Revista de Análisis Económico, Vol. 27, Nº 1, pp. 55-70 (Abril 2012)

THE DETERMINANTS OF HOUSEHOLD DEBT DEFAULT

LOS DETERMINANTES DE NO-PAGO DE DEUDA DE HOGARES

RODRIGO ALFARO*
BBVA Research

NATALIA GALLARDO
MIT Sloan School of Management

Abstract

In this paper, we study household debt default behavior in Chile using survey data. Previous research in this area suggests financial and personal variables help estimate individual and group probabilities of default. We study mortgage and consumer default separately, as the default decisions and overall borrower behavior are different for each type of debt. Our study finds that income and income-related variables are the only significant and robust variables that explain default for both types of debt. Demographic or personal variables are affected by only one type of debt but not more. For example, the level of education is a factor that affects mortgage default, whereas the determinants of consumer debt default include the age of the household head, and the number of people within the household that contribute to the total family income. We find that the probability of default decreases as the family income increases, and that our estimations are consistent with other studies similar to ours.

Keywords: *Credit risk, mortgage default.*

JEL Classification: *G14, G17.*

* The authors wish to thank Daniel Oda, Luis Opazo, Andrés Sagner, and the participants of the Research Committee of the Central Bank. All remaining errors are our own.

Resumen

En este artículo estudiamos el comportamiento de no-pago de deudas de los hogares en Chile utilizando datos de encuesta. Investigaciones anteriores sobre este tema sugieren variables de tipo financieras y personales que pueden ayudar a estimar la probabilidad de no-pago para individuos y grupos de personas. Estudiamos el no-pago de deuda hipotecaria y de consumo por separado, ya que tanto la decisión de no-pago como el comportamiento del deudor es distinta en los dos casos. Nuestro estudio encuentra que el ingreso y las variables relacionadas con este son las únicas robustas y significativas que explican el no-pago de ambos tipos de deuda, mientras que las variables demográficas o personales tienden a estar relacionadas con uno u otro tipo de deuda, pero no con ambas. Por ejemplo, el nivel de educación es un factor que afecta el no-pago hipotecario, mientras que los determinantes del no-pago de deuda de consumo incluyen la edad del jefe del hogar y el número de personas en el mismo hogar que contribuyen al ingreso total de la familia. Encontramos que la probabilidad de no-pago disminuye a medida que el ingreso del hogar aumenta, y que nuestras mediciones son consistentes con las obtenidas por otros autores.

Palabras Clave: *Riesgo de crédito, no-pago hipotecario.*

Clasificación JEL: *G14, G17.*

I. INTRODUCTION

Studies that look into debt default at the household level are mostly empirical in nature and oriented towards credit scoring, in which the main objective is to develop ratios such that lenders could discriminate between good and bad payers (DeVaney and Lytton, 1995). A remarkable exception to the empirical approach is Andrade and Thomas (2007), where a structural model is proposed following the structure of Morton's (1975) paper. Similar to the corporate model, default occurs when some key variable is below predetermined threshold. The authors define that variable as creditworthiness and they use credit scoring as a proxy variable.

In the empirical analysis, some papers analyze mortgages defaults on one hand and on the other they look at credit card and others non-securitized debt default. Mortgages are considered securitized because they are backed by real estate as collateral, however under a housing price boom there is a positive probability of a decrease in the value of the collateral and therefore an additional source of risk in these types of loans. Jackson and Kasserman (1980) discuss two alternative scenarios that could describe home mortgage default behavior. The "equity theory of default" involves rational borrowers who attempt to maximize the equity position of the mortgaged property at each point in time. They cease payments if the market value of the mortgaged property declines sufficiently in relation to the outstanding mortgage loan balance at any time. The alternative explanation is based on cash flows, and termed the “ability to pay” theory of default. Under this theory, debtors will avoid defaulting on their debts as long as their income flows are sufficient to cover the mortgage payments without undue stress. Wong *et al.* (2004) attempt to identify the main determinants of mortgage default behavior in terms of these two theories, and state that under the profit maximization theory the current loan-to-value ratio, LTV (the ratio between the amount lent and the current value of the property), should be the most important factor in the borrower’s decision to default. On the other hand, under the ability to pay paradigm, the current debt service ratio, DSR (the proportion of income that is used to pay off debt), should play a major role in the decision to default. Although this insight contributes an identifying condition to discern between the two proposed models, Wong *et al.* are unable to find support for either theory as these variables are insignificant in their study.

These variables have been studied in previous research. Campbell and Dietrich (1984), Vandell and Thibodeau (1985), Lawrence and Smith (1992), Mills and Lubuele (1994) and Deng *et al.* (1995) all conclude that the LTV ratio is a strong determinant of mortgage loan default risk and also show that their relationship is positively correlated. On the other hand, Stansell and Millar (1976), Vandell (1978) and Ingram and Frazier (1982) confirm the importance of DSR as an explanatory variable of this type of default. Aside from financial variables, various authors conclude that personal characteristics such as education, income and gender are as important in explaining default (if not more so) as those described above (see for example, Morton (1975), Ingram and Frazier (1982), Webb (1982), Aylward (1984), Waller (1988), Canner *et al.* (1991) and Lawrence and Arshadi (1995)). Indeed, simulation results from Vandell and Thibodeau (1985) show that several nonequity factors dominate the equity effect on default, which helps to explain why some households with zero or negative equity may not default, while others with positive equity may do so. Finally, Campbell and Cocco (2010) develop a structural model of mortgage default that is able to replicate 2007-2008 crisis.

In terms of non-securitized loans, Avery *et al.* (2004) find that longtime married individuals have lower rates of default than recently married or divorced individuals. This is because married couples are less sensitive to income shocks, perhaps because they tend to have two incomes. Regarding gender, male subjects tend to have higher probabilities of default. Sharma and Zeller (1997) argue that females are less likely to default because they choose less risky projects. This is also confirmed by Stavins (2000), who tests the determinants of credit card delinquency and default, and finds that married couples, older individuals, better educated and higher income individuals all have a lower probability of default.

In the present paper we study the determinants of debt default at the household level in Chile, following the empirical approach presented in Pham and Lensink (2008). Using a dataset obtained from the Survey of Household Finances (*Encuesta Financiera de Hogares*, EFH, 2007), we estimate the various specifications of a probit model in search of the characteristics, both personal and financial, that have the highest impact on the overall probability that a household will default on its outstanding debt. We test a range of explanatory variables that have been identified by previous theoretical and empirical studies as being influential in a household's decision to stop debt repayments. Since the very structure of the types of debt differs and thus so do the determinants of default, we choose to analyze securitized (mortgage) and non-securitized (consumer) debt separately. We find that, for both types of debt, income is a significant and robust predictor of default risk, be it as a direct continuous variable, an indicator for income-quintile groups, or as other variables that are highly correlated with income and therefore act as proxies for it, like owning a bank account. For mortgage debt the level of education of the head of the family is a significant determinant, while for consumer debt the age and age squared of the household head are also factors. Debt service ratio is also tested as an independent variable and is found to be of importance in determining consumer debt risk only, as are various controls for the number of people who contribute to the family income.

The main contributions of this paper are to test and validate various variables with readily available information as potential determinants of household debt default in Chile and, through econometric analysis performed on the household debt dataset, establish that the larger portions of outstanding debt in Chile are in the hands of borrowers that are less vulnerable to macroeconomic or systemic shocks, indicating that the Chilean financial system is relatively robust to these risks.

II. EMPIRICAL ANALYSIS

In order to study personal default behavior in Chile, we use data from the EFH-2007 which is similar in design to surveys regularly carried out in the U.S. (the Federal Reserve's Survey of Consumer Finances, SCF), and various European countries, for example, the EFF (*Encuesta Financiera de las Familias*) in Spain, and the SHIW (Survey of Households' Income and Wealth) in Italy. Taken in Chile for the first time in 2007, the survey contemplates various areas that include personal and household data, information regarding employment, income, assets, debt, insurance, savings and investments, amongst others. The sampling design is skewed towards households with higher incomes mainly for two reasons: first, to provide a more precise estimate of wealth in general and of narrowly held assets and, second, to better compensate for nonresponse, which is differentially higher amongst the wealthy as can see in Kennickell (2008) and Barceló (2006). Therefore, expansion factors are used in all statistics and estimates to make results representative at the national level (Madeira, 2011). Financial information from the survey is aggregated at the household level. However, when we use individual data as part of our analyses, this information corresponds to the head of the surveyed household, who is defined as the main provider of household income.

2.1. Data Description

To analyze the differences between different income levels, the sample is divided into income quintiles. As can be seen in Tables 1 and 2, group Q1 includes homes with the lowest levels of income, while Q5 contains those with the highest sampled incomes.

TABLE 1

INCOME PER QUINTILE WITHOUT IMPUTED BASES (1)

Quintile	Number of Homes	Minimum	Maximum	Mean	Median
Q1	977,410	24	633	405	420
Q2	941,365	634	1,160	876	860
Q3	856,824	1,164	1,907	1,495	1,478
Q4	716,097	1,913	3,640	2,576	2,528
Q5	407,464	3,644	106,400	8,085	5,269
Total	3,899,160	24	106,400	1,959	1,190

(1) Amount of income in US$.

TABLE 2

INCOME PER QUINTILE WITH 3 IMPUTED BASES (1)

Quintile	Number of Homes	Minimum	Maximum	Mean	Median
Q1	979,042	24	648	407	420
Q2	938,740	646	1,180	892	880
Q3	877,915	1,176	1,960	1,523	1,500
Q4	713,414	1,941	3,736	2,672	2,605
Q5	401,465	3,727	106,600	8,386	5,531
Total	3,910,576	24	106,600	2,006	1,200

(1) Amount of income in US$.

Using the available data and survey question's format, we define "default" in the following way: (i) Mortgage default: the information for this classification is obtained from the survey question: "Are you up to date with your mortgage payment?" a family is considered to be in mortgage debt default if the head of the household replies the he (or she) is delinquent in his (her) payments or has stopped them altogether;

(ii) Consumer default: a family that declares not to have outstanding mortgage debt, but declares itself delinquent in payments of consumer ("all purpose") loans (credit cards, department store credit cards, bank consumer loans, car loans or other forms of consumer related debt). In this case, the survey question considered is: "Approximately, in the last 12 months and for each outstanding form of debt, how many times have your credit payments fallen into delinquency?" We define default as payments that are late by the standards set in the contract of each form of debt. Unfortunately, the answers to this question do not allow us to distinguish which debt a household has defaulted on if it has both types of debt. We therefore study consumer debt default in a subsample of homes without mortgage debt.

In Table 3, panel A shows total debt per income quintile, while panels B and C report results that contemplate mortgage and consumer debt respectively. Since there is an overlap in the sample of families that report having both mortgage and consumer debt, Panel D summarizes the data for consumer debt for families without mortgage debt. As we can see in Table 3, although Q5 represents a smaller number of homes than the others, the group represents a large portion of the total outstanding debt.

Table 4 contains the totals of defaulted debt per income quintile, both for mortgage debtors (panel A), and for consumer debtors without mortgage debt (panel B). In column 1 we see the levels of total defaulted debt for each quintile and each type of debt. It is interesting to note that the amounts of defaulted debt are similar across quintiles, while the number of homes with defaulted debt (in column 2) becomes smaller as the income level increases. In fact, as can be seen in column 4, the total amount of defaulted debt in the financial system is nearly evenly distributed between income quintiles. From Table 1 we know that higher income quintiles have more debt outstanding, which means that the amount of defaulted debt as a percentage of outstanding debt per quintile (a measure of credit risk itself) also shows a monotonic decrease as the level of income increases. As an example, the ratio results in 37% of all mortgage debt being in default for Q1, while the same statistic for Q5 results in barely 4%. This analysis is in line with the main conclusion drawn by Fuenzalida and Ruiz-Tagle (2009) mainly that the larger portion of outstanding debt in Chile is in the hands of people with a relatively lower incidence of default.

2.2. Methodology

In order to study the determinants of household debt default we have to consider two choices of the households: having debt and being in default. In this way, we analyze two types of default equations: conditional on having debt and unconditional. For the latter we follow the literature on selection bias, in which our selection equation is the decision of the household to have debt.

Given the information available from the survey, we are able to perform both types of analysis for the case of mortgages but we have to restrict the conclusions to the case of consumer default. We note that in our sample of households with mortgages 83% of them also have consumer debt, which shows that most households have both kinds of debts. In addition to that we include as explanatory variable the DSR includes

TABLE 3

DEBT PER QUINTILE

	Amount of Debt (1)	Numbers of Homes with Debt (2)	Average Debt (3)	Percent of Total Debt (4)
Panel A: All household debt				
Q1	1,244	472,237	2,634	5
Q2	3,666	616,369	5,948	14
Q3	4,864	547,038	8,892	18
Q4	8,136	478,937	16,987	30
Q5	8,824	246,703	35,767	33
Total	26,734	2,361,284	70,228	
Panel B: Mortgage debt				
Q1	529	44,726	11,830	2
Q2	1,404	106,401	13,199	5
Q3	2,407	109,793	21,926	9
Q4	4,212	146,478	28,754	16
Q5	5,828	102,393	56,916	22
Total	14,381	509,791	132,626	
Panel C: Consumer debt				
Q1	688	453,357	1,519	3
Q2	2,132	586,597	3,634	8
Q3	2,224	509,398	4,366	8
Q4	3,649	438,702	8,317	14
Q5	2,719	216,555	12,554	10
Total	11,412	2,204,609	30,390	
Panel D: Consumer debt without mortgage debt				
Q1	621	420,044	1,478	2
Q2	1,911	498,303	3,835	7
Q3	1,726	411,408	4,195	6
Q4	2,494	318,912	7,820	9
Q5	1,543	132,742	11,622	6
Total	8,294	1,781,409	28,950	

(1) Amount of debt in US$ million.
(2) Number of homes reporting outstanding debt.
(3) Average amount of debt per quintile in US$.
(4) Percentage of quintile amount of debt versus total debt.

TABLE 4

DEFAULTED DEBT (DD) PER QUINTILE

	Amount of DD (1)	Numbers of Homes with DD (2)	Average DD (3)	Percent of Quintile of Debt (4)
Panel A: Mortgage debt				
Q1	198	17,075	11,581	19
Q2	141	9,231	15,263	13
Q3	264	18,770	14,067	25
Q4	207	6,256	33,070	20
Q5	235	2,833	82,931	22
Total	1,045	54,165	156,912	
Panel B: Consumer debt without mortgage debt				
Q1	185	132,625	1,394	12
Q2	353	141,384	2,500	24
Q3	316	70,354	4,496	21
Q4	398	51,298	7,767	27
Q5	241	11,354	21,238	16
Total	1,494	407,015	37,394	

(1) Amount of defaulted debt in US$ million.
(2) Number of homes reporting defaulted debt.
(3) Average amount of defaulted debt per quintile in US$.
(4) Percentage of quintile amount defaulted debt versus total quintile debt.

all the monthly payments that households should pay. For the case of consumer default we consider only households without mortgage. We think that this constraint implies an interesting group of study given that consumer loans do not have collateral.

Considering the previous discussion we define X as a binary variable that takes the value one if the household reports debt and zero otherwise. For the case of default we use the variable Y which is equal to one if the household reports being in default, and zero otherwise. We use a first stage equation where the probability of having debt (PX) is used to determine the probability of default. Heckman (1979) shows that the two stage method is equivalent to solving the maximum likelihood multivariate normal approach. It is clear that the restriction of normality is strong, for which reason researchers tend to prefer the use of two stage methods. The key condition for this method is that the effect of the parameters from the first stage that are in the second stage be non-linear. In the case of the multivariate normal this non-linearity comes from the truncated distribution and it is a function of the density and the cumulative distribution functions.

Keeping in mind this mechanism we follow the empirical approach in this area (Vella, 1998; Angrist, 2001) which relies on the use of non-linear functions of the probability computed in the first stage, which in our case is represented by PX. In this way the effect of the first stage on second equation of the ith household is represented by $g_i = g(PX_i)$, where g() is the logistic transformation we also include it's square in the second stage equation. Specifications with higher order expression of this transformation showed non-significant effects on the explanatory variables nor on the overall effect. It is important to note that empirical applications tend to use polynomials of PX including higher order terms which are considered in our case given the non-linearity of the logistic function.

In addition to the inclusion of non-linear transformation of PX it is necessary to adjust the standard errors appropriately. Because we are using weights in the estimation we report the standard errors obtained by a bootstrapping procedure with 2000 replications. Also, for the case of the probability function we consider the probit model. Results using the logit function do not change qualitatively; however, those are not reported in this paper.

III. RESULTS

In Table 5, the effect of income in the probability of mortgage default has the expected sign and is quiet robust, whether expressed as a continuous variable or as quintile groups. The interpretation of the coefficients follows intuition: the higher the total family income, the lower the probability that the family will default on its mortgage debt. In Chile, access to bank accounts is far from universal and, recent market expansion notwithstanding, having one is still a sign that the user has a minimum income level (with all the related benefits of access to credit at better rates and terms). As stated in Morales and Yáñez (2006), in 2006 there were a little over 1.5 million checking accounts in Chile, indicating that only about 15% of the country's workforce had access to one. In terms of income cutoff, most banks consider a person to be eligible to open a bank account if his/her income is at or above CLP 400,000 (USD 800), which the EFH2007 shows to be the median income in Chile. We therefore control for such a borrower who has a bank account as an indicator of his/her socioeconomic status, as well as his/her relative access to credit (and the characteristics of this credit). Since banks apply their own credit and background checks, filters and models, a person that has a bank account can generally be expected to be at lower risk of default than someone who does not, all else being equal. Our results show that having a bank account is a significant and robust component of the probability of default.

Education is correlated with income and, therefore, one can expect that a higher level of education implies a higher income. Also, the level of education is sometimes included in banks' evaluation of an individual's credit worthiness, and can therefore constitute a barrier to obtaining mortgage loans. In this way, having a higher level of education is a personal characteristic that both provides access to mortgage debt, and characterizes the debtor as a relatively lower risk investment than a comparable person without the education credentials. Since this "bank filter" is not a factor for consumer (i.e.: non-securitized) debt, this variable is not significant in those regressions, as we will be seen below.

TABLE 5

PROBIT ESTIMATIONS OF MORTGAGE DEFAULT (1) (2)

Variable	No selection bias correction					Selection bias correction				
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5
Gender (1 if male)	–0.0633			0.0749		–0.4875			–0.2626	
	(0.2666)			(0.2632)		(0.3267)			(0.3059)	
	[0.3017]			[0.3135]		[0.3305]			[0.3298]	
Married	–0.1772			–0.1928		0.4877			0.5103	
	(0.2747)			(0.2675)		(0.3467)			(0.3628)	
	[0.3253]			[0.3182]		[0.4062]			[0.4847]	
Age	0.083			0.0876		0.1607			0.1581	
	(0.0874)			(0.0849)		(0.1105)			(0.1063)	
	[0.119]			[0.1251]		[0.1321]			[0.1803]	
Age (squared)	–0.0008			–0.0009		–0.0012			–0.0013	
	(0.0009)			(0.0009)		(0.0012)			(0.0012)	
	[0.0012]			[0.0013]		[0.0016]			[0.002]	
High school	–1.0819	–0.9372	–0.9718	–1.1027	–0.9139	–0.3728	–0.8827	–0.7101	–0.3161	–0.5522
	(0.4146)	(0.4256)	(0.4130)	(0.3978)	(0.4005)	(0.3970)	(0.4706)	(0.4277)	(0.3717)	(0.4114)
	[0.6768]	[0.7413]	[0.7271]	[0.6281]	[0.715]	[0.5173]	[1.4486]	[1.3934]	[1.348]	[1.5282]
College	–1.0856	–1.0732	–0.9132	–0.9810	–0.7429	–0.5433	–1.1077	–0.8039	–0.4122	–0.5956
	(0.4128)	(0.4271)	(0.4217)	(0.4174)	(0.4077)	(0.4175)	(0.4878)	(0.4375)	(0.4090)	(0.4291)
	[0.6731]	[0.7446]	[0.7325]	[0.6579]	[0.7207]	[0.5231]	[1.4693]	[1.4015]	[1.373]	[1.5535]
Bank account	–0.3694	–0.6023	–0.301	–0.5041	–0.5634	–0.4020	–0.6948	–0.4386	–0.6847	–0.6847
	(0.2742)	(0.2489)	(0.2682)	(0.2628)	(0.2508)	(0.2478)	(0.2504)	(0.2569)	(0.2322)	(0.2412)
	[0.3339]	[0.275]	[0.3117]	[0.3361]	[0.2875]	[0.2591]	[0.297]	[0.3011]	[0.3159]	[0.2972]
Total income (log)	–0.3044		–0.4091			–0.5471		–0.3861		
	(0.1643)		(0.1890)			(0.2228)		(0.1975)		
	[0.1915]		[0.2064]			[0.2829]		[0.2214]		
DSR	0.3737	0.6015		0.3307	0.321	–0.0589	0.3121		–0.2602	0.0294
	(0.2663)	(0.3158)		(0.2622)	(0.2636)	(0.2947)	(0.2288)		(0.3544)	(0.2365)
	[0.429]	[0.4512]		[0.4631]	[0.4176]	[0.5985]	[0.3862]		[0.4874]	[0.3691]
LTV	–0.0971			–0.0443		0.0921			0.0813	
	(0.2449)			(0.2300)		(0.0448)			(0.0522)	
	[0.3246]			[0.3485]		[0.1079]			[0.1847]	
PX (logit)						0.5071	0.4405	0.3463	0.4735	0.3961
						(0.2605)	(0.1826)	(0.1594)	(0.2496)	(0.1772)
						[0.3800]	[0.2414]	[0.2322]	[0.3559]	[0.2674]
PX (logit–squared)						–0.0941	–0.2018	–0.1761	–0.0938	–0.1961
						(0.0747)	(0.0803)	(0.0731)	(0.0719)	(0.0797)
						[0.0835]	[0.1017]	[0.0953]	[0.0855]	[0.1064]
Quintile_dummies	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Constant	1.916	–0.3684	5.2210	–1.4807	0.3378	1.4735	–0.5045	4.5719	–4.4354	0.2702
	(3.0245)	(0.4099)	(2.4822)	(2.0359)	(0.4587)	(4.0680)	(0.4309)	(2.6741)	(2.5715)	(0.5186)
	[4.021]	[0.7493]	[2.7957]	[3.0768]	[0.9144]	[4.8482]	[1.4286]	[3.3937]	[4.732]	[1.757]
Number of obs. (unweighted)	522	548	548	522	548	599	651	651	599	651
AIC (3)	355,440	355,930	353,630	360,040	360,340	266,140	287,490	282,590	255,480	275,650
BIC	308,600	334,400	332,090	300,430	321,580	323,280	318,840	313,940	325,810	324,920
Chi2				12.8**	12.5**				17.18***	12.48**

(1) The probit regressions are run on samples composed of a non-imputed dataset (Imp = 0).
(2) The first number indicates the coefficient, the second, in parenthesis, the standard error and the third, in brackets, the bootstrapped standard error.
(3) AIC and BIC are Akaike and Schwarz information criteria, respectively.
(4) ***, ** and * represent statistical significance at the 1%, 5% and 10% levels respectively.

Gender of the person who contributes the highest amount to the family income has no significant effect on the probability of mortgage default. Marital status does not have an effect on the probability of default either. Pham and Lensink (2008) state that married couples seem to have a lower risk of mortgage default than single people. In the case studied by authors (rural Vietnam), husband and wife tend to have paid jobs, which constitutes a sort of diversification of risk in that, if one loses his or her source of income, the partner can temporarily help make up for the shortfall until the second income is restored. This mitigation of risk through diversification is a very significant result in their paper, although it does not seem to be a factor in our study. We believe that the effect of the number of people who actually contribute to the family income is more important than the marital status of the head of the family. For that reason we construct additional variables to control for there being more than one person who works in a given family, employed > 1, as well as variables to separate the effects of having just two income earners in one home employed = 2, versus having three or more people contributing to the household income employed > 2. Tests with these variables show no interesting results, and we therefore do not reproduce them here. However, these variables do provide interesting information in the consumer debtors' case, which we discuss below.

Age and age squared are included to capture life cycle variations in behavior. These life cycle variables are not significant in almost any specification. This pattern follows the risk associated with increasing debt as a person ages and makes bigger investments (a larger family requires bigger home, and implies higher expenses), and then a decrease in risk as the debt is paid off and expenses reduced after a certain age peak.

We find that neither DSR nor LTV are significant for mortgage debtors. In the case of LTV, this could be due to the fact that the "value" component in the ratio is obtained from an uninformed estimation (the actual question in the survey is "what do you think you would be paid if you sold your property today?"). We tested other sources of data to calculate the LTV ratio, such as the original purchase price of the property, the price the owner believes the property is worth, and the inflation-indexed original purchase price, but none of these definitions resulted in any meaningful contribution to the analysis. On the other hand, if this ratio is an indicator that the "benefit maximization" model of default decision is true, then not finding it a significant component of the probability of default confirms our intuition that the general public does not consider debt default as a strategic decision, but simply an unavoidable situation brought on by insolvency. Finally, both DSR and LTV are functionally related to income, as well as between each other, implying a high degree of multicolinearity.

Table 6 show the results of the models estimated to characterize consumer credit default. As with mortgage debtors, we find that the financial variables are robust in that they seem to be significant predictors of default in all specifications. Income, whether it be a continuous variable or grouped by quintiles, is significant and its coefficient has a negative sign, indicating that the higher the level of a household's income, the lower its probability of falling into financial distress. The coefficient for bank account is negative and significant and, although it is correlated with income, it does include an additional quality of having passed a bank's "due diligence" process, which certifies that the respondent has a minimum level of credit-worthiness.

TABLE 6

PROBIT ESTIMATIONS OF CONSUMER DEFAULT (1) (2)

Variable	No selection bias correction					Selection bias correction				
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5
Gender (1 if male)	–0.0252					–0.1727				
	(0.1062)					(0.1193)				
	[0.1101]					[0.1615]				
Married	–0.103					–0.1872				
	(0.1023)					(0.1035)				
	[0.0831]					[0.0907]				
Age	0.0562	0.0503	0.0512	0.0503	0.0513	0.0678	0.0576	0.0594	0.0547	0.0566
	(0.0251)	(0.0247)	(0.0246)	(0.0244)	(0.0243)	(0.0274)	(0.0279)	(0.0275)	(0.0277)	(0.0274)
	[0.0202]	[0.0259]	[0.0255]	[0.0251]	[0.0253]	[0.0324]	[0.0284]	[0.029]	[0.028]	[0.0283]
Age (squared)	–0.0007	–0.0006	–0.0007	–0.0007	–0.0007	–0.0008	–0.0007	–0.0007	–0.0007	–0.0007
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
	[0.0002]	[0.0003]	[0.0003]	[0.0003]	[0.0003]	[0.0004]	[0.0003]	[0.0003]	[0.0003]	[0.0003]
Bank account	–0.4056	–0.4061	–0.3754	–0.4144	–0.3832	–0.2885	–0.2988	–0.2557	–0.3711	–0.3296
	(0.1429)	(0.1422)	(0.1402)	(0.1355)	(0.1340)	(0.1516)	(0.1505)	(0.1483)	(0.1429)	(0.1413)
	[0.1177]	[0.1468]	[0.1423]	[0.1425]	[0.1388]	[0.1765]	[0.1535]	[0.1568]	[0.1428]	[0.1463]
Total income (log)	–0.3859	–0.3968	–0.4151			–0.2447	–0.2789	–0.3113		
	(0.0792)	(0.0799)	(0.0811)			(0.0790)	(0.0802)	(0.0812)		
	[0.0703]	[0.0816]	[0.0825]			[0.079]	[0.0866]	[0.0881]		
DSR	0.2020	0.1961	0.1956	0.2074	0.2063	0.3070	0.3053	0.3122	0.3240	0.3308
	(0.1119)	(0.1109)	(0.1100)	(0.1089)	(0.1077)	(0.1211)	(0.1214)	(0.1218)	(0.1195)	(0.1196)
	[0.1119]	[0.1297]	[0.1268]	[0.1266]	[0.126]	[0.1612]	[0.1419]	[0.1435]	[0.1435]	[0.1486]
Employed>1 (3)	0.2813	0.2858		0.2691		0.3370	0.3496		0.3196	
	(0.1080)	(0.1082)		(0.1085)		(0.1083)	(0.1059)		(0.1066)	
	[0.1032]	[0.1078]		[0.11]		[0.0971]	[0.1087]		[0.1105]	
Employed=2 (4)			0.2433		0.2247			0.2794		0.2499
			(0.1135)		(0.1135)			(0.1133)		(0.1139)
			[0.1139]		[0.1173]			[0.1177]		[0.117]
Employed>2 (5)			0.4123		0.3991			0.5820		0.5494
			(0.1532)		(0.1549)			(0.1495)		(0.1500)
			[0.1553]		[0.1637]			[0.1568]		[0.1560]
PX (logit)						0.9632	0.9248	0.9289	0.9242	0.9256
						(0.1411)	(0.1388)	(0.1418)	(0.1386)	(0.1407)
						[0.2639]	[0.1786]	[0.1864]	[0.1895]	[0.1875]
PX (logit–squared)						–0.0964	–0.1163	–0.1112	–0.1146	–0.1083
						(0.1039)	(0.1026)	(0.1037)	(0.1043)	(0.1051)
						[0.1888]	[0.1262]	[0.128]	[0.131]	[0.1269]
Quintile_dummies	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Constant	3.3097	3.5221	3.7293	–1.3738	–1.3952	0.3946	0.9153	1.2737	–2.4567	–2.4968
	(1.1208)	(1.1112)	(1.1322)	(0.5343)	(0.5321)	(1.0848)	(1.0752)	(1.1118)	(0.5940)	(0.5832)
	[1.1621]	[1.1537]	[1.1491]	[0.5498]	[0.5457]	[1.0214]	[1.1621]	[1.2311]	[0.6032]	[0.602]
Number of obs. (unweighted)	1,659	1,659	1,659	1,659	1,659	2,439	2,439	2,439	2,439	2,439
AIC (6)	1,737,780	1,725,370	1,730,220	1,753,430	1,758,120	1,678,020	1,677,610	1,677,290	1,707,920	1,707,800
BIC	1,689,050	1,687,480	1,686,900	1,699,290	1,698,570	1,614,230	1,625,410	1,619,300	1,638,330	1,632,410
Chi2				22.8***	23.9***				13.02***	15.36***

(1) The probit regressions are run on samples composed of a non-imputed dataset (Imp = 0).
(2) The first number indicates the coefficient, the second, in parenthesis, the standard error and the third, in brackets, the bootstrapped standard error.
(3) Two or more persons employed in the household.
(4) Two persons employed in the household.
(5) Three or more persons employed in the household.
(6) AIC and BIC are Akaike and Schwarz information criteria, respectively.
(7) ***, ** and * represent statistical significance at the 1%, 5% and 10% levels respectively.

With respect to the default theory "indicator" ratios, LTV is omitted from these regressions, since this ratio pertains to mortgage debtors only. On the other hand, DSR results are in line with expectations, that is, a positive coefficient, which is interpreted as the higher the debt service compared to total income, the likelier it is for households to default. The reason why DSR is significant for consumer credit debtors and not for mortgage debtors is that, as mentioned, DSR is correlated with income, since income is the denominator of the DSR ratio and, therefore, in the mortgage regressions, DSR is only significant when income is not present (and, in fact, the significance of income increases when DSR is not present). Unlike the case of mortgage debtors, total consumer debt (the numerator of the DSR ratio) is composed of debt that cannot be monitored by a bank, or aggregated as a whole. An example of this are department store "credit cards", which can only be used at the issuing store or a few partner businesses at most, are easily obtained (hardly credit checks are needed) and the debts incurred with one issuer are not "visible" to another, nor are they reported into the financial system. Other examples include bank credit cards and overdraft lines, loans from family and friends, etc. Since this is the case, the information obtained in the EFH survey, which allows the DSR ratio to be constructed, is not freely available in the financial system, which means that, depending on the composition of their debt, highly leveraged individuals can choose to incur additional debt and, therefore, DSR is not a close proxy for income as in the mortgage case, and thus is far less likely to be significant in determining the probability of default. We now turn to the demographic variables used in previous research. As with the mortgage case, gender and marital status are insignificant.

The life cycle is significant and robust in all specifications, indicating that default risk in this case is sensitive to the changes in debt as a person ages. Based on the fitted coefficients in the table, we estimate that the default risk peaks at around 42 years of age, after which is begins to slowly decline. Unlike with mortgage debtors, the probability of default for consumer credit debtors does not seem to be affected by the level of education. This lends support to our view that a reason for it to be significant in the mortgage, or securitized debt, case is due to bank monitoring and access to credit criteria. Since consumer lending standards are far more lax than for mortgage lending, education does not provide the "accreditation" effect it does for mortgage debtors. Finally, in order to ascertain the importance of the number of people who contribute to the total family income within a household, we test variables that indicate whether there is more than one income provider in the household, employed > 1, and two variables to separate this "more than one" effect into "exactly two", employed = 2, income providers and "three or more", employed > 2. The intuition behind these tests is, in part, the same as the justification given by Pham and Lensink (2008) for the significance of the marriage variable: there is a diversification of risk if there is more than one provider of income in the household. We also have a prior belief that the higher the number of people that contribute to the household income, the higher that income should be and, as we've seen, higher income tends to reduce the risk of default. We therefore expect the occupation controls to have negative coefficients. However, the results show that the coefficients that are significant and robust in every specification are nevertheless positive in sign. We believe that this is the result of two unobserved effects: relative job security and the motivation for the number of people working in a household. In the lower income quintiles people tend to have a lower level of education and are therefore able to obtain work only at a non-professional (or unskilled) level. This means that they are the most vulnerable to macroeconomic shocks that impact labor, thus making their source of income more uncertain, and their debt more risky. On the other side of the spectrum, people in the higher quintiles tend to have professional jobs, and tend to have much lower probabilities of being laid off, a situation described in Fuenzalida and Ruiz-Tagle (2009) in their analysis of the probability of job loss obtained from panel data. It is therefore not necessary for higher quintile families to have more members with paying jobs. On the other hand, due to the inequality in income distribution in Chile, a higher number of people working in the household does not imply a larger combined income than that of a single person in a higher income quintile, meaning that a relatively large number of people in a family who contribute to the total income is a necessity and probably equates to a barely adequate total income. This can be seen in the low and middle quintiles, where a comparatively large number of people contribute to the family income and help diversify the job-loss risk as well, but the families are nevertheless classified into these low income quintiles, and their default risk is comparatively high. These considerations make the positive coefficients obtained a logical result of the country's labor and income conditions.

IV. CONCLUSIONS

In the present paper, we study the determinants of debt default at the household level in Chile, using data obtained from the Survey of Household Finances performed in 2007 (EFH 2007). We find that the main determinants of mortgage debt default are income, and proxies for income such as income quintile indicator variables, having a bank account and even an education level beyond high school. In the case of consumer debt, we find that the main determinants are also income and related variables, but we also find statistical support for the DSR as well as for the number of people in the household who contribute to the total income.

The results shown here open up new avenues for research in the areas of household finance and aggregate financial stability. Future research can also hope to develop from further instances of the EFH, when a panel study will be possible. F-RT have taken a first step, there are various forms of stress testing that can be applied to this data to better understand the possible effects of various changes in the prevalent market conditions, and how these might affect the stability of the Chilean financial system. Given the risks involved, the results of these tests might have important policy implications in terms of lending practices, credit scoring and screening.

REFERENCES

ANDRADE, F. and L. THOMAS (2007). "Structural models in consumer credit", *European Journal of Operational Research* 183, pp. 1569-1581.

ANGRIST, J. (2001). "Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors: Simple Strategies for Empirical Practice", *Journal of Business & Economic Statistics* 19(1), pp 2-16.

AVERY, R., P. CALEM and G. CANNER (2004). "Consumer credit scoring: do situational circumstances matter?", *Journal of Banking and Finance* 28(4), pp. 835-856.

AYLWARD, F. (1984). "Anatomy of the Residential Mortgage", *Real Estate Today*, pp. 23-25.

BARCELO, C. (2006). "Imputation of the 2002 Wave of the Spanish Survey of Household Finances (EFF)", Working Paper, Central Bank of Spain.

CAMPBELL, J. (2006). "Household Finance", *The Journal of Finance* 61(4), pp. 1553-1604.

CAMPBELL, J. and J. COCCO (2010). "A Model of Mortgage Default", Harvard University, Working Paper.

CAMPBELL, S. and J. DIETRICH (1984). "The Determinants of Default on Insured Conventional Residential Mortgage Loans", *The Journal of Finance* 38(5), pp. 1569-1581.

CANNER, B., S. GABRIEL and J. WOOLLEY (1991). "Race, Default Risk and Mortgage Lending: A Study of the FHA and Conventional Loan Markets", *Southern Economic Journal* 58(1), pp. 249-262.

DENG, Y., J. QUIGLEY and R. ORDER (1995). "Mortgage Default and Low Downpayment Loans: The Cost of Public Subsidy", Working Paper NBER.

DEVANEY, S. and R. LYTTON (1995). "Household Insolvency: A Review of Household Debt Repayment, Delinquency, and Bankruptcy", *Financial Services Review* 4(2), pp. 137-156.

FUENZALIDA, M. and J. RUIZ-TAGLE (2009). "Riesgo Financiero de los Hogares", *The Chilean Economy* 12(2), pp. 35-53.

HECKMAN, J. (1979). "Sample Selection Bias as a Specification Error", *Econometrica* 47(1), pp. 153-162.

INGRAM, F. and E. FRAZIER (1982). "Alternative Multivariate Tests in Limited Dependent Variable Models: An Empirical Assessment", *Journal of Financial and Quantitative Analysis* 17(2), pp. 227-240.

JACKSON, F. R. and D. L. KASSERMAN (1980). "Default risk on home mortgage loans: a test of competing hypotheses", *Journal of Risk and Insurance* 47(4), pp. 678-690.

KENNICKELL, A. (2008). "The Role of Over-Sampling of the Wealthy in the Survey of Consumer Finances", *Irving Fisher Committee Bulletin* 28, pp. 403-408.

LAWRENCE, C. and N. ARSHADI (1995). "A Multinomial Logit Analysis of Problem Loan Resolution Choices in Banking", *Journal of Money, Credit and Banking* 27(1), pp. 202-216.

LAWRENCE, E. and L. SMITH (1992). "An Analysis of Default Risk in Mobile Home Credit", *Journal of Banking and Finance* 16, pp. 299-312.

MADEIRA, C. (2011). "Computing Population Weights for the EFH Survey", Working Paper 632, Central Bank of Chile.

MILLS, E. and L. LUBUELE (1994). "Performance of Residential Mortgages in Low and Moderate-Income Neighborhoods", *Journal of Real Estate Finance and Economics* 9(3), pp. 245-260.

MORALES, L. and A. YAÑEZ (2006). "La bancarización en Chile, concepto y medición", Working Paper, Superintendency of Banks and Financial Institutions of Chile.

MORTON, G. (1975). "A Discriminant Function Analysis of Residential Mortgage Delinquency and Foreclosure", *AREUEA Journal* 3, pp. 73-90.

PHAM, T. and R. LENSINK (2008). "Household Borrowing in Vietnam: A Comparative Study of Default Risk of Informal, Formal and Semi-Formal Credit", *Journal of Emerging Market Finance* 7(3), pp. 237-261.

SHARMA, M. and M. ZELLER (1997). "Repayment performance in group-based credit programs in Bangladesh: An empirical analysis", *World Development* 25(10), pp. 1731-1742.

STANSELL, S. and J. MILLAR (1976). "An Empirical Study of Mortgage Payment to Income Ratios in a Variable Rate Mortgage Program", *The Journal of Finance* 31(2), pp. 415-425.

STAVINS, J. (2000). "Credit Card Borrowing, Delinquency, and Personal Bankruptcy", *New England Economic Review,* pp. 15-30.

VANDELL, K. (1978). "Default Risk Under Alternative Mortgage Instruments", *The Journal of Finance* 33(5), pp. 1279-1296.

VANDELL, K. and T. THIBODEAU (1985). "Estimation of Mortgage Defaults Using Disaggregate Loan History Data", *AREUEA Journal* 15(3), pp. 292-317.

VELLA, F. (1998). "Estimating Models with Sample Selection Bias: A Survey", *The Journal of Human Resources* 33(1), pp. 127-169.

WALLER, G. (1988). "Residential Mortgage Default: A Clarifying Analysis", *Housing Finance Review* 7, pp. 321-333.

WEBB, B. (1982). "Borrower Risk Under Alternative Mortgage Instruments", *The Journal of Finance* 37(1), pp. 169-183.

WONG, J., L. FUNG, T. FONG and A. SZE (2004). "Residential mortgage default risk and the loan-to-value ratio". *Hong Kong Monetary Authority Quarterly Bulletin*, pp. 35-45.