, Matus, Flores and Silva, 2011, for a review of the structure and legal framework of different credit providers). Differences across lenders - such as screening costs, asymmetric information, brand preferences, marketing, search and travel costs, or tied products - can create substantial frictions for customers (Nevo, 2011), which leads us to model loan choice and repayment as a differentiated product choice.

The model of loan choice and repayment behavior has three main components: i) a categorical choice between having no debt, wanting debt but being credit constrained, and five different types of lenders, ii) the choice of loan amount, and iii) a categorical outcome of whether the household defaulted or not on at least one payment over the previous year. The five lender types in this categorical model correspond to: Banks, Banks and Retail Stores, Retail Stores, Union Credit (i.e., loans provided by credit and labor unions), and Other Loans (which includes auto loans, educational debt, pawn shops and some informal lending). Banks and Retail Stores are the two major lenders in Chile, therefore using both lenders is treated as a separate choice than the option of using just one type of lender. Other types of lenders represent a small proportion of the population and therefore I do not model the interaction of those lenders with other types of debts. Furthermore, there are two options for the families that do not report a consumer loan, which are "No Access to Debt" and "No wish to apply for Consumer Loans". The option of "No Access to Debt" represents families with credit constraints. These are families who applied for consumer loans but were denied credit, plus those who wished to apply for credit but did not do so because they expected to be refused. "No wish for Loans" represents the outside option for all agents, comprising the families who report not having consumer debt and no interest in applying for loans.

All three endogenous variables - the choice of type of lender, loan amount, and repayment outcome - are affected by observable and unobservable factors. The observables include income, education, labor income risk (measured by unemployment risk, wage volatility and the replacement ratio of income during unemployment), and demographic characteristics such as the age of the household head and household size. Also, the choice of lender is affected by the motives behind the indebtedness and unobservable factors of the household for each lender type. Unobservable factors can be thought as a mix of different effects, including idiosyncratic tastes of households or bargaining conditions with random-effects that denote "tastes" for each lender and a propensity to default, plus idiosyncratic terms for each agent over time. Assuming a parametric distribution for the unobservables, the model is estimated by Simulated Maximum Likelihood (Train, 2009).

The results show that households with different characteristics tend to sort themselves among different lenders. Households with No desire for Consumer Loans have the lowest wage volatility and the lowest unemployment and job separation rates among all groups. This result confirms that consumer debt is related to smoothing income shocks, therefore households with few income shocks have low demand for consumer debt. Banks are the institution that applies credit scoring and customer specific interest rates on a wider basis, therefore Banks capture the households of highest income and with the lowest unemployment rates among loan applicants. Also, households with loans in Banks suffer the lowest income falls during unemployment. Households with loans in Bank plus Retail and Other debts have both the largest loans in absolute amount and the larger ratios of debt relative to income. Households with No Access to Debt have the lowest income levels and also suffer the strongest income falls during unemployment spells.

Unemployment risk increases the probability of households opting for all loans, but its impact is highest for clients of Retail Stores, Union Credit and Other Loans. Wage volatility is strongly associated with opting for Union Credit, Other Loans and No Access to Debt. Loan amounts increase with income, unemployment risk and wage volatility, therefore it is possible that households use consumer loans to smooth income shocks. The default probability decreases with income and increases with high levels of debt amount and debt service (debt service includes both amortization and interest payments) relative to income, unemployment risk and wage volatility. Bank debtors have a significantly lower probability of default even after controlling for observable variables. Since banks resort more to credit scoring and risk-adjusted interest rates, then one should expect banks to capture the customers with lowest risk (Edelberg, 2006), confirming the economic theory of lender equilibrium with adverse selection (Jaffee and Russell, 1976, Jaffee and Stiglitz, 1990). Health needs are positively associated with default behavior, confirming the predictions of economic models for health expenses that are unpredictable and uninsurable for households (Chatterjee et al., 2007).

Finally, the probability of having a loan and the choice of loan amount is increasing in the number of household members and quadratic in age, first increasing with the age of the household head and then falling in its later years. Therefore the demand for consumer debt has an age profile that resembles the findings of life-cycle consumption in the literature (Attanasio and Weber, 2010). In terms of unobservable factors, I find that households with higher income and education are less heterogeneous in their tastes, and that their choice of loan amount is less persistent over time.

This paper is related to a recent and growing literature of empirical models of loan choice and default behavior which measures the impact of observable risk factors and adverse selection (Roszbach, 2004, Adams, Einav and Levin, 2009, Einav, Jenkins and Levin, 2012). It extends that literature in three ways: i) it applies a similar framework for loan choice and default to a developing economy such as Chile, ii) it introduces a wider range of loan options and unobserved preferences by using tools from the applied product choice models in the field of industrial organization (Train, 2009, Nevo, 2000, 2011), and iii) it uses a more diverse characterization of labor income risk by separating overall risk into different variables such as unemployment risk and wage volatility.

This paper is organized as follows. Section 2 describes the consumer credit environment in Chile and the applied model of loan choice and default. Section 3 summarizes the Chilean Household Finance Survey dataset (2007-2011) and the main characteristics of Chilean families. Section 4 describes the sorting of households across different types of lenders in terms of loan amount, income and labor market risk. Section 5 presents the results of the joint model of lender choice, loan amount and default. Finally, section 6 concludes with implications for policy and future research.

2 Credit environment and empirical model of consumer behavior

2.1 The structure of consumer loan providers in Chile

This section starts with a review of the structure of Chilean credit markets and the differences among lenders, whether caused by differentiated product lines or by legal regulations (see Marinovic, Matus, Flores and Silva, 2011, for a review). In Chile all lenders have public access to a commercial registry of debtors who defaulted on payments¹, however this registry is limited only to negative events and therefore lenders' information sets on the positive characteristics of loan applicants differ substantially, implying agents' have different relationships with each lender (Jaffee and Stiglitz, 1990). Banks represent one type of consumer loan provider in Chile, as well as in other countries.

¹See www.dicom.cl/.

Chilean banks have access to a common credit registry with information on all loan amounts and debt default within the banking system², but they do not observe loans from non-banking institutions. Banks also make a strong use of credit scoring, according to agents' history of credit and other financial products such as direct deposit of wages, automatic bill payment or mortgages.

Retail stores are another kind of credit provider, with a strong brand image and their own credit cards³, and which have access to their own private databases of customers. Retailers provide few cash advances, but their credit cards are widely accepted in several stores, including by utility companies and several merchants. Another type of loan providers are credit unions (denoted as Savings and Loans' Cooperatives⁴) and labor unions (denoted as Family Compensation Funds⁵) which are regulated as providers of "social credit". By legislation all Chilean companies must register their workers in one among several Family Compensation Funds, which provide social credit and other services to their affiliates. These labor unions or Family Compensation Funds represent 67.6% of the aggregate "social credit". Family Compensation Funds are chosen by each employer for all its workers and therefore workers do not choose their institution directly. Union credit providers must offer the same conditions to all of their affiliates, therefore they can change interest rates according to loan size and maturity, but are unable to discriminate against characteristics of the debtors such as their income. Also, Family Compensation Funds benefit from being able to deduct loan payments directly from their clients' wage payroll and therefore face little risk of default. Even in the case of a debtor losing its job, Family Funds may deduct a substantial payment from the workers' severance pay and therefore their risk is limited even in the face of unexpected unemployment events. Finally, there are lenders with more specific goals, such as auto loans at car dealers, education loans, pawn shops⁶, and consumer loans provided by insurance companies⁷.

⁶See www.dicrep.cl.

²See the General Law of Banks of the Chilean Superintendency of Banks and Financial Institutions, www.sbif.cl.

³The norms for non-banking credit card providers are detailed in the Chapter III.J.1 of the Compendium of Financial Norms of the Central Bank of Chile.

⁴See the Chilean Government Department of Cooperatives, www.decoop.cl, the General Law of Cooperatives, DFL 5 (2003), www.bcn.cl, and Chapter III.C.2 of the Compendium of Financial Norms of the Central Bank of Chile.

⁵ These institutions are regulated by the Chilean Superintendency of Social Security. Each Family Compensation

Fund is associated with one of the five labor unions registered at the Confederation of Production and Trade. See the General Statute of Family Compensation Funds, articles 29 to 31 of the Law N°18.833 of 1989.

⁷The regulation of credit by insurance companies is detailed in several norms of the Chilean Superintendency of Assets and Insurance, such as norms NCG 152 of 2002, NCG 208 of 2007 and NCG 247 of 2009.

In Chile, during the year of 2006, banks represented 62.1% of the total amount of consumer credit, while social credit institutions represented 13.1% and retail stores 24.8% of the market, respectively.⁸ However, market presence in terms of customers differs from the aggregate loan amounts, since there are 3.5 million debtors with banking loans, while social institutions and retail stores reach around 2.5 million and 7 million customers, respectively. Therefore retail stores are actually the largest provider of small consumer loans and reach the widest number of customers. Over the last half-decade the market size of each type of lender has differed substantially. The aggregate amount of consumer loans in banks at the end of 2013 was 233% as large as at the beginning of 2006 (Banco Central de Chile, 2013). Aggregate consumer credit by social institutions was 245% as large in 2013 as in 2006, but retail store credit grew only 57% in the same period.

2.2 An empirical model of choice of lender, loan amount and debt default

Economic theory predicts that better loan conditions (such as larger loan amounts, longer maturities for payment, and lower interest rates) are associated with observable characteristics of lower risk (such as higher income and more secure jobs). Some types of high risk agents will not be a profitable loan market segment and will be credit constrained, due to either legal restrictions (such as usury laws and interest rate ceilings) or fixed costs for loan evaluation (Jaffee and Stiglitz, 1990). However, unobservable risk characteristics (such as a taste for higher loan amounts) may create adverse selection and will be associated with larger loan demand and more frequent default. In summary, according to the economic theory of loan markets we should expect three results (Edelberg, 2004, Einav, Jenkins and Levin, 2012): i) lenders will offer better and larger loans to agents with observable characteristics of low risk, ii) unobservable characteristics of high risk will still be associated with both larger loan amounts and default, and iii) agents with very high risk will be credit constrained and without access to loans.

The consumer choice model considers three endogenous variables: i) a categorical choice between having no debt, wanting debt but being credit constrained, and five different types of loans, ii) the choice of loan amount, and iii) a categorical outcome of whether the household defaulted or not

⁸The aggregate amount of other loans (such as automotive and informal lending) is not entirely known, since credits of smaller and unregulated institutions do not need to be registered for statistical purposes.

on at least one payment over the previous year. The five lender types in this categorical model correspond to the major loan providers described in the previous section: Banks, Banks and Retail Stores, Retail Stores, Union Credit, and Other Loans (which includes mainly auto loans, educational debt, plus pawn shops and informal lending). It is possible that some consumers have more than one debt type, say debt at Banks and Other Loans (for example, an educational loan), but except for retail store credit (which reaches around 7 million people in Chile) there are few observations with such interactions. For simplicity, I classify the observed lender choice of each household as the one corresponding to the largest loan amount reported by each family. Banks and Retail Stores are the two major lenders in Chile, therefore using both lenders is treated as a separate choice when the household has a positive amount of loans with both lenders.

Families with no consumer loans are classified in two categories: "No Access to Debt" and "No wish to apply for Consumer Debt". "No Access to Debt" represents families with credit constraints, including those who applied for credit but were denied and the ones who did not apply for credit because they expected to be refused. "No wish for Debt" is the outside option for all agents, comprising the families who report no consumer debt and no interest in applying for loans. To be succinct, these options from now on will be denoted simply as "No Access" and "No Debt".

The modelling of a multivariate choice model with several options and many periods incurs into a problem of multidimensionality, since with P possible products there are P^T possible choices in a panel of T periods (Nevo, 2011). Therefore it is useful to apply a parsimonious model that can summarize the choice among the different options in terms of a restricted number of observable and unobservable factors. This is done in terms of a fully specified maximum likelihood model.

Let $U_{i,b,t}$ denote the utility of household *i* from the option *b* in period *t*, with $b \in \{1 \text{ "Bank"}, 2 \text{ "Bank & Retail"}, 3 \text{ "Retail"}, 4 \text{ "Union"}, 5 "Other Loans", 6 "No Access"}. Furthermore, let$ $us standardize the utility of the outside option, "No wish for Debt", as zero, <math>U_{i,0,t} = 0$. This standardization is made without any loss of generality, since all that matters for the agents' choice is the difference in utility from each option relatively to the outside option (Nevo, 2000). Consumer chooses the option $Y_{i,t} = b$ of highest utility $(\max(U_{i,0,t}, U_{i,1,t}, ..., U_{i,B,t}))$ and then a loan-amount $L_{i,t}$, which are affected by observable characteristics, $x_{i,t}$, plus unobservable preferences for each loan type b, $\varepsilon_{i,b,t}$, and loan-amount, $\zeta_{i,t}$. For simplicity, let us assume the utility of each loan type is both an additive and linear function of the observables and the error term:

1)
$$U_{i,b,t} = \alpha_{b,t} + \beta_b x_{i,t} + \varepsilon_{i,b,t}$$

At the same time as the consumer decides its loan type (options 1 to 5) instead of either "No wish for debt" (option 0) or "No Access to Debt" (option 6), he chooses a log-loan amount which is an additive and linear function of the observable factors, $x_{i,t}$, plus an unobservable preference $\zeta_{i,t}$:

2)
$$\ln(L_{i,t}) = \pi_t + \delta x_{i,t} + \zeta_{i,t}$$
.

The decision of defaulting at time $t, D_{i,t} \in \{0, 1\}$, is then given by whether a latent propensity to default is positive, $d_{i,t} > 0$. The latent propensity for defaulting on loans is again given by an additive and linear function of the observable characteristics, $z_{i,t}$, plus an unobserved shock $\nu_{i,t}$:

3)
$$d_{i,t} = \mu_t + \lambda z_{i,t} + \nu_{i,t}$$
.

Note that the vector of observable variables that explains default, $z_{i,t}$, differs from the vector of observable variables that explains the choice of the type of loan and the loan-amount, $x_{i,t}$. This is an intentional feature of the model and it is necessary for identification. The reason is because choice models that include an endogenous variable (for example, default in this model) affected by sample selection into different groups (for example, the type of loan chosen by agents in this model) are ill-identified if the same exact vector of variables explains both the endogenous variable choice and the sample selection choice (Vella, 1998). Therefore it is useful if there are at least a few variables that affect sample selection (the loan choice, in this case), but do not affect the default decision directly. In our application there are some valid candidates for this role of instruments that affect loan choice, but not default. Note that although for simplicity of exposition all the variables are indexed as being observed at the survey time t, in fact loans have a maturity of several periods (typically, around 1 or 2 years) and therefore the decision of loan choice happened before the repayment period. Therefore it is natural to use the lagged value of some variables as an explanation for loan choice and loan amount (for example, unemployment in the past year), but use the contemporary value of the same variables as an explanation for default. This choice of instrumental variables for loan choice is quite intuitive in economic terms and the validity of this identification approach is often recommended for panel data estimators (Vella, 1998).

To estimate the model it is necessary to specify the distribution of the unobservable random terms, which has some degree of subjectivity since there are several possible distributions that may provide a plausible fit. Unobservable factors can be thought as a mix of several elements who matter for families' decisions, but are not measured in the data. Such unobserved factors may include idiosyncratic preferences, the effects of geographical distance or marketing influence, the burden for the loan applicant to provide the information and legal documents requested by certain lenders, contractual costs of loans such as penalty charges or insurance fees, and even the result of negative events during the bargaining process between the loan applicant and the lenders. However, it is desirable that the distribution of the error term satisfies four characteristics: i) it allows for the unobserved preferences of each agent to be correlated over time, with some families being persistent in their behavior; ii) it accounts for some loan types being closer substitutes to each other, therefore the utilities of different options are correlated; iii) the agents' choice of all the distinct outcomes such as lender type, loan amount and default must be correlated, which is predicted by the theory of adverse selection of debtors (Einav, Jenkins and Levin, 2012); and, iv) the distribution should allow for an appropriate degree of heteroscedasticity, since groups are not equally affected by the unobserved shocks. A flexible way for achieving these desired properties is to assume the unobserved tastes for each option $\varepsilon_{i,b,t}$ are given by the sum of an independent extreme valued component plus a normal random-effect that is heteroscedastic and correlated over several choices and time periods (McFadden and Train, 2000, Nevo, 2000, 2011):

$$\begin{array}{l} 4.1) \ \varepsilon_{i,b,t} = \bar{\varepsilon}_{i,b} + \tilde{\varepsilon}_{i,b,t}, \\ \\ 4.2) \ \bar{\varepsilon}_{i,b} = 1(1 \le b \le 5)\eta_{i,1} + 1(1 \le b \le 2)\eta_{i,2} + 1(2 \le b \le 3)\eta_{i,3} + \omega_{i,b}, \end{array}$$

with $\tilde{\varepsilon}_{i,b,t} \sim EV(0,1)$, $\eta_{i,a} \sim N(0, \sigma_{\eta_a}(x_i))$ and $\omega_{i,b} \sim N(0, \sigma_{\omega_b}(x_i))$. 1(.) is the indicator function, assuming the value 1 if the condition is met and 0 otherwise. $\bar{\varepsilon}_{i,b}$ is the random-effect that represents the time-invariant tastes of the agent for each choice. Equation 4.2) for $\bar{\varepsilon}_{i,b}$ has a simple interpretation in terms of its distinct components, with $\eta_{i,1}$ representing a random factor denoting agent *i*'s taste for any type of loan, $\eta_{i,2}$ being a random factor denoting agent *i*'s taste for both the Bank and Bank plus Retail loan options, and $\eta_{i,3}$ denoting his taste for the options of Bank plus Retail or just Retail. Finally, the random effect $\omega_{i,b}$ is agent *i*'s specific taste for option *b*. The distribution of all the random-effects is heterocedastic in the vector x_i , which represents the time invariant characteristics of the agent and differs from $x_{i,t}$ which includes time-varying variables.

In a similar way, I assume the unobserved terms for loan amount, $\zeta_{i,t}$, and the propensity to default, $\nu_{i,t}$, are correlated with the unobserved tastes for loan types:

$$4.3) \zeta_{i,t} = \bar{\zeta}_i + \theta \bar{\varepsilon}_{i,b} + \tilde{\zeta}_{i,t},$$

$$4.4) \nu_{i,t} = \bar{\nu}_i + \rho [\bar{\varepsilon}_{i,b}, \bar{\zeta}_i] + \tilde{\nu}_{i,t}.$$

with $\tilde{\zeta}_{i,t} \sim N(0, \sigma_{\tilde{\zeta}}(x_i))$, $\bar{\zeta}_i \sim N(0, \sigma_{\bar{\zeta}}(x_i))$, $\bar{\nu}_i \sim N(0, \sigma_{\bar{\nu}}(x_i))$, and $\tilde{\nu}_{i,t} \sim EV(0, 1)$. The log-loan amount is a continuous variable and for this reason the contemporary shock that each agent faces can be heteroscedastic. Note that the unobserved propensity of default is correlated with both the unobserved factors of the loan types and the unobserved factors affecting the loan amount $\bar{\zeta}_i$.

The model includes random-effects, which requires panel data to identify the parameters. However, the EFH data contains some purely cross-sectional samples and it is inefficient to ignore such observations. For this reason the likelihood function includes both the panel and the cross-section samples, which is a specific case of a Full Information Maximum Likelihood (FIML) model. Let $\sum_i \equiv \{Y_{i,t}, L_{i,t}, D_{i,t}, Y_{i,t+s}, L_{i,t+s}, D_{i,t+s} \mid x_i, x_{i,t}, z_{i,t}, x_{i,t+s}, z_{i,t+s}\}$ be the vector containing agent *i*'s choices of type of loan, loan amount and default at both time *t* and *t* + *s*, conditional on the observables of both years. Also, let $\varepsilon_i \equiv \{\eta_{i,1}, \eta_{i,2}, \eta_{i,3}, \omega_{i,1}, ..., \omega_{i,B}, \overline{\zeta}_i, \overline{\nu}_i\}$ be the vector of all the unobservable random-effects. All the random-effects in vector ε_i are independent of each other, therefore the pdf of ε_i is given by $f(\varepsilon_i) = \phi(\frac{\eta_{i,1}}{\sigma_{\eta_1}(x_i)})...\phi(\frac{\eta_{i,3}}{\sigma_{\eta_3}(x_i)})...\phi(\frac{\overline{\zeta}_i}{\sigma_{\overline{\zeta}}(x_i)})\phi(\frac{\overline{\nu}_i}{\sigma_{\overline{\nu}}(x_i)}).$ This is assumed without any loss of generality, since the same random-effects affect different endogenous variables and therefore the endogenous variables are correlated with each other.

For simplicity of exposition it is easier to write the likelihood of the three endogenous variables given in equations 1), 2) and 3) conditional on the fixed-effects ε_i and then multiply it by the pdf $f(\varepsilon_i)$. Let $\tilde{U}_{i,b,t} = \alpha_{b,t} + \beta_b x_{i,t} + \bar{\varepsilon}_{i,b}$, $\ln(\tilde{L}_{i,t}) = \pi_t + \delta x_{i,t} + \bar{\zeta}_i + \theta \bar{\varepsilon}_{i,b}$, and $\tilde{d}_{i,t} = \mu_t + \lambda z_{i,t} + \bar{\nu}_i + \rho[\bar{\varepsilon}_{i,b}, \bar{\zeta}_i]$, represent the expected means for the latent variables of equations 1), 2) and 3), assuming ε_i is known. The likelihood of observing \sum_i can then be written as a simple product of the multivariate probability of the observed loan option b (given by the traditional multivariate logit ratio), with the probability of loan amount $(L_{i,t})$ and subsequent default $(D_{i,t})$ in both periods:

5)
$$\Pr(\sum_{i}) = \int \cdots \int f(\varepsilon_{i}) \frac{\exp(\tilde{U}_{i,b,t})}{\sum_{d} \exp(\tilde{U}_{i,d,t})} \phi(\frac{\ln(L_{i,t}) - \ln(\tilde{L}_{i,t})}{\sigma_{\tilde{\zeta}}(x_{i})})^{1(L_{i,t}>0)} \frac{\exp(\tilde{d}_{i,t})^{D_{i,t}}}{1 + \exp(\tilde{d}_{i,t})}$$
$$\frac{\exp(\tilde{U}_{i,b',t+s})}{\sum_{d} \exp(\tilde{U}_{i,d,t+s})} \phi(\frac{\ln(L_{i,t+s}) - \tilde{L}_{i,t+s}}{\sigma_{\tilde{\zeta}}(x_{i,t+s})})^{1(L_{i,t+s}>0)} \frac{\exp(\tilde{d}_{i,t+s})^{D_{i,t+s}}}{1 + \exp(\tilde{d}_{i,t+s})} \partial \varepsilon_{i}.$$

For the cross-sectional sample, let the vector $\sum_{i,t} \equiv \{Y_{i,t}, L_{i,t}, D_{i,t} \mid x_i, x_{i,t}, z_{i,t}\}$ represent agent *i*'s choices at time *t*, conditional on the observables $\{x_i, x_{i,t}, z_{i,t}\}$. If one assumes the panel and cross-sectional samples have the same representation in the population, then the likelihood function can be integrated for the same distribution of random-effects as the panel data observations. Note that this does not imply the model is unidentified, since the panel sample allows the model to identify the complete distribution of the unobservables. Therefore this approach is valid as long as the panel data sample is large enough. The likelihood of vector $\sum_{i,t}$ is therefore written as:

6)
$$\Pr(\sum_{i,t}) = \int \cdots \int f(\varepsilon_i) \frac{\exp(\tilde{U}_{i,b,t})}{\sum_d \exp(\tilde{U}_{i,d,t})} \phi(\frac{\ln(L_{i,t}) - \ln(\tilde{L}_{i,t})}{\sigma_{\zeta}(x_{i,t})})^{1(L_{i,t}>0)} \frac{\exp(\tilde{d}_{i,t})^{D_{i,t}}}{1 + \exp(\tilde{d}_{i,t})} \partial \varepsilon_i$$

The log-likelihood of the model is then given by the sum of the log-likelihood of the panel and cross-sectional samples, where $i \in P$ denotes whether the observation is in the panel sample or not:

7)
$$LL = \sum_{t=1}^{T-1} \sum_{i=1,i\in P}^{N_t} \sum_{b=1}^B \ln(\Pr(\sum_i)) + \sum_{t=1}^T \sum_{i=1,i\notin P}^{N_t} \sum_{b=1}^B \ln(\Pr(\sum_{i,t})).$$

Besides the time-varying error terms, this model has 11 unobserved random-effects which form the vector ε_i and influence the correlation of different choices and periods. This implies that the likelihood function of equations 5) and 6) is based on a high dimensional integral and it is computationally difficult to calculate precisely. For this reason the choice probabilities are not calculated exactly, but rather based on an approximation which averages a limited number of draws, R, from the distribution of $f(\varepsilon_i)$. This Simulated Maximum Likelihood (SML) method is asymptotically consistent if R increases proportionally with N (Train, 2009). In this application I use 100 draws to simulate the probability of each observation, with the multivariate draws chosen by a Modified Latin Hypercube Sampling (MLHS) method (Hess, Train and Polak, 2006).⁹ In general, the MLE asymptotic distribution is also valid for the SML method, but this asymptotic

⁹I choose the MLHS method, because it chooses pseudo-random draws equally spaced in each dimension of the integral and then randomly paired across dimensions. The reason why MLHS can perform better than standard uniform draws is because uniform draws can have too much randomness and there is a certain probability of obtaining

distribution is invalid if the model is not exactly true and if the number of draws R does not converge to infinity (Train, 2009). Therefore the model's standard-errors are estimated from 100 bootstrap replicas, which is asymptotically valid under a general set of conditions (Horowitz, 2001).

This model of loan decisions has certain implicit assumptions into it, since it assumes choices are well approximated by a function of known characteristics and randomly distributed unobserved preferences. One could assume other models for debt choice, such as an explicit multi-period optimization where agents choose the best option for maximizing expected lifetime utility based on an explicit evaluation of uncertain future paths and punishment costs for defaulting (see for instance, Chatterjee et al., 2007). However, an explicit lifetime optimization framework requires several assumptions about the agents' utility functions, their discount rates relative to future consumption and the knowledge agents have about their uncertain future outcomes. Empirical evidence of agents' cognitive limitations disputes assumptions such as rational expectations, time-consistency and revealed preference (Bertrand and Morse, 2009, Kahneman, 2011). Therefore simple behavioral models are not necessarily less realistic than structural models based on complete optimization. For this reason, the choice model in this paper is more closely related to other works who approximate agents' decisions in a flexible way, such as Edelberg (2006) and Einav, Jenkins and Levin (2012).

draws too close to each other, while some areas of the integral have few or no draws at all. In this sense MLHS guarantees that all the areas of each dimension are represented with at least one draw and therefore the simulated draws have a wider coverage. Some simulation studies show that 100 MLHS draws can be as efficient as more than 1000 uniform draws (Hess, Train and Polak, 2006). The MLHS method to obtain R multivariate draws basically starts with an equal spaced sequence of values, $\varphi(j) = \frac{j-1}{R}$ for j = 1, ..., R, in each dimension. Then a scrambled Halton pseudo-uniform number x is added to the draws of each dimension to get $\tilde{\varphi}(j) = \varphi(j) + \frac{x}{R}$ for j = 1, ..., R. The draws are then transformed using the inverse normal cdf and multiplied by the standard-deviation of the univariate distribution of the integral, $\sigma \Phi^{-1}(\tilde{\varphi}(j))$, to obtain an univariate normal draw. The draws of each dimension are then randomly paired with the R draws from the other dimensions to obtain R multivariate normal draws.

3 Data

3.1 The Chilean Household Finance Survey (EFH)

The main source of information for the characterization of the financial behavior of Chilean households is the Chilean Household Finance Survey (in Spanish, Encuesta Financiera de Hogares, hence on EFH). The EFH is a representative survey with detailed information on households' assets, debts, income and financial behavior, and is broadly comparable to similar surveys in the United States and Europe (Eurosystem, 2009). In 2007 and 2011 the EFH interviewed 3828 and 4059 urban families nationwide. In the years 2008 to 2010 the EFH was only implemented in the capital city of Santiago (which represents over 40% of the total national population), therefore the sample size is smaller for those waves. The EFH has a rotating sample, in which part of the sample is re-interviewed. Therefore there are 1792 families which were interviewed both in 2007 and 2011, while 947 families were interviewed both in 2008 and 2009. In total there are 6790 cross-sectional observations (i.e., families interviewed only once) plus 2739 panel observations (Table 1).

Table 1: Panel and cross-sectional sample size of the Household Finance Survey (EFH)

EFH	Panel	Cross-Section	Total
2007	1,792	2,036	$3,\!828$
2008	947	207	1,154
2009	947	243	1,190
2010		2,037	2,037
2011	1,792	2,267	$4,\!059$
Total	2,739	6,790	12,268

The EFH has a particularly detailed focus of the loans and debt commitments of each household. It asks for the largest 3 debts that each household has for each type of loan, among a total of 13 categories of loans: Banking Credit Card Debt, Banking Line of Credit, Banking or Financial Agency Consumer Credit Loan, Retail Store Credit Card, Retail Store Consumer Loan, Auto Loans, Union Credit, Education Loans, Loans from relatives, Loans from usurers, Pawn shops, Grocery and Shopping on credit (i.e., store tabs), and Other Debts. Therefore the survey may ask up to a total of 39 debts that the household has at the moment, although obviously few agents will report having debts with all the possible categories of loans.

For two reasons it is easier to work with just 5 types of lenders (or 5 types of loans), therefore my analysis is limited to options that sum all the loans for a given lender type and with each family classified discretely with the lender type representing the largest loan amount: Banks, Banks and Retail Stores (for the families reporting the use of both kinds of loans). Retail Stores, Union Credit. and Other Debts. The first reason is that it is desirable to eliminate the irrelevant alternatives from the choice model (Train, 2009), with a classic example being the inclusion of options such as "red bus" and "blue bus" for agents that do not care about the color of public transport. Several of the 13 types of loans elicited by the survey are similar products and are often offered by lenders to the same customers and for similar purposes (for instance, many customers use Credit Cards and Lines of Credit for similar reasons, although their choices may depend on the specific convenience of the occasion). This is a strong reason for aggregating all the options for credit cards, lines of credit and consumer contracts for each lender, instead of treating them separately. The second reason is related to the curse of multidimensionality, since the number of parameters in the model increases with the number of options and it is difficult to make a reliable analysis of too many options, particularly if some options have few or no observations (for example, loans from usurers are reported by less than 0.07% of the families). For this reason, Other Debts represents the sum of Auto Loans, Education Loans, Loans from relatives, Loans from usurers, Pawn shops, Grocery and Shopping on credit (i.e., store tabs), and Other Debts. Note that this category is largely composed of Auto Loans, Education Loans and Other Debts, with the remaining options representing negligible numbers.

Table 2 shows the proportion of households that chose each of the 5 lender types, plus households with either No Consumer Debt (because the family does not want debt) and No Access to Debt (if the family applied for loans, but was refused). The proportion of households without a wish for consumer debt represents 27% of the Chilean population, while those with No Access to Debt represent close to 13% of the population. Retail Stores are the most popular choice among households, representing more than 40% of the population, with 29% being Retail Store only users and 13% being users of both Bank and Retail Store Loans.

For each debt the EFH survey registers its loan amount, maturity, and other details such as the motivation for contracting the loan. The survey asks about the loan's interest rates, but less

Table 2: Population^{*}, Maturity (months), Loan amounts (thousands of Chilean pesos), motivation (share of total consumer debt destined for a given purpose), interest rate and morosity rates (EFH)

Type of Debtor Population Maturity Morosity Amount: mean/percentiles Interest Durables Pay debts Health

		Mean	Mean	Mean	p25	p50	p75	Mean	Mean	Mean	Mean
Bank	7.8%	25.28	10.2%	2,549	416	1,110	$2,\!649$	19%	21.7%	14.3%	5.0%
$\operatorname{Bank+Retail}$	12.9%	20.45	21.3%	3,015	859	1,703	3,520		15.6%	16.9%	6.0%
Retail Store	28.9%	12.17	19.0%	492	102	216	486	47%	2.9%	3.7%	3.3%
Union Debt	5.6%	27.07	11.4%	1,124	307	590	$1,\!131$	21%	15.8%	18.1%	13.8%
Other Debts **	4.6%	32.25	21.5%	4,101	1,207	2,425	4,274		6.6%	8.0%	3.2%
No Consumer Debt	27.3%										

No Access to Debt 13.0%

* % of the total Chilean households in urban areas. ** Maturity for Other Debts is for Auto Loans only.

than half the respondents report to remember these.¹⁰ The question used for measuring default is "Approximately, in the last 12 months have you fallen into morosity or late payments for each one of your loans?". I define default as a dummy variable denoting one or more events of morosity.

Table 2 shows the loan amount, maturity and morosity rates for each lender type using the pooled EFH sample, that is all the cross-sectional samples available. I also report the average loan interest rates of different lender types, from statistics of the Chilean Superintendency of Banks and Financial Institutions and the Superintendency of Social Security. There is no information on interest rates from users of several lenders, such as Banks and Retail or Other Debts. Table 2 also reports the share of the consumer loan destined for a given purpose of the household, more specifically "Purchase of Durables or Household Investments", "Pay previous debts" and "Health needs". Other motivations are classified as "General Consumption" so that the total motivations sum to 100% of the debt. Households with Bank, Bank plus Retail, and Union Debt are more likely to have motivations of "Purchase of Durables", "Pay previous debts" or "Health needs", which also motivate a significant part of unsecured debt in the USA (Chatterjee et al., 2007).

Users of Bank credit only have a morosity rate of 10%, which is half the value reported by users of both Bank and Retail credit (Table 2). Also, Bank users have much larger loan amounts

¹⁰This memory problem is explained by debtors tendency to remember the payment amount better than their contract's interest rate. Also, it is difficult to recover an estimate of the implicit interest rate from the financial formula for the present value of the payments of a loan. This is due to ommitted variables in the financial formula, such as loan commission fees, and measurement error in the reporting of the loan amounts, payments and maturities.

and longer maturities than the users of Retail Stores. In Chile neither Retail Stores or institutions of Union Credit are able to offer heterogeneous interest rates to their customers, only Banks offer customer specific interest rates (Marinovic, Matus, Flores and Silva, 2011), so the economic theory predicts that Banks will get the best observable risk types by offering better loan terms such as lower interest rates, larger loan amounts and longer maturities. While Union Debt lenders are unable to risk price their offers, these institutions are able to garnish their clients' wages easily. therefore this high punishment cost should explain their low morosity rates. However, households with both Bank and Retail Store debt have morosity rates as high as the customers of Retail Stores only. Perhaps this can be explained because such debtors have an unobservable taste for high loan amounts. Table 2 shows that households with both Bank and Retail Store debt have much higher loan amounts than the debtors of Bank and Retail Store separately, which could be a sign that these are debtors with particularly high needs for liquidity. The households with Other Debts also have high loan amounts and morosity rates, but perhaps this can be explained by special characteristics of these debtors. For example, education loans are granted to younger agents, who may be more subject to unemployment risk and unstable income. Also, perhaps education and auto loans have lower punishment costs for morosity, since lenders cannot deduct payments and punishment fees from their clients' bank accounts (as Banks do) or their wages (as Union Credit institutions do).

Table 3 shows the percentage of the population, median loan amounts and morosity rates in the years 2007 and 2011. The biggest changes observed between 2007 and 2011 are that users of only Banks and Union Debt increased respectively to 8.2% and 7.8% of the population. Loan amounts for users of Union and Other Debts increased substantially, while the median loan amount at Banks increased less. It is also noticeable that the morosity rate of Bank users increased somewhat.

Tables 4 and 5 summarize the changes to income and use of consumer loans in Chilean households, using information from the EFH panel sample (2007-2011). In Table 4 I report the transition probabilities from one household income quintile $(Q_{i,t})$ to another between 2007 and 2011, $\Pr(Q_{i,2011} = q \mid Q_{i,2007} = q')$, where 1 denotes the families with the 20% lowest income. The conclusion is that household income has some persistence, but there is substantial income volatility in Chile. The probability that a household of the lowest income (quintile 1) will remain at the bottom of the distribution is 40%, while the probability of a household remaining at the top income level (quintile 5) is 53%. Among the middle income levels (quintiles 2 to 4), mobility is even higher and there is

Type of Debtor	Popu	lation	Loan a	$mount \ (median)$	Morosity rate	
	2007	2011	2007	2011	2007	2011
Bank	6.5%	8.2%	968	$1,\!176$	8.8%	11.7%
$\operatorname{Bank+Retail}$	13.6%	11.8%	$1,\!435$	1,826	18.9%	24.6%
Retail Store	31.9%	25.9%	232	177	21.1%	19.5%
Union Debt	3.8%	7.8%	484	748	12.1%	12.2%
Other Debts	4.6%	4.9%	$1,\!511$	2,866	25.2%	20.5%
No Consumer Debt	26.6%	28.7%				
No Access to Debt	13.0%	12.7%				

Table 3: Population of debtors, loan amounts (thousands of pesos) and morosity over time (EFH)

a high chance that households will move into either a higher or a lower income level.

In Table 5 I show the transition probability of a household changing from one lender type to another or towards having either no consumer debt or no access to debt, $\Pr(Y_{i,2011} = b \mid Y_{i,2007} = b')$. The last column in the table replicates the share of the population in each debt status over the whole period of 2007 to 2011. If one compares the diagonal values of the transition matrix, which represent the probability of a debtor keeping the same status as previously, with the average debt status of the population, then one gets an idea of how persistent agents are in their choices. It is clear that the probability of an agent keeping the same debt status is above the average rate in the total population and this happens for all categories, therefore choices tend to be persistent. In particular, debtors of Union Debt, Banks or of Bank plus Retail Store are more than twice as likely to keep their choices relative to the average probability in the population. Also, it is striking that debtors of Retail or Bank plus Retail have a probability of only 20% and 16% respectively of moving into a state of No Debt. Therefore these debtors are systematically in need of debt, whether with the same lender or a different one. This confirms the previous results that debtors of Bank plus Retail appear to be agents with higher needs for liquidity relative to other households.

4 The sorting of income risk across different types of loans

The EFH survey collects detailed information on the income, education, age and other characteristics of each household member, but it has limited data on some aspects, such as their income volatility

	Quintile 2011						
Quintile 2007	1	2	3	4	5		
1	40%	27%	19%	7%	7%		
2	22%	$\mathbf{29\%}$	27%	15%	7%		
3	11%	23%	$\mathbf{26\%}$	25%	15%		
4	10%	14%	22%	30%	24%		
5	7%	9%	12%	20%	53%		

Table 4: Transition of families across different income quintiles (EFH Panel, 2007-11)

Table 5: Transition of households across different debtor types (EFH Panel, 2007-11)

Debt Status in 2011 Pop								Population
Debt in 2007	No Debt	Bank	Bank+Retail	Retail	Union	\mathbf{Other}	No Access	in 2007-11
$\mathbf{No} \ \mathbf{Debt}$	40.9%	7.0%	6.7%	27.8%	4.2%	2.3%	11.1%	27.3%
Bank	27.8%	18.1%	20.5%	19.1%	3.9%	2.1%	8.4%	7.8%
Bank+Retail	16.4%	18.2%	30.0%	25.1%	2.2%	1.5%	6.7%	12.9%
Retail Store	20.9%	7.5%	13.6%	$\mathbf{39.1\%}$	5.2%	0.4%	13.3%	28.9%
Union Debt	36.2%	8.1%	3.5%	34.0%	14.1%	0.0%	4.1%	5.6%
Other Debts	34.7%	12.7%	28.4%	12.6%	0.0%	5.5%	6.1%	4.6%
No Access	33.4%	4.9%	7.8%	28.9%	6.9%	0.6%	17.6%	13.0%

or stability of employment. For this reason I estimate the income and employment risks of the EFH workers based on the mean statistics for workers with the same characteristics in another dataset.

Based on the quarterly Chilean Employment Survey, ENE, which covers 35,000 households, Madeira (2014) estimated three measures of risk in employment status for the period 1990 to 2012: the unemployment rate $(u_{k,t} = \Pr(U_{k,t} = 1 \mid t, x_k))$, the separation rate $(\lambda_{k,t}^{EU} = \Pr(U_{k,t+1} = 1 \mid t, U_{k,t} = 0, x_k))$ defined as the probability of being unemployed given that one was employed in the previous quarter, and the job finding rate $(\lambda_{k,t}^{UE} = \Pr(U_{k,t+1} = 0 \mid t, U_{k,t} = 1, x_k))$ defined as the probability of being employed after being unemployed in the previous quarter. The vector x_k is composed of 540 mutually exclusive groups, given by $x_k = \{$ Santiago Metropolitan city or Outside, Industrial Activity (primary, secondary, terciary sectors), Gender, Age (3 brackets, $\leq 35, 35 - 54$, ≥ 55), Education (less than secondary schooling, secondary or technical education, college), and Household Income quintile}. Madeira (2014) also computed these groups' labor income volatility even if no job is lost, $\sigma_{\zeta,t}(x_k) = \sqrt{E\left[(Y_{k,t} - E[Y_{k,t} \mid Y_{k,t-1}, x_k])^2 \mid t, U_{k,t} = U_{k,t-1}, Y_{k,t}, x_k],$ and the income loss caused by going into unemployment, $R_{k,t}(x_k) = \frac{E\left[Y_{k,t} \mid t, U_{k,t} = 1, x_k\right]}{E\left[Y_{k,t} \mid t, U_{k,t} = 0, x_k\right]}$.

Using these labor risk measures I calculate the expected income $P_{i,t}$ of each EFH household i as the sum of their non-labor income, a_i , and its expected labor income, $P_{i,t}$: $\bar{P}_{i,t} = a_i + P_{i,t}$, where $P_{i,t} = \sum_k P_{k,t}$ is the sum of expected labor income of each household member k. $P_{k,t} = W_{k,t}(1 - u_{k,t}) + W_{k,t}R_{k,t}(u_{k,t})$ is each worker k's average labor income during the employed and unemployed states. The employment risk of each household is then given by a weighted average of the rates of each member using their labor income relative to the total household labor income: $\bar{u}_{i,t} = \sum_k \frac{P_{k,t}}{P_{i,t}} u_{k,t}, \, \bar{\lambda}_{i,t}^{UE} = \sum_k \frac{P_{k,t}}{P_{i,t}} \lambda_{k,t}^{UE}$ and $\bar{\lambda}_{i,t}^{EU} = \sum_k \frac{P_{k,t}}{P_{i,t}} \lambda_{k,t}^{EU}$. Similarly, the household's weighted labor income volatility (even if no job is lost) and the replacement ratio during unemployment are given by $\bar{\sigma}_{i,t} = \sum_k \frac{P_{k,t}}{P_{i,t}} \sigma_{\zeta,t}(x_k)$ and $\bar{R}_{i,t} = \sum_k \frac{P_{k,t}}{P_{i,t}} R_{k,t}(x_k)$.

Figure 2 shows the cumulative distribution function of the loan amounts (in logarithm) and the consumer debt to annual income ratio $\left(\frac{L_{i,t}}{12 \times \bar{P}_{it}}\right)$, where \bar{P}_{it} is the expected monthly income) in the pooled EFH survey (2007-11). Retail only debtors are the ones with the highest probability of having low loan amounts, since their cdf is stochastically dominated by either Union and Bank debtors. Bank plus Retail debtors and Other debtors have the greatest probability of having high loan amounts (or the lowest probability of having low loan amounts). One question is if the difference in loan amounts is entirely explained by income, since higher income households may



Figure 1: The Cdf of the loan amounts chosen by debtors of different loan types

also pay larger loans. The answer is given by the empirical cdf of the consumer debt to annual income ratio. In Figure 2 it is shown that clearly Retail only debtors have lower debt to income values in relation to both Union and Bank debtors. Also, Bank plus Retail and Other Debts users have the highest debt to income ratios. Therefore the differences in the sorting of loan amounts across lender types remains even if we take into account household income.

Table 6 reports the mean values of the household' measures for the unemployment rate $(\bar{u}_{i,t})$, the separation rate $(\bar{\lambda}_{i,t}^{EU})$ and the job finding rate $(\bar{\lambda}_{i,t}^{UE})$ across different loan choices. The groups with No Consumer Debt or only Bank loans are the ones with the lowest unemployment and separation rates. Households with Other Debts are the ones with the highest average unemployment rates, perhaps because of their younger age. The mean job finding rate is between 31% to 37% for all groups. Table 6 also reports the means values for the log household expected income $(\ln(\bar{P}_{i,t}))$, the labor income volatility $(\bar{\sigma}_{i,t})$ and its replacement ratio of income during unemployment $(\bar{R}_{i,t})$. Bank only customers are the group of highest income, while those with Retail Store loans or with No

Debtor Type	$\bar{u}_{i,t}$	$\bar{\lambda}_{i,t}^{EU}$	$\bar{\lambda}_{i,t}^{UE}$	$\ln(\bar{P}_{i,t})$	$\bar{\sigma}_{i,t}$	$\bar{R}_{i,t}$
Bank	4.8%	2.0%	33.8%	13.56	18.4%	25.8%
$\operatorname{Bank+Retail}$	5.3%	2.3%	35.4%	13.46	18.2%	25.5%
Retail Store	5.5%	2.6%	36.6%	13.01	16.5%	23.5%
Union Debt	5.0%	2.0%	30.9%	13.14	17.4%	22.7%
Other Debts	6.1%	2.2%	34.2%	13.47	20.7%	26.1%
No Consumer Debt	4.2%	1.9%	30.6%	13.13	16.2%	23.0%
No Access to Debt	5.4%	2.2%	31.0%	12.77	17.6%	21.3%

Table 6: Mean values of labor market risk and household earnings across debtor types (EFH)

Access to Debt have the lowest mean income. Unemployment represents a strong income reduction for Chilean households, since the mean values of $\bar{R}_{i,t}$ imply that agents only keep 21% to 26% of their working income during an unemployment spell. The households with No Consumer Debt appear to be the group least susceptible to shocks, since they are the group with the lowest unemployment rate, lowest separation rate and lowest labor income volatility. The permanent income theory of consumption predicts that agents should use debt to smooth temporary income shocks (see Chatterjee et al., 2007, or Dynan and Kohn, 2007), therefore it makes sense that households with the lowest income risk also have the lowest demand for consumer loans.

While Table 6 reports the mean values of households' income, employment risks and income volatility, it is also useful to analyze how heterogeneous households are and how each group deviates from the mean. Figure 3 shows the cdf of the households' expected income $(\ln(\bar{P}_{i,t}))$, unemployment rate $(\bar{u}_{i,t})$, labor income volatility $(\bar{\sigma}_{i,t})$, which can also be denoted as the standard deviation of wage shocks) and replacement ratio of income during unemployment $(\bar{R}_{i,t})$ for debtors and non-debtors. For simplicity, I use only 4 groups in the graphical comparison instead of the 7 groups used in Table 6 and the previous tables. Basically, I classify households in the same two options for non-debtors as before (No Consumer Debt, No Access to Debt), but use only two classifications for the groups of debtors: i) users of Retail Store loans only, which represent 29% of the household population (Table 2) and are the largest group with consumer debt; and, ii) users of Bank, Union Debt and Other Debts, which represent 30.9% of the Chilean population (this figure is obtained by summing the distinct categories of this group in Table 2). Another simplification concerns the problem that often households have a lot of heterogeneity at the extreme margins, but one is mostly concerned

with the heterogeneity that affects most of the population and not its extreme points (which could eventually be outliers due to measurement error). Therefore to make the graphs easier to read the cdfs are plotted only in the range of 20% to 90% probability.

Figure 3 shows that in terms of income there is a clear stochastic dominance among the different groups, with households with No Access to Debt having lower income than those with Retail loans and those with Retail loans having lower income than both the households with No Consumer Debt and the households with Bank, Union and Other Debts. Also, it is clear that households with No Consumer Debt have the lowest unemployment rates, which is another confirmation that a partial motivation for consumer loans is to smooth temporary income shocks. Households with Bank, Union and Other Debts also have lower unemployment rates relative to those with Retail loans only or No Access to Debt. Labor income volatility ($\bar{\sigma}_{i,t}$) is highest for the households with Bank, Union and Other Debts, which may imply that consumer debt is used for smoothing income shocks in this group. The replacement ratio of income during unemployment is the lowest for those with No Access to Debt, followed by the users of Retail loans only and those with No Consumer Debt. Users of Banks, Union and Other Debts have the highest replacement ratios during unemployment, therefore this is the group that suffers the lowest loss of income from job loss.

Figure 3 shows that income and labor experiences have a lot of heterogeneity in the population. Unemployment rates can range from as low as 2% to as high as 8%. Labor income volatility has a range between 11% to 27%, while replacement ratios can vary between 18% and 34%.

Besides analyzing unemployment rates, it is also appropriate to look at the employment separation $(\bar{\lambda}_{i,t}^{EU})$ and job finding $(\bar{\lambda}_{i,t}^{UE})$ rates. The reason is because unemployment rates has a different interpretation if it is driven by high separation rates (lots of workers losing their jobs) or by low job finding rates (which implies that unemployed workers have difficulties finding jobs and therefore unemployment spells last a long time). Both of these employment transition rates play a role in explaining labor market shocks in the United States (Shimer, 2012) and in Chile (Madeira, 2014).

Figure 4 shows the cdf of the separation $(\bar{\lambda}_{i,t}^{EU})$ and job finding $(\bar{\lambda}_{i,t}^{UE})$ rates for Chilean households. There is a lot of heterogeneity in these variables, with the separation rate ranging from as low as 1% to as high as 4% and the job-finding rate varying between 25% and 55%. The separation rate has the most clear differences between debtor and non-debtor groups. Households with No Consumer Debt have lower separation rates than users of Bank, Union and Other Debts, and these



Figure 2: The Cdf of labor market characteristics of debtors versus non-debtors



Figure 3: The Cdf of employment transition probabilities for different loan types

last ones have lower separation rates than those with No Access and the users of Retail loans only. The differences in job-finding rates are less clear. Users of Retail loans only have both the highest separation rates and the highest job finding rates, which implies that employment mobility is high in this group. However, the groups with No Consumer Debt, No Access and users of Bank, Union and Other Debts have a similar distribution for the job-finding rate.

Figure 5 shows the differences in income and labor market characteristics of different debtor groups. Users of Retail and Union Debt are the ones with the lowest income, while the users of Bank and Other Debts have the highest income. Also, Bank users have a lower unemployment rate than all the other debts, with Union Debt users being the second group with the lowest unemployment rates and users of Other Debts having the highest unemployment. Retail and Union Debt users, however, have the lowest labor income volatility (or standard-deviation of wage shocks), while users of Bank and Other Debts have the highest wage risk. Users of Retail and Union Debt are the ones with the lowest replacement ratios and therefore suffer the most during a jobless spell.

Overall, Figures 2, 3 and 4 portray a clear picture of different income and labor market characteristics across non-debtors and different groups of debtors. Households with No Access to Debt have the lowest income, highest unemployment rates and lowest replacement ratios of income, therefore it is the group most subject to low income and income fluctuations. Households with No Consumer Debt (because of a lack of demand for such loans) have the lowest unemployment rates, separation rates and labor income volatility, therefore it is the group least subject to income shocks. The users of just Bank loans are the ones with the highest income, highest replacement ratio and lowest unemployment rate, but they suffer from substantial wage volatility which may create a demand for smoothing consumption. Users of Other Debts have high income and high replacement ratios in the same way as Bank users, but they are the debtor group most subject to both high unemployment rates and high labor income volatility, therefore it could be seen as a riskier segment relative to Bank users. Finally, users of Retail loans are the ones with the lowest income among debtors (although they have higher income than the group with No Access to Debt), and also have a high unemployment rate and low replacement ratio, which could make them a riskier debt segment. However, Retail users have a low standard-deviation of wage shocks, therefore their income is relatively stable during their employment experience. Users of both Bank plus Retail loans are a segment somewhat in between the exclusive users of either Banks or Retail loans.

5 Results

5.1 The role of demographics, income profile and unobserved preferences

Now I discuss the results from the consumer loan choice and default model exposed in section 2. As explained before, the model requires some variables that affect loan choice, but not default. Since unemployment risk, labor income volatility are measured for several time periods (all the quarters from 1990 to 2012) for each type of worker, then it is possible to create these variables for each EFH household for previous periods than the survey date. It is natural to assume that households were driven by labor market effects that happened at the time of the loan contract, which was a substantial time before the current period t. Consumer loans typically have a maturity of 12 to 24 months, therefore it is reasonable to assume that the labor market conditions that influenced loan



Figure 4: The Cdf of labor market characteristics of debtors of different loan types

choice happened 4 quarters or more before the current period. For this reason the vector affecting loan choice includes expected income $(\ln(\bar{P}_{i,t-4}))$, unemployment risk $(\bar{u}_{i,t-4})$ and labor income volatility $(\bar{\sigma}_{i,t-4})$ with a lag of 4 quarters (although a shorter or longer lag could be used). Note that expected income is a weighted sum of each household member's labor income, its unemployment probability and its replacement ratio, therefore it can be estimated for previous time periods. In addition the vector that affects loan choice includes the education, age and structure of the household (whether it is a couple or a family with many members), and the loan's motivation:

 $x_{i,t} = \begin{cases} \ln(\bar{P}_{i,t-4}), \text{ unemployment risk } \bar{u}_{i,t-4}, \text{ labor income volatility } \bar{\sigma}_{i,t-4}, \\ \text{years of education of household head, age of the household head, age squared,} \\ \text{dummies for each year, dummy for residence being out of the Santiago capital,} \\ \text{dummy for 2 members and dummy for 3 or more members in household,} \end{cases}$

Share of debt justified by "Durables", "Pay Previous Debts" and "Health".

In a similar way I assume that the vector $z_{i,t}$ that explains loan morosity or default at time t includes some variables that do not necessarily affect loan choice. One variable is the ratio of consumer debt to the annual income $(RDI_{i,t} = \frac{L_{i,t}}{12 \times \bar{P}_{i,t}})$, which can be seen as a measure of long-term solvency of the household. Households with larger loans may feel more stressed about their long term commitments and choose to default on their loans. In the same way, some households may be more worried about this month's specific commitments instead of their long-term expenses. For those households, the current monthly debt service (the debt service of a loan, $DS_{i,t}$, includes both the amortization and the interest payments) relative to this month's income $(Y_{i,t})$ may provide a liquidity motive for defaulting or simply paying a loan with some delay. For this reason I also include the ratio of monthly Debt Service to Income $(RDSI_{i,t})$ as a possible factor affecting household default. The overall vector of observables that explain default includes the financial ratios $RDI_{i,t}$, $RDSI_{i,t}$, plus the current expected income $(\ln(\bar{P}_{i,t}))$, unemployment risk $(\bar{u}_{i,t})$ and labor income $(\bar{\sigma}_{i,t})$, and the same demographic variables that affect loan choice:

 $z_{i,t} = \begin{cases} \frac{L_{i,t}}{12 \times \bar{P}_{i,t}}, \frac{DS_{i,t}}{Y_{i,t}}, \ln(\bar{P}_{i,t}), \text{ unemployment risk } \bar{u}_{i,t}, \text{ labor income volatility } \bar{\sigma}_{i,t}, \\ \text{years of education of household head, age of the household head, age squared,} \\ \text{dummies for each year, dummy for residence being out of the Santiago capital,} \\ \text{dummy for 2 members and dummy for 3 or more members in household,} \\ \text{Share of debt justified by "Durables", "Pay Previous Debts" and "Health".} \end{cases}$

Finally, I need to specify the degree of heteroscedasticity in the unobserved tastes for loan choice,

	1 = Bank	2=Bank+Retail	3=Retail Store	4=Union Debt	5=Other Debts	6=No Access
2007	-9.565 (0.564)***	-6.409 (0.610)***	$4.662 \ (0.727)^{***}$	$4.427 (1.073)^{***}$	-6.351 (3.809)*	8.357 (0.824)***
2008 / 09	-9.734 (0.550)***	-6.975 (0.592)***	$4.506 \ (0.789)^{***}$	$3.976\ (1.081)^{***}$	-7.044 (4.068)*	8.558 (0.82)***
2010	-9.545 (0.574)***	-7.133 (0.601)***	$4.739 \ (0.706)^{***}$	$5.008 \ (1.064)^{***}$	-6.265 (3.783)*	$8.84 \ (1.038)^{***}$
2011	-9.388 (0.571)***	-6.66 (0.594)***	$4.742 \ (0.682)^{***}$	$5.416 \ (1.079)^{***}$	-5.801 (1.047)***	8.847 (0.953)***
Income: $\ln(\bar{P}_{i,t-4})$	$0.323 \ (0.039)^{***}$	$0.078 \ (0.04)^{**}$	-0.495 (0.156)***	-0.544 (0.158)***	$0.131 \ (0.052)^{**}$	-0.741 (0.156)***
Education	$0.065 \ (0.014)^{***}$	$0.035 \ (0.011)^{***}$	-0.116 (0.053)**	-0.076 (0.027)***	$0.039 \ (0.009)^{***}$	-0.078 (0.017)***
Unemployment $\bar{u}_{i,t-4}$	$2.245 \ (1.162)^*$	$2.762 \ (0.965)^{***}$	$4.484 \ (1.452)^{***}$	$4.136 \ (0.316)^{***}$	$4.695 (2.005)^{**}$	$2.307 \ (1.385)^*$
Wage volatility $\bar{\sigma}_{i,t-4}$	-0.011 (0.057)	-0.041(0.074)	-0.223 (0.380)	$1.200 \ (0.378)^{***}$	$3.213 \ (0.599)^{***}$	$1.471 \ (0.804)^*$
Out of Santiago	-0.181 (0.052)***	-0.307 (0.138)**	-0.295 (0.091)***	-0.504 (0.036)***	-0.267 (0.024)***	-0.246 (0.110)**
Age of home head	$0.108 \ (0.029)^{***}$	$0.153 \ (0.046)^{***}$	$0.112 \ (0.056)^{**}$	-0.003 (0.022)	$0.047 \ (0.014)^{***}$	$0.043 \ (0.012)^{***}$
Age squared/ 100	-0.133 (0.059)**	-0.187 (0.104)*	-0.138 (0.067)**	0.013(0.014)	-0.087 (0.051)*	-0.052 (0.013)***
2 members in home	$0.323 \ (0.123)^{***}$	$0.620 \ (0.151)^{***}$	$0.786 \ (0.231)^{***}$	$0.051 \ (0.063)$	$0.766 \ (0.463)^*$	$0.005\ (0.026)$
3 or more members	$0.598 \ (0.124)^{***}$	$1.16 \ (0.308)^{***}$	$1.391 \ (0.671)^{**}$	$0.318 \ (0.172)^*$	$1.079 \ (0.327)^{***}$	$0.298 \ (0.168)^*$
Durables	$5.329 \ (2.657)^{**}$	$5.061 \ (2.521)^{**}$	$2.544 \ (1.432)^*$	$4.814 \ (2.367)^{**}$		
Pay previous debts	$5.352 (2.148)^{**}$	$5.513 (2.153)^{**}$	$3.081 \ (1.507)^{**}$	$5.157 (2.089)^{**}$		
Health needs	$5.619 (2.52)^{**}$	6.014 (2.525)**	$4.628 (2.497)^*$	$6.561 \ (2.483)^{***}$		

Table 7.1: Coefficients for the mean value of the Utility of each type of Loan

loan amount and default, that is the standard-deviation of each element of the normally distributed vector $\{\tilde{\zeta}_{i,t}, \varepsilon_i\}$, where the vector of random-effects of tastes (i.e., tastes that are constant over time for each agent) is given by $\varepsilon_i \equiv \{\eta_{i,1}, \eta_{i,2}, \eta_{i,3}, \omega_{i,1}, ..., \omega_{i,B}, \bar{\zeta}_i, \bar{\nu}_i\}$. In this case I assume that all the standard-deviations are exponential functions of a linear-index, $\sigma = \exp(\beta x_i)$, which guarantees that all standard-deviations are positive. The vector x_i that models heteroscedasticity includes a constant, a dummy for the 2008/09 panel, the years of education of the household head, plus the average labor market characteristics of the household:

 $x_i = \left\{ \text{constant, dummy for 2008/09, years of education of head, } \frac{1}{T} \sum_{t=2007}^{2011} \left\{ \ln(\bar{P}_{i,t}), \bar{u}_{i,t}, \bar{\sigma}_{i,t} \right\} \right\}.$ Table 7.1 shows the estimates for the coefficients of loan choice, β_b . The coefficients of a multivariate logit model sometimes have a difficult interpretation (Train, 2009), because the agents' choice is made over a multivariate set with B + 1 choices, $\max(U_{i,0,t}, U_{i,1,t}, ..., U_{i,B,t})$, with the first choice being standardized to have value zero. Let us think of a generic variable x and its coefficient on choice b, β_b , which is assumed to be positive. Then $\beta_b > 0$ implies the odds ratio of the probability of option b relative to option 0 is increasing in x, meaning that larger x makes b more likely to be chosen relative to option 0. However, at the same time there could be another option c which has a larger coefficient than b, implying x decreases the chance of b being chosen relative to option c. Therefore in the multivariate case $\beta_b > 0$ does not always increase the probability of b being chosen with larger x. Such is the case only if $\beta_b \ge \max(\beta_1, ..., \beta_B)$. This interpretation of the multivariate logit coefficients must be kept in mind while reading Table 7.1.

The coefficient for the lagged household expected income $(\ln(P_{i,t-4}))$ is the largest (i.e., the most positive) for the option of Bank loans, while it is the lowest (i.e., the most negative) for the No-Access option (Table 7.1). This implies that larger income unambiguously increases the option of a Bank loan and decreases the option of No-Access. The impact of income on the choice of Retail Store loans and Union Debts is negative, therefore larger income increases the likelihood of No Debt in relation to these options. The coefficient of education is largest for the Bank option and lowest for the Retail Store option, which implies that education increases the probability of a Bank loan and decreases the option of Retail Store loans. The coefficient of lagged unemployment ($\bar{u}_{i,t-4}$) increases the probability of all the loan options and the No Access option in relation to No Debt. However, unemployment has the effect of increasing more the probability increases the probability of choosing Union Debt and Other Debts options. Wage volatility increases the probability of choosing Union Debt, Other Debts and No Access. The positive impact of unemployment risk and wage volatility on the probability of choosing loans can be seen as evidence in favor of debt providing consumption smoothing in the face of increase shocks.

Being outside of the Santiago capital city lowers the probability of all loan options, with its strongest effect on Union Debt. With the exception of Union Debt, all loan options are quadratic in age. This result may be evidence in favor of loans performing a role in consumption smoothing over the life cycle, with older households starting to save for retirement and engaging in less debt. Households with more members are more likely to choose all kinds of loan options, but its impact is strongest for the options of Retail Store and Other Debts. Finally, the motivations for undertaking a loan (Durables, Pay previous debts or Health) have a special standardization, because families who report positive values for the loan motivation must have chosen one of the loan options 1 to 5 and therefore I standardize the loan motivation coefficients for the last option Other Debts as being 0. The loan motivations of Purchase of Durables, Pay Previous Debts and Health needs increase the probability of choosing the options of Bank, Bank plus Retail and Union Debt.

Table 7.2 shows the heteroscedasticity for the random effects that denote the unobserved tastes for each loan type. The heterogeneity for the tastes of all loan options decreases with household income. Also, education decreases the heterogeneity of unobserved factors for the options of Bank,

Table 7.2: Coefficients for the standard-deviation (in log) of the random-effect of each type $\omega_{i,b}$

1=Bank	2=Bank+Retail 3=Retail Store	4=Union Debt 5=Other Debts	6=No Access
--------	------------------------------	----------------------------	-------------

$\mathbf{constant}$	-0.346(0.312)	-0.065(0.092)	$1.614 \ (0.724)^{**}$	$0.146\ (0.394)$	-0.058(0.097)	$0.179\ (0.136)$
2008/09	$0.028\ (0.082)$	$0.018 \ (0.157)$	$0.289 \ (0.083)^{***}$	-0.961(0.683)	-0.099(0.108)	$-0.266 (0.161)^*$
$\frac{1}{T}\sum_{t=2007}^{2011}\ln(\bar{P}_{i,t})$	$-2.746(1.342)^{**}$	$-0.684 (0.324)^{**}$	$-0.129 (0.008)^{***}$	$-2.576 (1.531)^*$	$-0.414 \ (0.133)^{***}$	$-2.99(1.534)^*$
Education	$-1.924 \ (0.886)^{**}$	$-0.516 (0.151)^{***}$	$0.025\ (2.308)$	0.379(0.411)	-0.414 (0.203)**	-3.024 (1.146)**
$\frac{1}{T}\sum_{t=2007}^{2011} \bar{u}_{i,t}$	-0.056(0.084)	-0.005 (1.002)	$-2.264 (0.706)^{***}$	$0.181 \ (0.253)$	-0.003(0.132)	$0.132 \ (0.050)^{***}$
$\frac{1}{T} \sum_{t=2007}^{2011} \bar{\sigma}_{i,t}$	-0.063(0.082)	-0.011(0.769)	$-1.92 (1.159)^*$	0.007(1.271)	0.009(0.422)	$0.344 \ (0.151)^{**}$

Table 7.3: Coefficients for the standard-deviation (in log) of the factors affecting several choices

	factor 1, $\eta_{i,1}$	factor 2, $\eta_{i,2}$	factor 3, $\eta_{i,3}$
	$(choices \ 1 \ to \ 5)$	$(choices \ 1 \ to \ 2)$	(choices 2 to 3 $)$
constant	$0.096\ (0.117)$	-1.451 (0.522)***	$-1.43 (0.302)^{***}$
2008/09	$0.812 \ (0.149)^{***}$	$0.532 \ (0.171)^{***}$	$0.427 \ (0.224)^{**}$
$\frac{1}{T}\sum_{t=2007}^{2011}\ln(\bar{P}_{i,t})$	$-0.848 (0.231)^{***}$	0.082(0.450)	$0.077 \ (0.093)$
Education	$-1.442 \ (0.786)^*$	$0.005\ (0.156)$	$0.012 \ (3.732)$
$\frac{1}{T} \sum_{t=2007}^{2011} \bar{u}_{i,t}$	-0.009(0.013)	$2.306 (0.704)^{***}$	$1.246 \ (0.571)^{**}$
$\frac{1}{T} \sum_{t=2007}^{2011} \bar{\sigma}_{i,t}$	$0.052\ (0.125)$	-0.024(0.121)	-0.515(0.727)

Bank plus Retail, Other Debts and No Access. Unemployment risk decreases the heterogeneity of preferences for Retail Store debt, but it increases slightly the heterogeneity of factors affecting No Access. Table 7.3 shows the heteroscedasticity of the random-effects that affect several loan options. The main conclusion is that the heterogeneity for the unobserved tastes for all loans $(\eta_{i,1})$ is also decreasing in income and education. Therefore higher income and more highly educated households have lower heterogeneity of unobserved factors. Unemployment risk, however, increases the heterogeneity of factors associated with Bank plus Retail $(\eta_{i,2})$ and Retail Store $(\eta_{i,3})$ choices.

Table 8.1 shows the results for the choice of loan amount $(\ln(L_{i,t}))$ and default $(D_{i,t})$. It is worth noting that the expected income $(\ln(\bar{P}_{i,t-4}))$, unemployment risk $(\bar{u}_{i,t-4})$ and wage volatility $(\bar{\sigma}_{i,t-4})$ that affect the loan amount decision have a lag of 4 quarters, while the variables affecting default correspond to the current period t. Basically, income, unemployment rates, wage volatility, households with more members, and loan motivations of "Pay previous debts" and "Health" are positively related to loan amounts. Education is negatively related to loan amounts, which may denote a tendency of highly educated households to better manage their finances over time and resort less to expensive consumer debt (Table 2 showed that interest rates for consumer loans are high in Chile). Also, the estimates show that households with strong unobservable tastes for Bank loans, Bank plus Retail and Union Debts are more likely to have higher loan amounts.

The propensity to default is positively related to high levels of consumer debt relative to annual income $(RDI_{i,t})$, higher debt service $(RDSI_{i,t})$, unemployment risk, wage volatility, households with more members, and to loan motivations of "Pay previous debts" and "Health". In the year 2010 there was a substantially lower rate of default (3rd column of Table 8.1) and also higher loan amounts (2nd column of Table 8.1), even after accounting for the other factors in the model. This could have been due to the Chilean earthquake of 2010 and the economic expansion that followed the efforts of reconstruction in the country. The choice of loan amount and default behavior are both quadratic in terms of age, first increasing with age and then falling. Default is negatively related to income and to the "Purchase of Durables", but it is not significantly affected by education. Recent studies for the USA show that education, apart from math skills, has no significant impact on debt repayment behavior (Brown, van der Klaauw, Wen and Zafar, 2013), which is a result similar to our finding with Chilean data. Default is also negatively related to households with a higher taste for Bank loans and positively related with households with a propensity for "Banks plus Retail" and "No Access". This results justifies the behavior of Chilean banks in terms of giving preference to customers with a longer and more exclusive credit history in the banking system, since those are the households of lower risk. Credit history is also related to default in other countries as well (Gross and Souleles, 2002, Roszbach, 2004, Edelberg, 2006). In Chile some economists have suggested that the lack of public knowledge of retail loans by other lenders hurts the financial system (Alvarez, Cifuentes and Cowan, 2011), since banks are unable to account for loan applicants that have higher risk due to their retail loans. Our model finds that indeed borrowers with loans in both Banks and Retail have an unobservable higher risk of default, which confirms that a public database with information of all lenders could be a positive policy as suggested in the literature (Pagano and Jappelli, 1993). The positive effect of the unobserved factors related to "No Access" on the default probability can be seen as a confirmation that lender screening in the Chilean system works, since households with unobservable high risk have no access to debt.

Both the motives "Pay previous debts" and Health needs have a positive effect on loan amount and the propensity to default. It is also interesting that Health needs has a larger impact on default than the motive "Pay previous debts", since it confirms the predictions of economic models.

Exogenous variables	Log-loan amount $(t' = t - 4)$	Propensity to morosity $(t' = t)$
2007	-0.027 (0.055)	1.534 (0.178)***
2008 / 09	-0.001 (0.681)	$1.535 (0.195)^{***}$
2010	$0.245 \ (0.147)^*$	$0.791 \ (0.171)^{***}$
2011	0.05 (0.044)	$1.638 \ (0.142)^{***}$
Ratio of Debt to Income, $RDI_{i,t}$		$1.457 \ (0.105)^{***}$
Ratio of Debt Service to Income, $RDSI_{i,t}$		$0.467 \ (0.136)^{***}$
Log-Income: $\ln(\bar{P}_{i,t'})$	$2.575 \ (0.634)^{***}$	-0.400 (0.089)***
Years of education of home head	$-3.372 (1.955)^*$	-0.062(0.126)
Unemployment $\bar{u}_{i,t'}$	$0.071 \ (0.027)^{***}$	$2.871 \ (1.585)^*$
Wage volatility $\bar{\sigma}_{i,t'}$	$0.057 \ (0.021)^{***}$	$0.425 \ (0.253)^*$
Out of Santiago	-0.02(0.369)	-0.044 (0.113)
Age of home head	-0.371(0.367)	$0.109 \ (0.016)^{***}$
Age squared/ 100	$0.011 \ (1.349)$	$-0.139 \ (0.045)^{***}$
2 members in home	$0.099 \ (0.031)^{***}$	$0.071 \ (0.042)^*$
3 or more members	$0.126 \ (0.061)^{**}$	$0.390 \ (0.236)^*$
Share of loan for "Durables"	$0.003 \ (0.007)$	$-0.290 \ (0.173)^*$
Share of loan for "Pay previous debts"	$0.247 \ (0.097)^{***}$	$0.375 \ (0.188)^{**}$
Share of loan for Health needs	$0.075 \ (0.023)^{***}$	$0.736 \ (0.345)^{**}$
RE of loan type: $\bar{\varepsilon}_{i,b}$	heta	ρ
1 = Bank	$0.614 \ (0.227)^{***}$	$-1.127 \ (0.679)^*$
$2 = \mathrm{Bank} + \mathrm{Retail}$	$0.849 \ (0.489)^*$	$0.909 \ (0.113)^{***}$
3 = Retail	$0.045 \ (0.088)$	-0.241(0.184)
4 = Union	$0.621 \ (0.211)^{***}$	0.009(1.021)
5 = Other	$0.701 \ (0.794)$	-0.397(0.335)
6 = No Access	$0.006 \ (0.765)$	$0.749 \ (0.157)^{***}$
RE of log-amount: $\bar{\zeta}_i$		$0.062 \ (0.072)$

Table 8.1: Coefficients for the mean loan amount (in log) and propensity for morosity

Economic models of default decision assume that Health expenses are less predictable than other expenses, since tastes for consumption and past loan commitments are already known to the household. Therefore health expenses are an unpredictable shock for households and one that often leads to default even for low amounts of debt (Chatterjee et al., 2007).

Table 8.2 summarizes the estimated heteroscedasticity of the unobserved tastes for loan amount and default behavior, plus a contemporary loan amount shock which is independent over time. Again, we can conclude that households of higher income and education are less heterogeneous in their tastes for loan amount and default behavior. However, the heteroscedasticity of the

Exogenous variables	Log-loan amount:	Log-loan amount:	Propensity to morosity:
	Random Effect $\bar{\zeta}_i$	Contemporary shock $\widetilde{\zeta}_{i,t}$	Random Effect $\bar{\nu}_i$
constant	-0.306(0.489)	$0.055\ (0.127)$	-0.05(0.225)
2008/09	-0.434 (0.518)	$0.207 \ (0.275)$	$0.004 \ (0.035)$
$\frac{1}{T} \sum_{t=2007}^{2011} \ln(\bar{P}_{i,t})$	$-4.522 \ (0.938)^{***}$	$1.839 \ (0.447)^{***}$	$-0.712 \ (0.181)^{***}$
Education	$-3.107 (0.712)^{***}$	$1.134 \ (0.487)^{**}$	$-0.588 \ (0.218)^{***}$
$\frac{1}{T} \sum_{t=2007}^{2011} \bar{u}_{i,t}$	$0.107 \ (0.291)$	0.078(0.441)	$0.001 \ (0.027)$
$\frac{1}{T} \sum_{t=2007}^{2011} \bar{\sigma}_{i,t}$	$0.043 \ (0.054)$	$0.116\ (0.663)$	-0.007(0.013)

Table 8.2: Coefficients of the standard-deviation (in log) of the random-effects of loan amount and morosity

contemporary unobserved shock for loan amount is increasing with income and education. This shows that higher income and highly educated households are less persistent in their indebtedness, since their loan amounts depend more on contemporary shocks than constant factors.

6 Conclusions

This paper shows how households' characteristics impact their choice of consumer loans and default behavior. Low labor market risk (as measured by unemployment risk, job separation rates and wage volatility) is correlated with having no desire for consumer debt, while low income is the strongest cause of a lack of access to credit. Unemployment risk increases the probability of households opting for all types of consumer loans, but it has a greater impact on lenders who do not apply credit scoring such as Retail Stores and Credit Unions. Loan amounts increase with income, unemployment risk and wage volatility, therefore consumer loans may help smooth income shocks.

The default probability decreases with income and increases with high levels of indebtedness relative to income, unemployment risk and wage volatility, confirming the existence of unobservable factors in the adverse selection among Chilean debtors. Bank debtors have the lowest risk levels, which is expected from a lender that applies credit scoring extensively (Edelberg, 2006). However, borrowers with a high unobservable preference for loans with both Banks and Retail Stores have a higher probability of default. This is relevant for Chile, because there is no public database of all the loans of borrowers and therefore banks cannot know whether their customers have debts outside of the banking system (Álvarez, Cifuentes and Cowan, 2011). This may imply that there could be benefits in creating a public database providing common knowledge for all lenders in Chile (Pagano and Jappelli, 1993). At the same time I estimate that households with lower access to consumer loans have unobserved factors that are positively correlated with default, which implies that Chilean lenders are somewhat successful in screening out negative risks. Health needs are positively associated with default behavior, which could denote these expenses are difficult to predict and insure (Chatterjee et al., 2007). Finally, the probability of getting a loan and the choice of loan amount is increasing in the number of household members and quadratic in age, resembling the profile of life-cycle consumption (Attanasio and Weber, 2010).

Finally, I show that households are heterogeneous in their loan tastes, especially among the low education and income families. This result may imply that economic shocks or loan regulation (or deregulation) initiatives will have an heterogeneous welfare impact across the population.

References

- Adams, William, Liran Einav, and Jonathan Levin (2009), "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, 99(1), 49-84.
- [2] Alvarez, Roberto, Rodrigo Cifuentes and Kevin Cowan (2011), "Analysis of the potential effects of the law project of information disclosure of economic and financial commitments", Central Bank of Chile.
- [3] Attanasio, Orazio, Pinelopi Koujianou Goldberg, and Ekaterini Kyriazidou (2008), "Credit Constraints In The Market For Consumer Durables: Evidence From Micro Data On Car Loans," *International Economic Review*, 49 (2), 401-436.
- [4] Attanasio, Orazio and Guglielmo Weber (2010), "Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy," *Journal of Economic Literature*, 48(3), 693-751.
- [5] Banco Central de Chile (2013), "Financial Stability Report," 2nd Semester.
- [6] Bertrand, Marianne and Adair Morse (2011), "Information Disclosure, Cognitive Biases, and Payday Borrowing," *Journal of Finance*, 66, 1865-1893.

- [7] Brown, Meta, Wilbert van der Klaauw, Jaya Wen and Basit Zafar (2013), "Financial Education and the Debt Behavior of the Young", Federal Reserve Bank of New York Staff Paper 634.
- [8] Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima and José-Víctor Ríos-Rull (2007), "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default," *Econometrica*, 75(6), 1525-1589.
- [9] Dynan, Karen and Donald Kohn (2007), "The Rise in U.S. Household Indebtedness: Causes and Consequences," Finance and Economics Discussion Paper 2007-37, Federal Reserve Board.
- [10] Edelberg, Wendy (2004), "Testing for Adverse Selection and Moral Hazard in Consumer Loan Markets," Finance and Economics Discussion Paper 2004-9, Federal Reserve Board.
- [11] Edelberg, Wendy (2006), "Risk-based Pricing of Interest Rates for Consumer Loans," Journal of Monetary Economics, 53, 2283-2296.
- [12] Einav, Liran, Mark Jenkins, and Jonathan Levin (2012), "Contract Pricing in Consumer Credit Markets," *Econometrica*, forthcoming.
- [13] Eurosystem Household Finance and Consumption Network (2009), "Survey Data on Household Finance and Consumption: Research Summary and Policy Use", ECB Occasional Paper 100.
- [14] Girouard, N., M. Kennedy, and C. André (2007), "Has the Rise in Debt Made Households More Vulnerable?", OECD Economics Department Working Paper 535.
- [15] Gross, David B. and Nicholas S. Souleles (2002), "An Empirical Analysis of Personal Bankruptcy and Delinquency," *Review of Financial Studies*, 15 (1), 319-347.
- [16] Hess, Stephane, Kenneth Train and John Polak (2006), "On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a Mixed Logit Model for vehicle choice," *Transportation Research Part B: Methodological*, 40(2), 147-163.
- [17] Horowitz, Joel L. (2001), "The Bootstrap," in Handbook of Econometrics, edition 1, vol. 5, chapter 52, 3159-3228, Elsevier.
- [18] IMF (2006), "Household credit growth in emerging market countries", Global Financial Stability Report, 46-73.

- [19] Jaffee, Dwight M. and Thomas Russell (1976), "Imperfect Information, Uncertainty, and Credit Rationing," *Quarterly Journal of Economics*, 90(4), 651–666.
- [20] Jaffee, D. and J. Stiglitz (1990), "Credit rationing," in B. M. Friedman and F. H. Kahn (eds.) Handbook of Monetary Economics, vol. 2, Amsterdam: North Holland.
- [21] Kahneman, Daniel (2011), "Thinking, Fast and Slow," Farrar, Straus and Giroux Press.
- [22] Pagano, Marco and Jappelli, Tullio (1993), "Information Sharing in Credit Markets," Journal of Finance, 43(5): 1693–1718.
- [23] Madeira, Carlos (2014), "Earnings dynamics of Chilean workers: employment transitions, replacement ratio, and income volatility," Central Bank of Chile, mimeo.
- [24] Marinovic, Alejandra, José Matus, Karla Flores, Nancy Silva (2011), "Systemic risk of households in Chile", *Revista de Economía Chilena*, 14(3), 5-39.
- [25] McFadden, Daniel and Kenneth Train (2000), "Mixed MNL models for discrete response," Journal of Applied Econometrics, 15(5), 447–470.
- [26] Nevo, Aviv (2000), "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," Journal of Economics & Management Strategy, 9(4), 513-548.
- [27] Nevo, Aviv (2011), "Empirical models of consumer behavior," Annual Review of Economics, 3, 51-75.
- [28] Roszbach, K. F. (2004), "Bank Lending Policy, Credit Scoring and the Survival of Loans," *Review of Economics and Statistics*, 86(4), 946–958.
- [29] Shimer, Robert (2012), "Reassessing the Ins and Outs of Unemployment," Review of Economic Dynamics, 15(2), 127-148.
- [30] Train, Kenneth (2009), "Discrete choice methods with simulation," Cambridge University Press.
- [31] Vella, Frank (1998). "Estimating Models with Sample Selection Bias: A Survey", The Journal of Human Resources, 33(1), 127-169.

Documentos de Trabajo Banco Central de Chile	Working Papers Central Bank of Chile
NÚMEROS ANTERIORES	PAST ISSUES
La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica:	Working Papers in PDF format can be downloaded free of charge from:
www.bcentral.cl/esp/estpub/estudios/dtbc.	www.bcentral.cl/eng/stdpub/studies/workingpaper.
Existe la posibilidad de solicitar una copia impresa con un costo de Ch\$500 si es dentro de Chile y US\$12 si es fuera de Chile. Las solicitudes se pueden hacer por fax: +56 2 26702231 o a través del correo electrónico: <u>bcch@bcentral.cl</u> .	Printed versions can be ordered individually for US\$12 per copy (for order inside Chile the charge is Ch\$500.) Orders can be placed by fax: +56 2 26702231 or by email: <u>bcch@bcentral.cl</u> .

DTBC - 837

The Persistent Effect of a Credit Crunch on Output and Productivity: Technical or **Allocative Efficiency?** Patricio Toro

DTBC - 836

Sectoral Transitions Between Formal Wage Jobs in Chile Rosario Aldunate, Gabriela Contreras, Matías Tapia

DTBC - 835

Misallocation or Misspecification? The Effect of "Average" Distortions on TFP Gains **Estimations**

Elías Albagli, Mario Canales, Antonio Martner, Matías Tapia, Juan M. Wlasiuk

DTBC - 834

Forecasting Inflation in a Data-rich Environment: The Benefits of Machine Learning **Methods**

Marcelo Medeiros, Gabriel Vasconcelos, Álvaro Veiga, Eduardo Zilberman

DTBC - 833

XMAS: An Extended Model for Analysis and Simulations

Benjamín García, Sebastián Guarda, Markus Kirchner, Rodrigo Tranamil

DTBC - 832

Extracting Information of the Economic Activity from Business and Consumer Surveys in an Emerging Economy (Chile) Camila Figueroa y Michael Pedersen

DTBC - 831

Firm Productivity dynamics and distribution: Evidence for Chile using micro data from administrative tax records

Elías Albagli, Mario Canales, Claudia De la Huerta, Matías Tapia y Juan Marcos Wlasiuk

DTBC-830

Characterization of the Recent Immigration to Chile Rosario Aldunate, Gabriela Contreras, Claudia De la Huerta y Matías Tapia

DTBC - 829

Shifting Inflation Expectations and Monetary Policy

Agustín Arias y Markus Kirchner

DTBC-828

Macroeconomic Effects of Credit Deepening in Latin America Carlos Carvalho, Nilda Pasca, Laura Souza y Eduardo Zilberman

DTBC - 827

Determinants of Household Position within Chilean Wealth Household's Distribution Felipe Martínez y Francisca Uribe

DTBC - 826

Revisiting the Exchange Rate Pass Through: A General Equilibrium Perspective Mariana García-Schmidt y Javier García-Cicco

DTBC - 825

An Econometric Analysis on Survey-data-based Anchoring of Inflation Expectations in Chile Carlos A. Medel

DTBC - 824

Can Economic Perception Surveys Improve Macroeconomic Forecasting in Chile? Nicolas Chanut, Mario Marcel y Carlos Medel

DTBC - 823

Characterization of the Chilean Financial Cycle, Early Warning Indicators and Implications for Macro-Prudential Policies Juan Francisco Martínez y Daniel Oda

DTBC - 822

Taxonomy of Chilean Financial Fragility Periods from 1975 Juan Francisco Martínez, José Miguel Matus y Daniel Oda DTBC - 821

Pension Funds and the Yield Curve: the role of Preference for Maturity Rodrigo Alfaro y Mauricio Calani



BANCO CENTRAL DE CHILE

DOCUMENTOS DE TRABAJO • Septiembre 2019