Madeira *Financial Innovation* (2023) 9:49
https://doi.org/10.1186/s40854-023-00458-6

Financial Innovation

RESEARCH Open Access

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

Carlos Madeira*

*Correspondence: carlosmadeira2009@u.northwestern.edu

Central Bank of Chile, Santiago, Chile

Abstract

Why do households use different types of loans? Which factors cause borrowers to default? Using a comprehensive survey dataset from Chile, I estimate a partial information model of consumer debt access, lender choice, loan amount and default. The model consists of a first-stage multinomial logit that explains the choice across the five loan types, plus the options of no access to debt due to credit constraints and a no wish for consumer debt. In the second and third stages, the model assumes a log-linear regression of the debt amount and a logit regression of the default behavior, accounting for the loan type selection probability. Identification is obtained using factors measured at different time periods for the default and the loan type choices. I find that households choose different lenders based on income, education and labor risks. Higher income and education decrease the probability of credit constraints, while increasing bank lending and debt amounts. Unemployment risk and household size increase the chances of all the loan types; however, unemployment decreases the debt amount. Age and wage volatility reduce the probability of all loans. Default decreases with income, education and age, whereas it increases with indebtedness, unemployment, household size, health shocks, and paying previous loans. Counterfactual exercises demonstrate that pension reform, higher requirements for borrowers' capacities, and financial literacy programs could substantially reduce default risk. Financial literacy could greatly reduce arrears, households with credit constraints, the number of debtors and the aggregate debt amounts, especially for non-bank lending.

Highlights

Chilean borrowers present heterogeneous adverse selection across lender types.

No Debt Access decreases with income, age, education, but it increases with risk.

Default is associated with income, unemployment, indebtedness and demographics.

Paying past loans and health needs are associated with indebtedness and default.

Financial literacy programs may be a powerful policy to improve the debt market.

Keywords: Consumer credit, Default risk, Unemployment, Financial literacy, Adverse selection, Credit constraints

Introduction

Household debt has consistently increased in recent decades, both in emerging economies and developed countries (Edelberg 2006), with strong financial innovation offering consumers a wide range of loan products (Agarwal et al. 2020). Many consumers can access credit from a variety of sources, such as credit cards, auto loans, or education loans, and with motives as diverse as health, vacations, purchase of durable goods, or a renegotiation of previous debts. However, the question of lender choice remains understudied, especially in developing countries where consumer credit is widely used, often at high interest rates (Madeira 2019b) and high default risk (Madeira 2018).

This paper examines the access to consumer credit, choice of lender, loan amount decision and repayment behavior of families in Chile. Consumer debt is defined as unsecured loans requested for consumption (unlike mortgages which are secured by real properties). 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 (in Spanish, *Encuesta Financiera de Hogares*; 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 among different lenders according to their observable characteristics.[1] Furthermore, household debt levels, income, and labor market risk significantly impact default behavior. As expected, lenders offer higher loan amounts to households with better observable characteristics, but unobservable risks (known to borrowers but not to lenders, such as debt motives) make riskier borrowers borrow and default more.

The loan market in Chile has many providers, which represent imperfect substitutes for consumers. These loan providers have access to different customer information and are subject to different regulations, which affects their loan terms and the ability to target borrowers. Loan screening costs, asymmetric information, brand preferences, marketing, search, travel costs, and tied products, can create substantial frictions for customers, so loan choice and repayment are modeled as a differentiated product choice (Hensher et al. 2015).

The model of loan choice and repayment behavior has three main components: (i) the categorical choice between having no wish for debt, wanting debt but having no access, and five lender types; (ii) the loan amount decision; and (iii) the categorical outcome of whether the household defaulted or not on their debt. The five lender types correspond to: Banks, Banks and Retail Stores, Retail Stores, Union Credit (i.e., loans provided by credit and labor unions), and Other Loans (including 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 from the option of using only one type of lender. Other types of lenders represent a small proportion of the population and therefore are not modeled. The option of "No Access to Debt" represents families with credit constraints. These are families who applied for consumer loans but were denied credit and those who wished to apply for credit but did not do so because

[1] Families are sorted among different lender types, either through self-selection according to their preferences or through screening (that is, selection from the lenders), most likely through a mix of both channels. In general terms, sorting is a methodology which groups individuals and solves a problem through a separating equilibrium. Sorting is especially useful in credit market problems, due to moral hazard and adverse selection (Jaffee and Stiglitz 1990).

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.

Debt choices are affected by both observable and unobservable factors. The observables include income, education, labor income risk (measured by unemployment risk and wage volatility), the motives behind the loans, and demographic characteristics such as the age of the household head and household size. Unobservable factors, such as random-effects, include idiosyncratic tastes of households or bargaining conditions that affects preferences for all types of debt and banking debt. The lender type choice, modeled as a Mixed Multinomial Logit Model, is estimated by Simulated Maximum Likelihood (Hensher et al. 2015). The subsequent loan amount decision and the default behavior are estimated as second-step regressions, with the inverse Mill's ratio of the lender type choice included as additional control variables (Vella 1998).

I demonstrate that households with different income, education, debt, unemployment risk and wage volatility are sorted among different lenders. Banks resort more to credit scoring and customer specific interest rates, therefore selecting the households of highest income and education and with the lowest unemployment rates among loan applicants. Bank plus Retail and Other debt borrowers are younger and have larger families, therefore they have the largest amount of debt. Their income is slightly lower and their unemployment risk a bit higher than those of the Bank-only borrowers. Households with No Access to Debt have the lowest income and education levels, while experiencing the highest unemployment risk and wage volatility, which excludes them from the consumer loan market. The Union Debt borrowers are a case in the middle of the debt and labor income risk profiles, because these households have a low income level but also a low wage volatility. Therefore, these households also have a moderate level of indebtedness. Finally, among the households with access to debt, Retail Store only borrowers show the lowest income and education coupled with the highest unemployment risk. Therefore, such households have the lower debt amounts.

Both education and income increase the probability of Banks, Banks-Retail and Other Debts, while decreasing the probability of No Access to Debt. Unemployment risk and household size increase the probability of opting for any loan type, while age and wage volatility decrease the probability of any loans. Unemployment risk has a greater effect on increasing the Retail Store and Other Debt probabilities. Wage volatility increases the probability of No Access to Debt.

Loan amounts increase with income, education and debt motivated by Durables, Pay Other Debts and Health needs, consistent with life cycle consumption patterns (Attanasio and Weber 2010). Loan amounts are negatively associated with unemployment risk, age and the share of household debt financed by a single lender type. The probability of default decreases with income, education and age, while increasing with high indebtedness ratios, unemployment risk, household size, loans motivated by Paying Previous Debts and Health needs.

The remainder of this paper is organized as follows. "Past research and credit market institutions in Chile" section summarizes the credit literature and Chile's regulatory environment. "An empirical model of choice of lender, loan amount and default" section describes the model of loan choice and default. "Data: the Chilean Household Finance Survey (EFH)"

section summarizes the Chilean Household Finance Survey. "Characterization of borrowers' loan options" section describes the households across loan types. "Results" section presents the results of the model of lender choice, loan amount and default. "Counterfactual simulations of policies" section estimates the counterfactual effect of different policies on households' choices. Finally, "Conclusions and policy implications" section concludes the study with a summary of the results and policy implications.

Past research and credit market institutions in Chile

Literature review

This study is related to a growing literature of empirical models of household finance. It relates to the study of borrower heterogeneity and its impact on credit constraints, loan choice and default behavior (Einav et al. 2012). It is especially related to empirical studies of household budgeting behavior (Habibah et al. 2018; Hamakhan 2020), default choices (Tang et al. 2020) and micro-founded applications of financial policies (Ahmed et al. 2018). It is also related to the studies that use household finance surveys to inform on credit frictions and policy (Eurosystem 2009; Gallardo and Madeira 2022), with these surveys being particularly important to obtain a complete view of borrowers' loans because many households rely on a diversity of funding sources, including bank and non-bank lenders (Cull et al. 2019; Madeira 2018). Due to their complexity and cost, complete household finance surveys (comparable to the Survey of Consumer Finances in the USA or the eurozone Household Finance and Consumption Survey) are especially scarce in emerging markets and developing countries, having only started in Chile, China, Uruguay and Mexico in 2007, 2011, 2014 and 2019, respectively.

Other studies have followed the goal of classifying consumers or SMEs (Small and Medium Enterprises) in terms of their bankruptcy, credit or fraud risks using large datasets from a financial services provider such as payment and transactions networks, banks or credit card managers (Li et al. 2022). It is relevant to note that the credit risk literature differs substantially for SMEs and households (which are the final consumers). SMEs default or bankruptcy decisions depend on the accounting profitability, cashflow, or net present value of their investments (see Kou et al. 2014 and references therein). However, households tend to be more influenced by short-term shocks such as unemployment or health (Eurosystem 2009; Gallardo and Madeira 2022). Unlike SMEs, which can decide to implement bankruptcy or exit a market, households in Chile cannot declare bankruptcy (similar as in many other countries). Furthermore, in many countries the SMEs can be much larger than households or consumers, with SMEs having up to 250 employees in the European Union and up to 500 employees in the USA (except for very rich households, most of the population tends to be employees, self-employed or owners of micro firms with 4 employees or less). This research article instead uses a small survey dataset with a range of variables that inform the factors behind the consumer decisions and which economic theories such as permanent income, precautionary behavior, lifecycle motives, or adverse selection, to explain consumer behavior.

Households borrow because of habit formation, life cycle changes, temporary fluctuations in income, consumption smoothing, and behavioral biases such as a lack of

financial literacy and temptation (Attanasio and Weber 2010; Agarwal et al. 2020), with excessive debt being associated with psychological and economic distress (Agarwal et al. 2020). Risk-based interest pricing, loan caps and higher down payments can reduce the adverse selection and moral hazard of borrowers (Einav et al. 2012). Risk-priced credit supply can reduce liquidity constraints and increase product demand, because riskier consumers are more sensitive to credit terms than to car prices (Einav et al. 2012). Finally, lender screening can control the borrowers' adverse selection and moral hazard by offering different contracts (Jaffee and Stiglitz 1990). In summary, loan markets should present three results (Einav et al. 2012): (i) lenders will offer better and larger loans to agents with observable characteristics of low risk; (ii) unobservable high risk characteristics will be associated with both larger loan amounts and default; and (iii) agents with very high risk will be credit constrained.

This study is relevant, because it includes a substantial amount of information about the borrower, its loan application, its family and the labor risk factors of its members. The article, therefore, provides a rich perspective of the credit markets in a developing middle income country. Research on consumer preferences between different alternative products and default risk is increasingly relevant in recent years. Technological development increased the menu of financial products and the available information for credit risk scoring (Thakor 2020), with several banks across the world investing in FinTech to expand their services and products (Kou et al. 2021).

This article extends the literature in five ways. First, it studies loan choice and default in a middle income country such as Chile (while most studies have been either with advanced economies, see Badarinza et al. (2016), or with low income countries, see Banerjee et al. (2015)). Second, it introduces a wider range of loan options, with borrowers being able to choose from banks, retail, credit unions or other lenders (such as auto sellers), whereas most previous studies have been limited to specific markets such as the banking sector, auto loans (Einav et al. 2012) and payday loans (Bertrand and Morse 2011), or a specific lender (Einav et al. 2012). Third, it uses a self-reported measured of households that do not wish for debt and it also includes households that expected to be rejected as being credit constrained. For instance, in administrative datasets, such as Einav et al. (2012), the researchers only observe the borrowers who applied for loans and were rejected, but there is no identification for those who did not apply thinking that a rejection was likely. Furthermore, such datasets often include only the loans of the borrowers with a single company and it is not measured whether the consumers obtained loans from other sources. Fourth, it measures unobserved preferences for loans by using tools from the applied product choice models in the field of industrial organization (Hensher et al. 2015). Fifth, it uses a more diverse characterization of labor income risk by separating overall risk into different variables such as unemployment risk and wage volatility, with such risks being measured for all the members of the household and not just the borrower. For instance, many studies of default risk only use the unemployment status of the borrower, although the spouse and other household members also present income shocks and risk factors. Also, some past studies such as Ampudia et al. (2016) consider unemployment, but not wage volatility.

The structure of consumer loan providers in Chile

In Chile all lenders have access to a commercial registry of debtors who defaulted on payments.[2] However, this registry is limited only to negative events, and therefore, lenders' information sets on the positive characteristics of loan applicants differ substantially (Cohen and Dijkman 2021).

Additionally, Chilean banks have access to a common credit registry with information on all loan amounts and debt default within the banking system. However, banks do not observe non-banking loans (Cohen and Dijkman 2021). Banks also make strong use of credit scoring, according to agents' history of credit and other products, such as direct deposit of wages, automatic bill payment or mortgages.

Credit unions (denoted as Savings and Loans' cooperatives[3]) and labor unions (denoted as Family Compensation Funds[4]) are another lender type, regulated as providers of "social credit". According to the legislation, all Chilean companies must register their workers with one among several Family Compensation Funds, which offer 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, therefore, workers do not choose their institution directly. 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 when a debtor loses its job, Family Funds may deduct a substantial payment from the workers' severance pay. Therefore, their risk is limited even in the face of unexpected unemployment events. Union credit providers must offer the same conditions to all members. Therefore, unions can change interest rates according to loan size and maturity, but are unable to differentiate debtors according to their characteristics such as income. Furthermore, Family Compensation Funds cannot offer other financial services (for example, insurance or checking accounts) like banks do and cannot sell ordinary consumer appliances such as retail stores. This may put unions at a disadvantage in approaching customers if the search frictions and travel costs are significant. Finally, Family Compensation Funds cannot grow their assets through deposits, corporate debt or equity, which limits their size relative to banks and retail stores.

Retail stores are another type of credit provider, with a strong brand image and their own credit cards,[5] and access their own private databases of customers. Retailers provide few cash advances, but their credit cards are widely accepted by several stores, utility companies and merchants. Finally, there are lenders with more specific goals, such as auto loans at car dealers, education loans, pawn shops[6], and consumer loans provided by insurance companies.[7]

[2] See www.dicom.cl/.

[3] See 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.

[4] 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.

[5] The 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.

[6] See www.dicrep.cl.

[7] The 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 2006, banks represented 62.1% of the total consumer credit, while social credit institutions represented 13.1% and retail stores 24.8% of the market, respectively.[8] However, the market presence in terms of the number of customers differs from the aggregate loan amounts. 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 for 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 (Madeira 2019c). Aggregate consumer credit by social institutions in 2013 was 245% the level of 2006, but retail store credit grew only 57% during the same period. Financial time series from the Central Bank of Chile show that, from the end of 2013 until the end of 2019, the real value of consumer loans from banks, retail stores and unions, education, vehicle sellers and other debts grew 4.2%, 4.4%, 10.2%, 5.5% and 1.4%, respectively. However, following the Covid pandemic in 2020 and 2021, the consumer credit for banks and retail-unions dropped by 11.3% and 4.4%, respectively, whereas the other lender types reported almost zero growth.

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

The consumer choice model considers three endogenous variables: (i) a categorical choice between having no debt, wanting debt but being credit constrained,[9] and five lender types; (ii) the loan amount decision; and (iii) a categorical outcome of whether the household defaulted on its debt. The five lender types correspond to: Banks, Banks and Retail Stores, Retail Stores, Union Credit, and Other Loans (which includes mainly auto loans, educational debt, pawn shops and informal lending). For simplicity, I classify the observed lender choice of each household as the one corresponding to the largest loan amount reported by each family.[10] Banks and Retail Stores are the two major lenders in Chile, therefore, using both lenders with positive debt is treated as a separate choice. After considering the households' largest loan amount among these 5 options, the remaining debt amount is quite negligible. For at least 75% of the borrowers, the choice among these 5 options represents 100% of its consumer debt. Therefore, there is little practical interest to increase the model by adding more options that mix specific lenders such as Banks with relatively rare options of "Other Loans" such as education or pawn loans (which represent a small number of debtors and share of the consumer debt).

Let $U_{i,b,t}$ denote the utility of household i from the option b in period t, with $b \in \{1$ "Bank", 2 "Bank & Retail", 3 "Retail", 4 "Union", 5 "Other Loans", 6 "No Access"$\}$, with the utility of the outside option, "No wish for Debt", standardized

[8] The 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.

[9] 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 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.

[10] Some consumers may 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 and such options represent a negligible amount of the consumer debt.

as zero, $U_{i,0,t} = 0$. Consumer chooses the option $Y_{i,t} = b$ with the highest utility ($U_{i,b,t} \geq \max(U_{i,0,t}, U_{i,1,t}, .., U_{i,B,t})$) and then a loan-amount $L_{i,t}$, which are affected by observable characteristics, $x_{i,t0}$, plus unobservable preferences for each loan type b, $\varepsilon_{i,b,t}$, and loan-amount, $\zeta_{i,t}$. The observed characteristics $x_{i,t0}$ correspond to a previous period $t0 < t$, because the debt choice was made in the past and not at the exact time of the survey interview. The utility of each loan type be a linear function of the observables and the error term is:

$$U_{i,b,t} = \alpha_{b,t} + \beta_b x_{i,t0} + \varepsilon_{i,b,t}. \tag{1}$$

Sometimes, it is difficult to interpret the coefficients of a multivariate choice model (Hensher et al. 2015), due to the agents' $B + 1$ possible alternatives. For a given choice b, $\beta_b > 0$ implies that the probability of option b relative to option 0 is increasing in x. However, 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 $\beta_b > 0$ only unambiguously increases the probability that b is chosen with larger x if $\beta_b \geq \max(\beta_1, \ldots, \beta_B)$.

The unobserved terms $\varepsilon_{i,b,t}$ may include a diverse set of factors, such as idiosyncratic preferences, geographical distance, marketing influence, the burden for loan applicants to provide the information and legal documents requested by certain lenders, contractual costs of loans such as penalty charges or insurance fees, and negative events during the loan bargaining process.[11] McFadden and Train 2000 show that under mild regularity conditions, any discrete choice model derived from random utility maximization has choice probabilities that can be approximated as closely as possible by a Mixed Multinomial Logit model. Therefore, the unobserved tastes $\varepsilon_{i,b,t}$ are specified as the sum of an independent extreme valued component and normal heteroscedastic random-effects that are correlated over different choices (McFadden and Train 2000):

$$\varepsilon_{i,b,t} = 1(1 \leq b \leq 5)\eta_{i,1,t} + 1(1 \leq b \leq 2)\eta_{i,2,t} + \tilde{\varepsilon}_{i,b,t}, \tag{2}$$

with 1(.) being the indicator function, $\tilde{\varepsilon}_{i,b,t} \sim EV(0,1)$ and $\eta_{i,a,t} \sim N(0, \sigma_{\eta_a})$. $\eta_{i,1,t}$ is a random factor denoting agent i's taste for any type of loan, while $\eta_{i,2,t}$ is a random factor denoting agent i's taste for both the Bank and Bank plus Retail options. This model is then estimated using Simulated Maximum Likelihood with 250 Hammersley integration points (Hensher et al. 2015).

After estimating the loan type choice probabilities, the corresponding inverse Mill's ratio ($\frac{f(x)}{\Pr(Y_{i,t}=b|x_{i,t0})} = \Pr(Y_{i,t} \neq b \mid x_{i,t0})$ in the Multinomial Logit) is used in the second step regression to correct for the loan type selection in the total consumer debt amount[12] model:

$$\ln(L_{i,t}) = \pi_t + \delta x_{i,t} + \sum_{b=1}^{B-1} \theta_b 1(Y_{i,t} = b) \Pr(Y_{i,t} \neq b \mid x_{i,t0}) + \zeta_{i,t}. \tag{3}$$

[11] Hensher et al. (2015) give further interpretation about unobserved preferences of multiple choice models.

[12] Considering the total consumer debt helps account better for the default risks in the final stage of the model.

These second step equations are identified semi-parametrically, with no need to specify the distribution of the unobservable factor $\zeta_{i,t}$ (Vella 1998). All that is required is at least one variable that affects loan choice but not the default (which is valid because $x_{i,t0}$ differs from $x_{i,t}$ and $z_{i,t}$ due to the lagged variables), and the assumption that the unobservables are uncorrelated with the vector variables $x_{i,t}$ after accounting for the loan type selection probability, $1(Y_{i,t} = b)\Pr(Y \neq b \mid x_{i,t0})$.

The decision to default at time t, $D_{i,t} \in \{0, 1\}$, is then given by whether the latent propensity to default is positive, $d_{i,t} > 0$, expressed by a linear function of the observable characteristics, $z_{i,t}$, the inverse Mill's ratio for the loan type decision $(1(Y_{i,t} = b)\Pr(Y_{i,t} \neq b \mid x_{i,t0}))$, and an unobserved shock $\nu_{i,t}$ (which, for simplicity, is assumed to be extreme-value distributed, $\tilde{\nu}_{i,t} \sim EV(0, 1)$):

$$d_{i,t} = \mu_t + \lambda z_{i,t} + \sum_{b=1}^{B-1} \gamma_b 1(Y_{i,t} = b)\Pr(Y_{i,t} \neq b \mid x_{i,t0}) + \nu_{i,t}. \tag{4}$$

The empirical variable measure of debt default is determined by whether the household is in arrears for any of its consumer loans, with arrears implying that the household is late on the payments of at least one of its consumer loans for one or more months. Some households can have multiple loans with the same lender or with different lenders. However, an event of arrears is considered to be a serious sign of delinquency of the borrower, with many borrowers ending up defaulting on their other unsecured consumer loans later on (Madeira 2018). For this reason, it is appropriate to treat any loan arrears as a delinquency event of the borrower. Finally, treating the consumer delinquency at the borrower level (a dummy for arrears in any of its consumer loans) is more appropriate econometrically, because the loan by loan decisions of the same borrower cannot be treated as statistically independent since the same factors affect all loans of the borrower.

The observable variables that explain default, $z_{i,t}$, differ from the vector explaining loan types, $x_{i,t}$. This is because loans have a maturity of several periods; therefore, the decision of loan choice and amounts happens before the repayment period, with loan choice being affected by the income and unemployment risk in the past year and default behavior by the current variables.

The explanatory variables for explaining the lender type choice include demographics, income and job risks (Attanasio and Weber 2010; Agarwal et al. 2020). Permanent income influences both household loan demand and its credit worthiness (Attanasio and Weber 2010; Agarwal et al. 2020). Age and household size account for life cycle components that affect the demand for durable goods and other expenditures (Attanasio and Weber 2010). Unemployment risk and wage volatility account for the large uninsurable shocks that can affect loan demand when the loan choice is made (Attanasio and Weber 2010) and delinquency behavior at the time of the survey measurement (Agarwal et al. 2020). Education proxies for financial literacy (Attanasio and Weber 2010; Badarinza et al. 2016), whereas the region proxies for differential access to lenders outside the capital and the survey wave dummies account for aggregate shocks (Wooldridge 2010).

The loan amount also includes the selection Mill's ratios (Vella 1998) to account for borrowers' self-selection with each lender and the debt motives known by the borrower but unseen by the lender. Finally, the final default behavior also controls for measures of

borrower liquidity and solvency (Einav et al. 2012), such as the debt service to monthly income ratio (DSIR) and the consumer debt to the annual permanent income ratio (CDPIR).

Similar results to those in this parametric partial information model are confirmed by the analysis of: (i) the households' mean statistics in section 5 of this article; (ii) the non-parametric probability density function (PDF) and cumulative distribution function (CDF) distributions of households' loan levels and labor risks[13]; (iii) a parametric full information model (see Madeira 2019c); iv) different time periods, such as just the 2007–2011 waves (Madeira 2019c).

Data: the Chilean Household Finance Survey (EFH)

This study uses data from the Chilean Household Finance Survey (in Spanish, *Encuesta Financiera de Hogares*, hence on EFH), which is a representative cross-sectional survey with detailed information on households' assets, debts, income and financial behavior. The EFH survey is broadly comparable to similar surveys in the United States and Europe (Eurosystem 2009). The EFH covered a total of 21,319 urban households from 2007 to 2017. The 2007, 2011, 2014 and 2017 waves were implemented at the national level, including all the regions of Chile, with an interval of three or four years between each wave, having interviewed 3828, 4059, 4502 and 4549 households, respectively. The 2008, 2009 and 2010 waves were implemented in the capital city of Santiago (which has over 40% of the national population); therefore, those waves only interviewed 1154, 1190 and 2037 households, respectively.[14] All sample statistics in the article use expansion factors and are representative.

The EFH has a particularly detailed focus of the loans and debt commitments of each household. It asks for the largest three debts that each household has for each type of loan among a total of 13 categories: 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 current debts of the household, although obviously few agents report having debts with all possible categories of loans. Although the survey indicates which household member signed each loan, this article's analysis follows the standard that considers that all household members share expenses and are co-responsible for debts (Eurosystem 2009; Attanasio and Weber 2010).

In addition to asking for the details (debt amount, monthly loan payment, maturity, interest rate) of each loan, the EFH also inquires about its credit applications over the previous year: (i) "Has the family made any loan request over the previous 12 months?", (ii) "How many requests for loans has the family made?", (iii) "How many loan requests have been rejected?", and (iv) "What is the main reason why your family did not make any loan requests? Options: (1) No need for credit. (2) Dislikes loans. (3) The family

[13] These results are available as figures from the author upon request. Similar figures are also available for the waves between 2007 and 2011 in the earliest draft of this article (Madeira 2019c).

[14] The study uses all the waves as a pooled cross-section to increase sample size and efficiency (Wooldridge 2010).

would be unable to repay the loan. (4) The family would be rejected for the loan requests or the loan would not be granted. (5) Other reasons.". The last question can separate the families with "No Wish for Consumer Debt" (those that report the options 1 and 2 for either no need for credit or disliking loans) and those with "No Access to Debt" (reporting wishing to have debt, but not making loan requests due to options 3 and 4 of being unable to repay the loans or likely to be refused for loans even if a request was made). Families are also classified as "No Access to Debt" if they made loan requests, but all of them were refused.

I represent the size of each household using two measures. The first is the total number of household members n_i. Other measures for the household size can account for scale economies in terms of the consumption of joint goods within the household. Therefore, I also use the OECD-modified scale for the household size (OECD 2008), which assigns a value of 1 to the household head, 0.5 for each additional adult member (above age 15) and 0.3 for each child: $ne_i^{OECD} = 1 + 0.5(adults_i - 1) + 0.3(children_i)$.

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

Using the quarterly Chilean Employment Survey, which covers 35,000 households, I obtain two measures of labor risk for the period 1990–2017 (Madeira 2015): the unemployment rate ($u_{k,t} = \Pr(U_{k,t} = 1 \mid t, x_k)$) and the labor income volatility even if no job is lost, $\sigma_{\zeta,t}(x_k) = \sqrt{E\left[(Y_{k,t} - E[Y_{k,t} \mid Y_{k,t-1}, x_k])^2 \mid V_{k,t}, x_k\right]}$, with $V_{k,t} = (t, U_{k,t} = U_{k,t-1}, Y_{k,t})$. The vector x_k creates 540 mutually exclusive groups, given by $x_k = \{$Santiago Metropolitan city or Outside, Industrial Activity (primary, secondary, tertiary sectors), Gender, Age (3 brackets, ≤ 35, $35 - 54$, ≥ 55), Education (less than secondary schooling, secondary or technical education, college), and Household Income quintile$\}$.

Using these labor risk measures, I calculate the expected income $\bar{P}_{i,t}$ 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 the 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 the worker k's average labor income in the employed and unemployed states. The employment risk and labor income volatility of each household are then given by a weighted average of the rates of each member (using as weights their labor income relative to the total household labor income): $\bar{u}_{i,t} = \sum_k \frac{P_{k,t}}{P_{i,t}} u_{k,t}$ and $\bar{\sigma}_{i,t} = \sum_k \frac{P_{k,t}}{P_{i,t}} \sigma_{\zeta,t}(x_k)$.

Characterization of borrowers' loan options

Borrowers according their type of lender

Table 1 illustrates the proportion of households that chose each of the five lender types, plus those 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). Retail Stores are the most popular choice among households, representing more than 40% of the population, with 26% being Retail Store only users and 14% being users of both Bank and Retail Store Loans. Banks are the second largest lender in terms of the number of borrowers,

reaching 23.4% of the households, with 9.4% being Bank only and 14% being Bank plus Retail borrowers. Finally, Union lenders and Other Debts reach, respectively, 6.1% and 6.4% of the households. Over the last ten years (that is, from the first wave in 2007 until the last wave in 2017), the users of Bank debt only and Other debts increased significantly in each year. Union borrowers fluctuated substantially over the years, increasing from 3.8% in 2007 to 8.1% in 2011, before falling in 2014 and increasing again to 7.2% in 2017. Retail Store debt only and No Access consumers decreased significantly over the last ten years, with Retail Store only declining mostly from 2007 to 2011 and again from 2014 to 2017, while the fall in No Access households was moderate from 2011 to 2014 and then had a stronger drop from 2014 to 2017. Households choosing Retail Stores only (that is, without bank loans) decreased over time from 31.9% in 2007 to just 19.4% in 2017. However, borrowers choosing loans at Banks only increased from 6.5% in 2007 to 10.9% in 2017. Other debts also increased steadily, from 4.6% in 2007 to 10% in 2017.

The proportion of households without a wish for consumer debt represents 27% of the Chilean population, while those with No Access to Debt represent 11% of the population. Over time, the fraction of household with No Access to Debt steadily decreased from 12.7% in 2007 to 8.9% in 2017. The consumers with No Wish for Debt fluctuated over time, increasing between 2007 and 2011, then dropping in 2014 and increasing again recently in 2017 to a higher level of 30.4%.

There is limited information in the EFH about the loan interest rates. The survey asks about the loan's interest rates, but less than half the respondents report to remember this information. Furthermore, it is not possible to include such information in the multiple choice model, because the survey does not ask the interest rates and maturities for the loan options not chosen by the consumers. Table 2 shows the distribution of interest rates reported by the borrowers from the waves between 2007–2011 and from 2014 to 2017. The analysis is separate into two periods, because a 2013 legislation in Chile reduced the maximum interest rate for consumer loans from 52% to just 37% (Madeira 2019b). Looking at the top interest rates (as expressed by the percentile 75), all lenders reduced their interest rates between 2007–2011 and 2014–2017, although only the Bank and Auto lenders reduced their median interest rates. Retail, Union and Auto lenders also reduced their loan maturity (in terms of median and mean) between 2007–2011 and 2014–2017. Median maturity fell by 12 months for Union and Auto Lenders. The worst fall in maturity was for Retail, which dropped their median maturities by 24 months. Banks kept similar median maturities before and after the 2013 interest rate, although the mean was lower by 8 months. Education lenders, however, increased their maturity substantially, perhaps due to government interest subsidization and guarantees against individual borrowers' default.

Before 2011 the most expensive consumer loans were those obtained from Retail Stores and Auto Lenders, while those for Education and Other Loans were the most expensive loan product after 2014.[15] Prior to 2011, Bank loans were the second cheapest

[15] One reason why Education and Other Loans changed interest rate conditions so dramatically between periods is that public education and a debt free college education was one of the major political flags of the Chilean government of Michelle Bachelet, which started in March of 2014. The government promised a debt free college education, which reduced demand for education loans. This government agenda was materialized by the laws number 21040 (November, 2017), 21091 (May, 2018) and the regulatory decree 333 (November, 2019).

Table 1 Fraction of households (in %) choosing each type of debt over the years

Type of debtor	All waves	2007	2011	2014	2017
Bank	9.4	6.5	8.4	10.7	10.9
Bank + Retail	14.0	14.1	11.9	16.8	13.2
Retail Store	25.8	31.9	25.9	26.8	19.4
Union Debt	6.1	3.8	8.1	5.7	7.2
Other Debts	6.4	4.6	5.1	5.8	10.0
No Wish for Debt	27.0	26.4	27.9	23.3	30.4
No Access to Debt	11.3	12.7	12.6	11.0	8.9

loan product jointly with Union debt,[16] while being the cheapest loan product since 2014. In terms of maturity, Bank and Auto consumer loans present the longest maturities for the borrowers, although other lenders also offer maturities close to four years. This table, however, only accounts for installment loans; therefore, it does not include many of the loans with banks and retail stores which are in terms of revolving debt, being renewed monthly and not presenting a maturity.

I summarize the overall household indebtedness in terms of three variables: (i) the debt service to monthly income ratio (DSIR$=\frac{DS_{i,t}}{Y_{i,t}}$), with the debt service including the loan amortization plus all the fees and interest to be paid in a given month; (ii) the consumer debt amount to annual permanent income ratio (CDPIR$=\frac{L_{i,t}}{12 \times \bar{P}_{i,t}}$); and (iii) the log of the real value of the consumer debt. The debt service ratio (DSIR) is a measure of the liquidity constraints faced by households at the end of the month (Madeira 2019b), while the consumer debt amount to the annual permanent income ratio (CDPIR) is a solvency measure that households may face in terms of their total debt amount. The consumer debt (in real terms) can be viewed as an overall measure of indebtedness.

Table 3 illustrates that both the Bank-Retail and the Other Debt borrowers are the most indebted in terms of any of the three measures (DSIR, CDPIR, and real debt amount). The data also demonstrate a substantial difference between Bank only and Bank plus Retail borrowers, with Bank plus Retail borrowers showing a much higher indebtedness with almost 50% higher consumer debt whether in level (in Table 3, exp(4.5)/exp(4.1)−1=49.1%) or CDPIR income ratio (30% for Bank plus Retail compared to 21.4% for Bank only borrowers). Bank plus Retail borrowers also present a much higher loan payment service, with a DSIR of 42.6% for Bank plus Retail compared to 29.7% for Bank only borrowers. Retail Store only and Union debt borrowers present the lowest indebtedness levels, whether in terms of debt service (DSIR), consumer debt to income (CDPIR) or in level (consumer debt in real terms). Retail Store only borrowers are the least indebted, with a consumer debt to income ratio (CDPIR) of just 10% (much lower than the 16.5% of CDPIR of Union debt borrowers) and a real consumer debt which is just one third of the level of the Union borrowers.

[16] Education and Other loans were the cheapest loan product before 2011, but this could have been explained due to strong government subsidization of education loans and an implicit public guarantee against borrower default.

Table 2 Distribution of the interest rate and maturity of each loan type (before and after the 2013 interest rate ceiling law)

Loan type	Interest rate (%)								Maturity (months)					
	Percentiles, Number of observations (N)								Median (P50), Mean, Observations (N)					
	2007–2011				2014–2017				2007–2011			2014–2017		
	P25	P50	P75	N	P25	P50	P75	N	Median	Mean	N	Median	Mean	N
Bank	9	20	40	511	13	18	26	355	36	47	1795	36	39	1426
Retail Store	9	24	42	95	16	23	29	79	36	44	463	12	22	541
Union Debt	13	19	27	200	19	24	27	138	36	45	984	24	33	909
Auto	18	30	39	66	13	21	29	31	48	48	305	36	36	337
Education, Other	9	10	27	77	13	26	27	205	24	37	292	84	104	415

Table 3 Debt service to monthly income ratio (DSIR), Consumer debt to annual permanent income ratio (CDPIR), and Real consumer debt (in log): average values per borrower

Type of debtor	Debt service to income (in %)	Consumer debt to annual permanent income (in %)	Consumer debt (real value in UF) (in log)				
			All waves	2007	2011	2014	2017
Bank	29.7	21.4	4.1	3.8	3.8	4.0	4.4
Bank + Retail	42.6	30.0	4.5	4.3	4.3	4.5	4.8
Retail Store	22.4	10.0	2.3	2.5	2.1	2.3	2.3
Union Debt	24.1	16.5	3.4	3.1	3.5	3.3	3.5
Other Debts	31.7	32.6	4.8	4.4	4.7	4.7	5.1
All debtors	29.2	19.3	3.4	3.2	3.2	3.4	3.8

Average values for borrowers with positive debt

Bank and Retail Store borrowers have the highest indebtedness in terms of the debt service ratio (DSIR), whereas Other Debt borrowers have a higher consumer debt amount to the annual permanent income ratio (CDPIR) and a higher real consumer debt amount. It is noteworthy that the average real consumer debt per borrower has increased for all debt types, except for the Retail Store only debtors, between 2007 and 2017. Looking at the entire population of borrowers, the strongest real debt increase was between 2014 and 2017, with an especially strong growth for the debtors of Banks, Bank-Retail and Other Debts. The real debt also dropped momentarily for the Union borrowers between 2011 and 2014. Households with both Bank and Retail Store debt have much higher loan amounts than the debtors of Bank and Retail Stores separately, which could be a sign that these are debtors with particularly high needs for liquidity. Similar conclusions can be obtained by analyzing the entire probability density function (pdf) and the cumulative distribution function (cdf) of the consumer loan amounts, the debt service to monthly income ratio, and the consumer debt to annual permanent income ratio (Madeira 2019c).

Between 2007 and 2017, the EFH survey asked the following question about loan arrears, "Approximately, in the last 12 months have you ever fallen into arrears or late payments for your loans?". This variable measures default as a dummy denoting one or more events of arrears in the previous year, whether for the current period's loans or loans during the previous year. Arrears imply that the borrower has been late with his loan payment for 30 days or more. Since 2010 the survey also asks for each of the current loans whether the households have been in arrears and for how many months. This allows to build a similar measure of borrower risk for arrears of one month or more for the current loans or a measure of arrears for a failure to repay after three months or more, which is a more common measure of debt risk used at the international level (Madeira 2018, 2019a). Note that this one month arrears measure differs slightly from the previous measure, which includes any arrears event in the previous 12 months, not just for the present period. Note also that these three definitions of loan arrears ("arrears in the past year", "arrears of the current loans: 1 month or more", "arrears of the current loans: 3 months or more") apply to any consumer debt loan, that is, the variable is a dummy with value one if any of the consumer loans is in arrears.

Table 4 illustrates that Bank plus Retail borrowers have the highest risk in all the three measures of arrears, with, respectively, a rate of 27.4% of arrears in the past year, 21.2%

of current arrears (one month or more), and 7.7% of current arrears (three months or more). Other Debt borrowers show the lowest risk of current arrears, with rates of 8.9% and 2.2% at the horizon of one month and three months, respectively. Across all borrowers, the arrears of three months or more are just 25% (that is, 4.2/17.1) of the level of the arrears of one month or more. This shows that most debtors recover a good financial health after being in arrears for one month. Retail Store only borrowers show a high risk of arrears at the horizon of one month with a rate of 19.6%, although a risk of just 2.2% after three months. Bank only, Union and Other debt borrowers show a low risk of arrears both at the horizon of one and three months. Indeed, Bank only borrowers report an arrears rate that is around half the value reported by users of both Bank and Retail credit, demonstrating arrears rates of just 11.6% and 4.2%, respectively, at the horizons of one month and three months or more.

In Chile neither Retail Stores or Unions offer heterogeneous interest rates to the customers. Banks offer customer specific interest rates (Madeira 2019c), therefore, 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. Union Debt lenders cannot risk price their offers, but are able to garnish their clients' wages easily; therefore, this high punishment cost should explain their low arrears rates. However, households with both Bank and Retail Store debt have arrears 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. Households with Other Debts also have high debt and a high value of arrears in the previous year. Perhaps this is because education loans are granted to younger agents, who may be more subject to unemployment risk and unstable income. Additionally, education and auto loans may have lower punishment costs for arrears, because lenders cannot deduct payments and punishment fees from clients' bank accounts (as Banks do) or their wages (as Union Credit institutions do). However, households report a lower risk of arrears above three months for other debts; therefore, these short term risks do not translate into long term default.

Table 4 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. The average borrower dedicates, respectively, 38%, 22%, 11% and 5% of its consumer debt to General Consumption, Durables, Pay Previous Debts, and Health Expenses. Other debt borrowers are mostly related to Durable goods purchases, with 56% of its debt dedicated to this purpose. All the other borrower types have General Consumption as the largest reason for their consumer indebtedness. Households with Bank, Bank plus Retail, and Union Debt are more likely to have motivations of "Pay previous debts", dedicating between 17 and 20% of their consumer debt to such purpose. Union borrowers have the highest share of Health Expenses, with a rate of 13.4%. In terms of the share of loan motives, Bank only borrowers are very similar to the Bank plus Retail borrowers, which makes sense, because the consumer debt loans of the Retail Stores are much lower than the Banks. For instance, Table 3 lists that the average Retail Store only borrower has a debt of just 16.5% of the level of Bank only borrowers.

Table 4 Consumer loan arrears (in %) for each debtor type (all waves) and shares of each loan motive as a fraction of the total consumer debt (in %): averages across borrowers (all waves)

Type of debtor	Arrears in past year	Arrears of the current loans		Shares of loan motives		
	(1 month or more)	(1 month or more)	(3 months or more)	Durables	Pay past debts	Health expenses
	2007–2017	2010–2017	2010–2017			
Bank	16.7	11.6	4.2	34.2	16.7	5.3
Bank + Retail	27.4	21.2	7.7	24.8	19.0	5.7
Retail Store	22.4	19.6	2.2	5.3	4.5	3.6
Union Debt	21.0	16.1	6.1	28.1	20.4	13.4
Other Debts	25.2	8.9	2.2	55.9	4.4	1.6
All debtors	22.8	17.1	4.2	21.6	11.2	5.1

Labor income risks

In addition to their indebtedness level, household risks can also depend on their demographics and labor market profile. Table 5 illustrates a clear sorting pattern of households across different debt types. For the average of all debtors, the household has 3.5 members (or 2.1 members in the OECD household size scale), log permanent income of 3.9, the household head's education is 12.8 years (just slightly above high school completion) and its age is 47 years. The unemployment and income volatility risk of the average debtor are 3.8% and 30.9%, respectively.

In terms of income, Bank debtors have the highest income, while showing the lowest unemployment risk. Other Debt households have the highest education and the highest income levels next to Bank and Bank plus Retail borrowers, but also the lowest age and one of the highest values of labor income volatility (only below the No Access households). On the other end, households with No Access to Debt have the lowest income and education, while showing the highest unemployment risk, age and labor income volatility. Next to the No Access households, Retail store borrowers are the ones with the lowest income and education, while showing the highest unemployment risk. In terms of the household size (whether measured as the number of household members or by the OECD modified scale), Bank plus Retail debtors represent the largest households, whereas those with No Wish for Debt and No Access to Debt represent the smallest households. It is relevant to note that Union debt borrowers have one of the lowest income (only above the No Access households), but also the lowest labor income volatility, possibly due to the types of occupations of their workers. Similar conclusions can be obtained by analyzing the probability density function (pdf) and the cumulative distribution function (cdf) of the permanent income, labor income volatility, and unemployment risk across different debtor types (Madeira 2019c).

Overall, this section portrays a clear picture of the household groups. Households with No Access to Debt have the lowest income and highest unemployment risk. The users of just Bank loans are the ones with the highest income and lowest unemployment rate; however, they experience moderate wage volatility, which creates a demand for debt to smooth consumption. Bank plus Retail and Other Debts have a slightly lower income than the Bank only borrowers, but they are subject to a considerably higher

Table 5 Demographic characteristics of the households according to their debt type: average values per borrower

Type of debtor	Permanent Income in UF (log)	Unemployment risk (%)	Labor income volatility (%)	Education (in years)	Age (years)	Number of household members	Household size (OECD scale)
Bank	4.2	4.1	30.9	14.1	46.5	3.3	2.0
Bank + Retail	4.1	4.4	28.8	13.5	46.0	3.6	2.2
Retail Store	3.6	5.3	31.0	11.7	48.2	3.4	2.1
Union Debt	3.7	4.8	27.6	12.1	52.7	3.4	2.0
Other Debts	4.1	4.8	38.6	14.3	42.3	3.5	2.1
No Wish for Debt	3.8	4.5	33.3	12.5	52.6	2.9	1.8
No Access to Debt	3.5	5.4	41.9	11.0	53.0	3.0	1.9
All debtors	3.9	4.8	30.9	12.8	47.3	3.5	2.1

The OECD equivalence scale for household size corresponds to a value of 1 for the household head plus 0.5 for each additional adult and 0.3 for each child, in order to account for economies of scale in the sharing of home goods (OECD 2008)

unemployment risk and, therefore, have a higher indebtedness level than Bank only borrowers. The Union Debt borrowers are in the middle of the debt and labor income risk profiles, because these households have a low income level but also a low wage volatility and a moderate level of unemployment risk. Therefore, these households also have a moderate level of indebtedness. Finally, Retail Store only borrowers among the households with access to debt have the lowest income and education coupled with the highest unemployment risk; therefore, such households also have the lower debt amounts.

Results

The role of demographics, income profile and unobserved preferences

This section discusses the results from the consumer loan choice and default model exposed in "An empirical model of choice of lender, loan amount and default" section. The model requires some variables that affect loan choice, but not default. Consumer loans typically have a maturity of 12–48 months, therefore, it is reasonable to assume that the labor market conditions influencing loan choice happened 4 quarters or more before the current period. The model also controls for the education, age, region, household size and debt motivations:

$$x_{i,t} = \left\{ \begin{array}{c} \ln(\bar{P}_{i,t-4}), \bar{u}_{i,t-4}, \bar{\sigma}_{i,t-4}, \\ \text{years of education and age of household head, dummies by year,} \\ \text{dummy for residence region, OECD household size scale.} \end{array} \right\}.$$

Similarly, I assume that the vector $z_{i,t}$ that explains loan arrears or default at time t includes some variables that do not necessarily affect loan choice, such as the Debt service to monthly income ratio (DSIR) and the Consumer debt to annual permanent income ratio (CDPIR):

$$z_{i,t} = \left\{ \begin{array}{c} DSIR_{i,t}, CDPIR_{i,t}, \ \ln(\bar{P}_{i,t}), \bar{u}_{i,t}, \bar{\sigma}_{i,t}, \\ \text{years of education and age of household head,} \\ \text{dummies for each year, dummy for residence region,} \\ \text{OECD household equivalence size scale,} \\ \text{Share of debt justified by “Durables”,} \\ \text{“Pay Previous Debts” and “Health”.} \end{array} \right\}.$$

Table 6 shows the estimates for the coefficients of loan choice, β_b, from the Mixed Multinomial Logit Model. The coefficients for both the lagged household expected income ($\ln(\bar{P}_{i,t-4})$) and years of education have a positive impact on choosing the Bank, Bank plus Retail, and Other Debts options, while having a negative effect on the No Access option. Both the income and unemployment coefficients have its most positive impact for the option of Bank loans. This implies that higher income and education unambiguously increase the option of a Bank loan, while decreasing the option of No-Access. The coefficient of lagged unemployment ($\bar{u}_{i,t-4}$) increases the probability of all the loan options relative to No Debt, especially for the Other Debt and Retail only options. This demonstrates that unemployment risk is a factor for choosing loans under a motivation of consumption smoothing (Attanasio and Weber 2010). It also represents that Retail Store and Other Debt borrowers are more subject to unemployment risk (Agarwal et al. 2020). Wage volatility decreases the probability of choosing any of the the debt alternatives, but it increases the probability of No Access, a sign of a precautionary savings motive to avoid debt (Attanasio and Weber 2010).

Older households are less likely to choose any debt option, which makes sense due to their savings needs at the end of life (Attanasio and Weber 2010). The age effect is lower for the Union and No Access alternatives. Larger households (measured by the OECD scale) are more likely to choose all kinds of loan options, although this has a small effect on the No Access alternative. This demonstrates that household size can be positively related to expenditure shocks, whereas age is negatively related to consumption shocks. The Santiago capital region dummy is not statistically significant for any alternative, which demonstrates that the Chilean lenders have an adequate presence across the entire country. Finally, the probability of Bank only, Union and Other Debts increased in 2017, whereas the probability of No Access decreased in the same year. This demonstrates that the Chilean households decreased their credit constraints and improved their loan options in recent years.

Table 7 shows the results for the choice of loan amount and arrears. Note 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, whereas the variables affecting default correspond to the current period t. Income, education, and loan motivations of “Durables”, “Pay previous debts” and “Health” are positively related to loan amounts. Unemployment risk and age are both associated with lower loan amounts. Wage volatility, household size and the region dummy have small coefficients, which are statistically insignificant. After controlling for observables, the loan amounts fell in 2010/2011 and 2014, which is consistent with the effects of the 2010 earthquake and the interest rate ceiling law of 2013 (Madeira 2019b). The coefficients for the Odds Ratios of selection indicate that the probability of making another choice besides the chosen alternative has a positive effect on the loan amount for the Bank

Table 6 Coefficients for the utility of each type of debt (Mixed Multinomial Logit Model)

Control variables	1 = Bank	2 = Bank + Retail	3 = Retail Store	4 = Union Debt	5 = Other Debt	6 = No Access
Age	− 0.107*** (0.0161)	− 0.109*** (0.0161)	− 0.102*** (0.0160)	− 0.0766*** (0.0159)	− 0.123*** (0.0163)	− 0.00340* (0.00183)
OECD equivalent household size	2.253*** (0.346)	2.508*** (0.347)	2.608*** (0.348)	2.643*** (0.350)	2.534*** (0.349)	0.340*** (0.0460)
Income: $\ln(\bar{P}_{i,t-4})$	1.094*** (0.183)	0.806*** (0.182)	− 0.0273 (0.178)	− 0.0636 (0.184)	0.958*** (0.186)	− 0.757*** (0.0591)
Unemployment $\bar{u}_{i,t-4}$	15.93*** (2.214)	16.29*** (2.186)	18.13*** (2.144)	17.25*** (2.286)	20.33*** (2.264)	1.873*** (0.694)
Wage volatility $\bar{\sigma}_{i,t-4}$	− 3.636*** (0.577)	− 3.788*** (0.577)	− 3.484*** (0.573)	− 3.796*** (0.581)	− 3.523*** (0.576)	0.252*** (0.0810)
Education (years)	0.123*** (0.0389)	0.0935** (0.0387)	− 0.00713 (0.0386)	0.0591 (0.0402)	0.110*** (0.0400)	− 0.0763*** (0.0118)
Santiago capital region	0.206 (0.217)	0.214 (0.215)	0.174 (0.213)	0.205 (0.221)	0.279 (0.221)	0.0467 (0.0601)
2008/2009	− 0.662* (0.384)	− 1.039*** (0.380)	− 0.830** (0.377)	− 0.805** (0.401)	− 1.121*** (0.397)	0.191* (0.106)
2010/2011	− 0.656** (0.302)	− 1.092*** (0.299)	− 0.475 (0.296)	0.246 (0.312)	− 0.636** (0.306)	0.530*** (0.0908)
2014	0.433 (0.353)	0.453 (0.351)	1.221*** (0.353)	1.796*** (0.371)	0.551 (0.359)	0.810*** (0.112)
2017	1.011*** (0.324)	0.281 (0.321)	− 0.122 (0.322)	1.029*** (0.338)	1.334*** (0.331)	− 0.505*** (0.100)
Constant	− 8.319*** (0.515)	− 3.676 (0)	7.217*** (1.899)	3.556* (1.981)	− 8.251*** (1.989)	9.179*** (0.655)
Observations: 21,319 households. McFadden's Pseudo R-squared: 0.071						

Robust Huber-White Standard-errors in ()

***, **, *Denote 1%, 5%, 10% statistical significance

plus Retail and the Other Debts, while having a negative effect on the loan amount for the Bank, Retail and Union Debt borrowers. This demonstrates that the borrowers of Bank-Retail and Other Debts have a strong preference for higher loan amounts, whereas the Bank, Union and, especially, the Retail Store borrowers prefer lower loan amounts.

Default behavior is negatively associated with age, income, education, and the share of the largest debt type in terms of the consumer debt (which makes sense, because it is easier to renegotiate debt with a single lender than with many lenders). Arrears are positively associated with larger households, unemployment risk, Debt service to monthly income (DSIR), Consumer debt to annual permanent income (CDPIR), the motives of "Pay previous debts" and "Health needs". Wage volatility and the region dummy have small coefficients, which are statistically insignificant. In the case of Arrears in the past year, there is a negative effect of the selection probability of not having a Bank, Union Debt and Other Debt alternative. However, although with the same sign, these effects are not significant for the Ordered Logit model of current Arrears above 1 and 3 months. This could be because these selection probabilities affect the short term liquidity of being in arrears or not, but their effect is perhaps less powerful for determining more

Table 7 Coefficients for the loan amount (in log) and propensity for arrears

Control variables	Consumer debt (in log) OLS	Arrears longer than 1 and 3 months Ordered Logit	Arrears in the past year Logit
Age	− 0.002*** (0.0009)	− 0.0249*** (0.00254)	− 0.0224*** (0.00193)
OECD equivalent household size	0.00188 (0.0176)	0.502*** (0.0506)	0.398*** (0.0376)
Income: $\ln(\bar{P}_{i,t})$	0.439*** (0.0211)	− 0.452*** (0.0705)	− 0.327*** (0.0486)
Unemployment $\bar{u}_{i,t}$	− 0.622** (0.296)	3.201*** (0.844)	2.304*** (0.642)
Wage volatility $\bar{\sigma}_{i,t}$	− 0.0220 (0.0359)	0.0680 (0.0846)	− 0.0568 (0.0708)
Education (years)	0.0308*** (0.00473)	− 0.0985*** (0.0138)	− 0.0686*** (0.0104)
Santiago capital region	0.00903 (0.0252)	− 0.102 (0.0689)	− 0.0637 (0.0536)
2008/2009	0.0375 (0.0485)		0.368*** (0.115)
2010 / 2011	− 0.277*** (0.0354)		0.711*** (0.0867)
2014	− 0.307*** (0.0406)	− 0.367*** (0.0966)	0.404*** (0.0984)
2017	0.0646 (0.0407)	0.0249 (0.0973)	1.252*** (0.0948)
Share of loan for "Durables"	0.906*** (0.0354)	− 0.182* (0.109)	0.154** (0.0769)
Share of loan for "Pay previous debts"	0.969*** (0.0419)	0.431*** (0.118)	0.531*** (0.0872)
Share of loan for health needs	0.416*** (0.0588)	0.355** (0.150)	0.522*** (0.116)
$1(Y_{i,t} = 1)\Pr(Y_{i,t} \neq 1 \mid x_{i,t})$	− 0.470*** (0.138)	− 0.359 (0.404)	− 1.107*** (0.302)
$1(Y_{i,t} = 2)\Pr(Y_{i,t} \neq 2 \mid x_{i,t})$	0.278** (0.137)	0.423 (0.390)	− 0.374 (0.299)
$1(Y_{i,t} = 3)\Pr(Y_{i,t} \neq 3 \mid x_{i,t})$	− 1.930*** (0.157)	0.402 (0.464)	− 0.471 (0.349)
$1(Y_{i,t} = 4)\Pr(Y_{i,t} \neq 4 \mid x_{i,t})$	− 1.098*** (0.130)	− 0.0540 (0.390)	− 0.704** (0.289)
$1(Y_{i,t} = 5)\Pr(Y_{i,t} \neq 5 \mid x_{i,t})$	0.382*** (0.131)	− 0.245 (0.390)	− 0.511* (0.288)
Ratio of Debt Service to Income, $DSIR_{i,t}$		1.267*** (0.148)	0.998*** (0.122)
Ratio of Debt to Income, $CDPIR_{i,t}$		0.797*** (0.164)	0.751*** (0.130)
Share of the largest loan type of the total consumer debt	− 0.951*** (0.0462)	− 0.578*** (0.131)	− 0.358*** (0.100)
Constant	8.449*** (0.286)	5.732*** (0.956)	3.870*** (0.654)
Cutoff 2		1.624* (0.953)	
Observations	12,075	9071	12,075
R-squared or Pseudo R2	0.498	0.106	0.104

Robust Huber-White Standard-errors in ()

***, **, *Denote 1%, 5%, 10% statistical significance

serious arrears above three months. The years 2010–2011 show a higher propensity of Arrears, which could have been due to the Chilean mega-earthquake of 2010.

In terms of the estimated heteroscedasticity of unobserved tastes for choosing different debt alternatives, the results demonstrate that the unobserved factor for the preference for all debts has a big unobserved standard-deviation and it has a significant effect on choices. However, the unobserved factor for bank loans has a small standard-deviation and it is statistically insignificant. These results are available from the author upon request and are reported with higher detail in the full information model estimated in the previous version of this article (Madeira 2019c). This demonstrates that the consumers with debt (choices 1 to 5) are different in some unobservable way from those with No

Access and No Wish for Debt, perhaps because of their preferences, previous credit history or a better overall impression these households make with lenders.

Counterfactual simulations of policies

I now use the pooled EFH (2007–2017) sample and the estimated models to develop four counterfactual exercises to analyze their effect on the debtor type choices $\Pr(Y_{i,t} = b \mid X_{i,t0})$, consumer loan amounts $E[\ln(L_{i,t}) \mid Z_{i,t0}]$, and arrears (past arrears, arrears above 1 or 3 months), $\Pr(D_{i,t} \geq 1 \mid Z_{i,t})$. I then obtain the aggregate consumer debt across all households ($aggD_t = \sum_i (\sum_{b=1}^{5} \Pr(Y_{i,t} = b \mid X_{i,t0})) \exp(E[\ln(L_{i,t}) \mid Z_{i,t0}])$), which accounts for each borrower having some positive probability of obtaining consumer debt, and the average debt per borrower with a positive debt amount, $E[\ln(L_{i,t}) \mid Z_{i,t0}]$. The exercise also computes the average borrower risk ($\Pr(D_{i,t} \geq 1 \mid Z_{i,t})$), the risk across all households which takes into account that some households have no debt ($\sum_{b=1}^{5} \Pr(Y_{i,t} = b \mid X_{i,t0}) \Pr(D_{i,t} \geq 1 \mid Z_{i,t})$), and the portfolio risk which also accounts for households having different amounts of debt ($\frac{\sum_i (\sum_{b=1}^{5} \Pr(Y_{i,t}=b|X_{i,t0})) \exp(E[\ln(L_{i,t})|Z_{i,t0}]) \Pr(D_{i,t}\geq 1|Z_{i,t})}{aggD_t}$).

The first counterfactual exercise is the Baseline, showing the average values of the variables obtained with the current exogenous variables observed in the EFH sample, $X_{i,t0}$, $Z_{i,t0}$, $Z_{i,t}$.

The second counterfactual exercise accounts for a large increase of the public pensions on January of 2022 (Law 21419), with a minimum noncontributory public pension of 185,000 pesos (roughly, 210 USD). This law, implemented owing to the social discontent caused by the Covid pandemic and the Social Explosion (Madeira 2022), provided all the retirees in households within the lowest nine deciles of income (therefore, almost universal coverage) with a monthly solidarity pension of 185,000 pesos for retirees with pensions below 630,000 pesos, with a decreasing linear amount until the benefit reaches zero pesos for pensions equal to or greater than one million pesos: $B_{k,t} = b_1 1(\tilde{p}_{k,t} \leq b_2) + b_1(1 - \frac{\tilde{p}_{k,t} - b_2}{b_3 - b_2}) 1(b_2 < \tilde{p}_{k,t} < b_3)$, with $b_1 = 185,000$, $b_2 = 630,000$, and $b_3 = 1,000,000$, with $\tilde{p}_{k,t}$ being the current contributory pension income of the k-th member of household i at time t. The households' new income level is: $\ln(\bar{P}_{i,t} + 1(\bar{P}_{i,t} \leq DY_9) \sum_k B_{k,t} 1(age_{k,t} \geq 65))$. The same adjustment is applied to $\ln(\bar{P}_{i,t-4})$. The adjustment accounts for households being within the 9-th income decile ($1(\bar{P}_{i,t} \leq DY_9)$), whether each member is above 65 and entitled to public pensions ($1(age_{k,t} \geq 65)$), and the pension income received by each member ($B_{k,t}$).

The third counterfactual exercise, labeled "Income test for repayment purposes", aims to simulate a scenario in which regulators require lenders to make stronger assessments of the repayment capacity of the borrowers, as suggested by the IMF-World Bank Financial Sector Assessment (FSA) mission to the Chilean government (Cohen and Dijkman 2021). This exercise is implemented by changing the coefficients β_b of the Mixed Multinomial Logit model, with the coefficients for unemployment ($\beta_b^{unemployment}$) and wage volatility ($\beta_b^{wage-volatility}$) for each debt alternative with loan access (options 1 to 5) being equal to the minimum across all lender types: $\beta_b^{unemployment} = \min_{b'=1,\ldots,5} \hat{\beta}_{b'}^{unemployment}$

and $\beta_b^{wage-volatility} = \min_{b'=1,\ldots,5} \hat{\beta}_{b'}^{wage-volatility}$.[17] The reasoning is that the regulator will require the same standards of screening for unemployment and wage volatility as those that are already implemented by the most robust lenders. Furthermore, I impose that the coefficients of unemployment risk and wage volatility have twice as much effect on the probability of household having No Access: $\beta_{b=B}^{unemployment} = 2\hat{\beta}_{b=B}^{unemployment}$ and $\beta_{b=B}^{wage-volatility} = 2\hat{\beta}_{b=B}^{wage-volatility}$.[18] Therefore, this exercise demonstrates the potential effects of requiring lenders to make more robust decisions regarding borrowers who may have repayment problems later on.

The fourth counterfactual exercise is related to the potential benefits that Chile could achieve if it invested in a large scale financial literacy program. This Financial literacy program consists in obtaining new values for the outcomes of interest $(\Pr(Y_{i,t} = b \mid X_{i,t0}), E[\ln(L_{i,t}) \mid Z_{i,t0}], \Pr(D_{i,t} \geq 1 \mid Z_{i,t}))$ by considering that: (i) all households increase their financial knowledge by the equivalent to one year of education, $education_i(new) = 1 + education_i$; (ii) the coefficient for years of education for getting No Access to Credit becomes twice as strong, $\beta_{b=B}^{education} = 2\hat{\beta}_{b=B}^{education}$[19]; (iii) the coefficients for the impact of education for choosing each loan alternative become less positive, owing to borrowers higher awareness of the median interest rate charged by those lenders, $\beta_b^{education} = \hat{\beta}_b^{education} - i_b$ for $b = 1, \ldots, 5$, with $i_1 = 0.15$, $i_2 = 0.195$, $i_3 = 0.24$, $i_4 = 0.22$, $i_5 = 0.24$.[20]

Table 8 shows the probabilities of each type of debt under the different counterfactual policies. The Baseline and the data statistics in Table 1 are quite close; therefore, the model is reliable for counterfactual analysis. Relative to the Baseline, all the counterfactuals increase the probability of the Bank debt type and decrease the probability of the Retail Store only debt type, with both effects being stronger in the case of the Financial Literacy program. The Bank plus Retail and Union debt types are highest for the Income test scenario, while being lowest for the Financial Literacy program. The Other Debts option is highest for the Pension law policy, while being lowest for the Financial Literacy program. Both the Pension Law and the Financial Literacy policies decrease the No Access and increase the No Wish for Debt option, with both effects especially strong in the case of the Financial Literacy program. The Pension law and Income test increase the Bank plus Retail option, while the Financial literacy decreases it substantially.

Overall, the Financial Literacy program shows the largest transformation. The probability differences between the Pension Law and the Baseline for each alternative vary between − 0.5% for the case of No Access to +0.5% for the No Wish for Debt option. The Income test policy has much stronger effects than the Pension Law. For the Income test, the differences relative to the Baseline go from − 1.7% in the case of a reduction of the Retail Store only option to an increase of +1.5% in the No Access option. However, the Financial Literacy program has much larger effects. Relative to the Baseline, the

[17] The coefficients for the unemployment risk and wage volatility in Table 6 become, respectively, 15.93 and − 3.796 for all the lender options with positive debt (that is, options 1 to 5).

[18] The coefficients for unemployment risk and wage volatility in the "No Access" column of Table 6 become 3.746 and 0.504, respectively.

[19] The coefficient for years of education on the column of "No Access" in Table 6 becomes − 0.1526.

[20] The coefficients for the years of education in Table 6 become − 0.027, − 0.1015, − 0.24713, − 0.1609 and − 0.13, for, respectively, the lender options of 1 "Bank", 2 "Bank & Retail", 3 "Retail", 4 "Union", 5 "Other Loans".

Table 8 Probabilities (in %) for each type of debt choice under different counterfactual policies

Counterfactual	Bank	Bank + Retail	Retail Store	Union Debt	Other Debts	No Wish	No Access
Baseline	9.6	13.2	26.0	5.9	5.9	28.0	11.3
Pension law of 2022	9.8	13.5	25.6	5.8	6.0	28.5	10.8
Income test	10.1	14.4	24.3	6.3	5.0	27.2	12.8
Financial literacy	16.9	12.6	14.7	4.4	2.9	41.9	6.5

Financial Literacy program increases the Bank only and the No Wish for debt options by, respectively, 7.3% and 13.9%. It also reduces the Retail Store only and the No Access options by, respectively, 11.3% and 4.8%. Furthermore, it decreases the Other Debts, Union and Bank plus Retail options by 3%, 1.5% and 0.6%.

Table 9 illustrates the policies' effects on the average debt probability (the sum of the probabilities for the choices 1 to 5 in Table 8), the aggregate consumer debt across all households and the average debt per borrower, with the debt amounts being standardized relative to the baseline (reported as 100%). The Financial Literacy program implies the largest reduction of the debt probability and of the aggregate consumer debt. However, the financial literacy program increases the average debt per borrower, although the number of borrowers decreases, implying an overall drop in the aggregate consumer debt. The Pension Law increases the debt probability, aggregate debt and average debt per borrower, while the Income test decreases all those variables. Again, the Financial Literacy program shows the largest transformations of the consumer debt choices. The Pension law increases aggregate consumer debt by 1%, whereas the Income test reduces aggregate debt by 0.9%. The Financial literacy program reduces the aggregate household consumer debt by 12.4%, more than ten times as much as the Income test policy.

Table 10 analyzes the policies' effect on default risk, as given by the probability of arrears in the previous year, arrears above 1 month or above 3 months. Every policy reduces (even if only slightly) the default risk, regardless of the measure of arrears used or the population (average borrower, all households, and portfolio risk). The Financial Literacy program is the policy with the strongest reduction in default risk. The Pension law differs from the baseline risk measures by at most 0.1% (for instance, it reduces portfolio arrears in the past year by 0.1%), whereas the Income test has a somewhat larger effect with a reduction in risk measures that reaches 0.2% (its effect on the portfolio arrears in the past year, for instance). The Financial Literacy program has a large effect in reducing arrears risk. It reduces the arrears risk in the past year by 1.5%, 2.7% and 2.7%, respectively, across all borrowers, all households and for the debt portfolio. It reduces the current arrears risk at the horizon of one month by 1.2%, 1.9% and 1.7%, respectively, across all borrowers, all households and for the debt portfolio. Finally, for the most standard risk measure which is the arrears at 3 months, the Financial Literacy program reduces risk by 0.4%, 0.5% and 0.5%, respectively, across all borrowers, all households and for the debt portfolio.

Overall, the Financial Literacy program would reduce the portfolio risk of arrears above 3 months from a Baseline of 2.5% to just 2%, a reduction of 20% in the overall risk faced by lenders.

Table 9 Effect of different policies on the consumer debt amount (for all households and the average borrower) relative to the baseline (in %) and on the overall probability (in %) of being a consumer debtor

Counterfactual	Any Debt Probability	Consumer debt (aggregate)	Consumer debt (average borrower)
Baseline	60.6	100	100
Pension law of 2022	60.7	101.0	100.7
Income test	60.0	99.1	99.9
Financial literacy	51.6	87.6	103.5

Table 10 Effect of different policies on the arrears risk (in %) of borrowers, all households and for the consumer debt portfolios

Counterfactual Default definition	Borrower risk			All households risk			Portfolio risk		
	Arrears past year	Arrears (months)		Arrears past year	Arrears (months)		Arrears past year	Arrears (months)	
		1	3		1	3		1	3
Baseline	24.1	18.2	4.8	14.3	9.7	2.5	16.7	9.6	2.5
Pension law of 2022	24.0	18.1	4.8	14.3	9.6	2.5	16.6	9.5	2.5
Income test	24.0	18.2	4.8	14.1	9.5	2.4	16.5	9.4	2.4
Financial literacy	22.6	17.0	4.4	11.6	7.8	2.0	14.0	7.9	2.0

All households' risk takes into account that households with no debts have zero risk. Portfolio risk takes into account the size of the loans held by each borrower

Table 11 summarizes the policies' effect on the aggregate debt and portfolio default risk across debt types. Every policy reduces the default risk for all the debt types, even if only slightly. The Pension law increases the aggregate debt of all the debt types, especially for the Unions. It increases aggregate debt in Unions by 2.1%, whereas its effect on other debt types is between 0.5 and 0.7%. The Income test reduces the aggregate debt of all the debt types, especially for the Retail Store only. The Retail Store only debt is reduced by 4.7%, the Other Debts experience no effect, while the remaining types are reduced by 0.3–1%. The effects of the Pension law and the Income test policies on reducing arrears at either the one or three months horizons are negligible.

Finally, the Financial Literacy program decreases the aggregate debt amount across all the debt types, with a particularly strong effect on Retail Store only loans. It also substantially decreases the Bank, Bank plus Retail, Retail and Union Debt default risks, with especially strong effects on the Retail Store only, Bank plus Retail Store, and Union risks. Its strongest effect is on reducing Retail Store only debt by 30%, whereas the other debt forms fall between 8.6% for Bank only and 14% for the Union option. The Financial Literacy program also decreases the risk of arrears substantially. It decreases the portfolio arrears risk at the one month horizon between 0.9% for the Bank only and 2.8% for the Retail Store only options. It also decreases the portfolio arrears risk at the three months horizon between 0.2% for the Bank only and 0.8% for the Retail Store only options.

This exercise demonstrates that, while all policies affect borrowers' choices, the Financial Literacy program presents the strongest impact on the debt market, whether in terms of the debt type choice, overall amount of debt, or the loan default. In particular,

Table 11 Effect of different policies on the loan amount aggregate relative to the baseline (in %) and the portfolio arrears risk (in %) across debtor types

Counterfactuals	Bank			Bank+Retail			Retail Store			Union Debt			Other Debts		
	Debt aggregate	Arrears (months)		Debt aggregate	Arrears (months)		Debt aggregate	Arrears (months)		Debt aggregate	Arrears (months)		Debt aggregate	Arrears (months)	
		1	3		1	3		1	3		1	3		1	3
Baseline	100	6.6	1.5	100	14.0	3.9	100	14.2	3.8	100	12.2	3.2	100	8.6	2.1
Pension law of 2022	100.6	6.6	1.5	100.5	14.0	3.9	100.7	14.1	3.8	102.1	12.1	3.2	100.6	8.6	2.1
Income test	99.7	6.5	1.5	99.0	13.8	3.8	95.3	13.9	3.7	99.5	12.1	3.2	100.0	8.5	2.1
Financial literacy	91.4	5.7	1.3	88.3	11.6	3.2	70.0	11.4	3.0	86.0	10.1	2.6	90.0	7.0	1.7

this program substantially increases the probability of No Wish for Debt and Bank loans, while decreasing the aggregate loan amounts and the Any Debt, No Access, and the default risk probabilities.

Conclusions and policy implications

This study demonstrates how households' characteristics affect their choice of lenders, consumer debt amounts and default behavior. I find that borrowers' characteristics differ substantially across lender types. Banks lend to the borrowers of highest income and education and the lowest unemployment rates. Households with No Access to Debt are at the precise opposite, with the lowest income and education and the highest unemployment risk. This makes sense because Banks represent the lender institutions that make the most intensive use of credit scoring and customer specific interest rates, therefore, being able to attract the safest borrowers. Bank plus Retail and Other Debt borrowers have slightly lower income and education than those with Bank loans only; however, they are younger and with larger debts. Retail store borrowers, among those with credit access, have the lowest income and education plus an unemployment risk almost as high as those with No Access to Debt; therefore, they have the lowest consumer debt amounts. Union borrowers have somewhat better conditions, with slightly higher income and education and also larger loans.

A multiple choice model demonstrates that education and income increase the probability of Banks, Banks-Retail and Other Debts, while decreasing the probability of No Access to Debt. Unemployment risk and household size increase the probability of opting for all loans (especially Retail Stores and Other Debts), whereas age and wage volatility decrease the probability of all loans. A second step regression demonstrates that loan amounts increase with income and education, while decreasing with unemployment risk, age and the share of debt with the main lender. The default probability decreases with income, education and age, while increasing with high indebtedness ratios, unemployment risk, household size, loans motivated by Paying Previous Debts and Health needs. Paying previous debts is a motive that proxies "moral hazard" (since it benefits previous lenders, not the current lender), while Health shocks can be viewed as an adverse selection variable that lenders do not observe.

The model can be used to simulate counterfactual results for a variety of policies. As examples, I analyze the recent increase in pensions, higher regulatory requirements from lenders to assess borrowers' repayment capacity, or a Financial Literacy program. The counterfactuals demonstrate that all of these policies reduce the borrowers' risk levels. Higher public pensions increase both the number of debtors and their debt amounts, whereas the repayment capacity test lowers them. A financial literacy program could have a great impact on reducing the number of debtors, the value of consumer debt and the share of households with No Access to Debt, while increasing the share of borrowers with Banks. It also substantially decreases the arrears risk of all the debt options.

Acknowledgements

I thank seminar participants at the Federal Reserve Bank of New York, Deutsche Bundesbank, Federal Reserve Bank of Philadelphia, LACEA, Universidad Carlos III, University of York, Bank of Spain, Bank of England and Central Bank of Chile. I thank Bernardita Piedrabuena for all the great conversations at an early stage of this project.

Author contributions

CM: conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft, writing—review & editing, visualization, supervision, project administration, funding acquisition. The author read and approved the final manuscript.

Funding
The author received no funding for this project.

Availability of data and materials
All the codes are freely available at the repository Mendeley data: https://data.mendeley.com/datasets/8hv5d3sm2h/1. The datasets used in this study are freely available from the Chilean Bureau of Official Statistics (INE) and from the Central Bank of Chile after a research project form is filled.

Declarations

Competing interests
The author declares that he has no competing interests.

Received: 9 August 2021 Accepted: 16 January 2023
Published online: 20 February 2023

References

Agarwal S, Qian W, Tan R (2020) Household finance: a functional approach. Palgrave Macmillan, London
Ahmed I, Socci C, Severini F, Yasser Q, Pretaroli R (2018) Forecasting investment and consumption behavior of economic agents through dynamic computable general equilibrium model. Financial innovation 4(1):1–21
Ampudia M, van Vlokhoven H, Zochowski D (2016) Financial fragility of euro area households. J Financ Stab 27:250–262
Attanasio O, Weber G (2010) Consumption and saving: models of intertemporal allocation and their implications for public policy. J Econ Lit 48(3):693–751
Badarinza C, Campbell J, Ramadorai T (2016) International comparative household finance. Annu Rev Econ 8(1):111–144
Banerjee A, Karlan D, Zinman J (2015) Six randomized evaluations of microcredit: introduction and further steps. Am Econ J Appl Econ 7(1):1–21
Bertrand M, Morse A (2011) Information disclosure, cognitive biases, and payday borrowing. J Finance 66:1865–1893
Brown M, van der Klaauw W, Wen J, Zafar B (2016) Financial education and the debt behavior of the young. Rev Financ Stud 29(9):2490–2522
Cohen C, Dijkman M (2021) Chile—financial sector assessment program, December 2021: Technical note—household indebtedness and financial consumer protection. World Bank and IMF, Washington, D.C
Cull R, Gan L, Gao N, Xu L (2019) Dual credit markets and household usage to finance: evidence from a representative Chinese household survey. Oxf Bull Econ Stat 81(6):1280–1317
Edelberg W (2006) Risk-based pricing of interest rates for consumer loans. J Monet Econ 53:2283–2296
Einav L, Jenkins M, Levin J (2012) Contract pricing in consumer credit markets. Econometrica 80(4):1387–1432
Eurosystem Household Finance and Consumption Network (2009) Survey data on household finance and consumption: research summary and policy use. ECB Occasional Paper 100
Gallardo S, Madeira C (2022) The role of financial surveys for economic research and policy making in emerging markets. Handbook Bank Finance Emerg Mark. https://doi.org/10.4337/9781800880900.00047
Habibah U, Hassan I, Iqbal M (2018) Household behavior in practicing mental budgeting based on the theory of planned behavior. Financ Innov 4(1):1–14
Hamakhan Y (2020) The effect of individual factors on user behaviour and the moderating role of trust: an empirical investigation of consumers' acceptance of electronic banking in the Kurdistan Region of Iraq. Financ Innov 6(1):1–29
Hensher D, Rose J, Greene W (2015) Applied choice analysis. Cambridge University Press, Cambridge
Jaffee D, Stiglitz J (1990) Credit rationing. In: Friedman BM, Kahn FH (eds) Handbook of monetary economics, vol 2. North Holland, Amsterdam
Kou G, Peng Y, Wang G (2014) Evaluation of clustering algorithms for financial risk analysis using MCDM methods. Inf Sci 275:1–12
Kou G, Akdeniz Ö, Dinçer H, Yüksel S (2021) Fintech investments in European banks: a hybrid IT2 fuzzy multidimensional decision-making approach. Financ Innov 7(1):1–28
Li T, Kou G, Peng Y, Yu P (2022) An integrated cluster detection, optimization, and interpretation approach for financial data. IEEE Trans Cybern 52(12):13848–13861
Madeira C (2015) Earnings dynamics of Chilean workers: employment transitions, replacement ratio, and income volatility, Working Paper 754, Central Bank of Chile, Chile
Madeira C (2019c) Adverse selection, loan access and default in the Chilean consumer debt market, Working Paper 838, Central Bank of Chile, Chile
Madeira C (2018) Explaining the cyclical volatility of consumer debt risk using a heterogeneous agents model: the case of Chile. J Financ Stab 39:209–220
Madeira C (2019a) Measuring the covariance risk of consumer debt portfolios. J Econ Dyn Control 109:21–38
Madeira C (2019b) The impact of interest rate ceilings on households' credit access: evidence from a 2013 Chilean legislation. J Bank Finance 106:166–179
Madeira C (2022) The impact of the Chilean pension withdrawals during the Covid pandemic on the future savings rate. J Int Money Finance 126:102650
McFadden D, Train K (2000) Mixed MNL models for discrete response. J Appl Econom 15(5):447–470
OECD (2008) Growing unequal? Income distribution and poverty in OECD countries, OECD, Paris. https://doi.org/10.1787/9789264044197-en

Sanroman G, Santos G (2021) The joint distribution of income and wealth in Uruguay. Rev Cuad Econ 40(83):609–642
Tang M, Mei M, Li C, Lv X, Li X, Wang L (2020) How does an individual's default behavior on an online peer-to-peer lending platform influence an observer's default intention? Financ Innov 6(1):1–20
Thakor A (2020) Fintech and banking: what do we know? J Financ Intermed 41:100833
Vella F (1998) Estimating models with sample selection bias: a survey. J Hum Resour 33(1):127–169
Wooldridge J (2010) Econometric analysis of cross section and panel data. MIT Press, Cambridge

Publisher's Note