



HOUSEHOLD SAVING AND LABOR INFORMALITY: THE CASE OF CHILE*

Alfredo Schclarek**
Mauricio Caggia***

I. INTRODUCTION

One of the main problems facing Latin American and Caribbean (LAC) countries is their low levels of saving, especially if compared with other regions of the world, such as East Asia and the Pacific (EAP), that have shown greater dynamism in saving rates. For example, Gutierrez (2007) presents evidence that, on average, the saving rate for LAC between 2000 and 2003 was 19.2%, while EAP countries had a mean saving rate of 34.5%. Further, Reinhardt (2008) provides evidence that the average gross domestic saving for LAC was 17.1% in the 1990s, compared to 24.8% for a group of 25 developing middle-income countries. In the 2000s, according to Pérez Monteiro et al. (2014), the average gross saving rate was 20% for LAC and 30% for EAP. Chile, meanwhile, despite its higher average saving rate than most other LAC countries, its average 22.3% for the 2000-2003 period is still low compared to EAP countries (Gutierrez, 2007).

Another important concern in the region is related to the high levels of informal employment. Although the 2000s have seen a reduction in the informality figures in comparison to the 1990s, informality affects between 37.7% and 88.4% of total workers in LAC (Tornarolli et al., 2014). In the case of Chile, its levels of informality are the lowest in the region and it has been following a mild downward trend from 40.6% in 1990 to 37.7% in 2009. According to Peticara and Celhay (2010), the informality was reduced from 39.5% in 1998 to 35.8% in 2006. However, when disaggregating between salaried workers and self-employed workers, the figures for 2006 were 24.9% and 71.6%, respectively. Evidently, by disaggregating by type of worker some heterogeneity appears, where it is evident that self-employed workers suffer a much higher level of informality than their salaried counterparts.

A natural question that arises is whether these two phenomena are interrelated and whether the prevalent high informality rates prevent the proper channeling of savings into the formal financial system. Clearly, this could have implications on the efficient allocation of surpluses to increase investment and economic growth.

* We would like to thank Eduardo Cavallo, J. Rodrigo Fuentes and five anonymous reviewers for comments and discussions; and seminar participants at the IDB Workshop Domestic Savings in Latin America and the Caribbean (Washington DC, USA) and Universidad Nacional de Córdoba, Argentina.

** Department of Economics and Finance, Universidad Nacional de Córdoba, Argentina; CONICET, Argentina, and CIPPES, Argentina. E-mail: alfredo@eco.unc.edu.ar; www.cbaeconomia.com.

*** Institute of Economics and Finance, Universidad Nacional de Córdoba, Argentina, and CIPPES, Argentina.

Thus, studying the relationship between household saving and informality could allow governments to develop adequate policies to influence aggregate saving rates. However, although there is plenty of research analyzing the main determinants of saving for the region¹, studies on how informality affects saving and how/why informal workers save are almost non-existent. A notable exception is the recent work by Granda and Hamann (2015), who build a theoretical occupational choice model to calibrate it with data for Colombia and analyze the effects of several formalization policies on saving. The underlying argument, following Browning and Lusardi (1996), is that if informal workers have a more uncertain and variable income stream, we would expect informal workers to save more for precautionary motives. Reaching opposite results, Lorenzo and Osimani (2001) find that lower-income households, which a priori could resemble informal households, have a lower saving rate than the other higher income households. In addition, Dupas and Robinson (2013) find that simply providing a safe place to keep money was sufficient to increase health-care saving by 66 percent in an experiment in Kenya. This study may support the argument that informal workers save less than formal workers because they are less financially included. Finally, Ogbuabor et al. (2013) use time-series analysis for Nigeria and find that informality hinders the growth of aggregate domestic saving. Although these studies reach potentially interesting conclusions, what is lacking is more micro-level empirical evidence on the link between informality and saving.

The objective of this paper is to empirically study the saving behavior of formal and informal workers in Chile, following the microeconomic approach of the pioneering studies of Attanasio and Székely (2000) and Butelmann and Gallego (2000). Following both OLS and probit estimation techniques, we econometrically test whether there are any differences in the saving behavior of formal and informal households. Although we do not econometrically study the potential reasons for the different saving behavior of informal and formal households, we present survey results that characterize how and why informal and formal workers save. In terms of data, we use the microdata from the Central Bank of Chile's Survey of Household Finance for 2007 and have between 2,533 and 1,740 observations of urban households from Chile, depending on the definitions of savings and informality that are used. A novelty of this survey for LAC is that it surveys not only income, expenditure and household characteristics, but also the structure and level of household assets and liabilities with a high degree of detail. Moreover, it surveys the restrictions to credit access for households, the expectations about households' future levels of saving, their access to insurance markets and various other determinants of saving. Thus, we are not only able to compare the saving rate of formal and informal households, but we also analyze the types of financial instruments selected by such households. Clearly, this database allows us to study in much more detail the saving pattern and financial behavior of households than previous studies, which use databases with a focus on labor and expenditure characteristics, such as the CASEN

1 For Latin America, the following articles may be mentioned: Attanasio and Székely (2000) for Mexico and Peru, Attanasio and Székely (1999) for Mexico, Lorenzo and Osimani (2001) for Uruguay, Butelmann and Gallego (2001) for Chile, and Sandoval-Hernandez (2010) for Mexico.



survey in Chile, the EPH survey in Argentina, the PNAD survey in Brazil or the ENAHO survey in Peru (Maurizio, 2012).

This study is structured as follows. In section II, we make a descriptive analysis of formal versus informal workers giving special attention to the socioeconomic characteristics and borrowing and saving behavior of such workers. Section III presents the data and econometric methodology that has been used to analyze whether informal households save less than formal households. Further, it discusses the main results and findings. Finally, section IV concludes.

II. CHARACTERIZING INFORMALITY IN CHILE

This paper uses data from the Survey of Household Finances (EFH) conducted by the Central Bank of Chile in 2007, which is a Chilean urban survey including demographic, economic, financial and social indicators.² While the 2007 survey covers 3,828 households across the country, for our estimations and calculations we use a subset of these observations because we discard those households that have no member working and those households that have both informal and formal workers, i.e. we kept only those pure households where all working members are either formal or informal. The reason for discarding “mixed” households is that we wanted to compare “pure” informal households against “pure” formal households to avoid any inconsistency that “mixed” households may bring about.

As for the definition of informality, this concept has certain ambiguities, in terms of both the theoretical definition and the empirical measurement. In the case of the theoretical definition, it is not completely clear what exactly we mean when we discuss “informality”. In terms of the empirical measurement, the difficulty rests on how to structure the different questions to measure each particular definition of informality and the content of the already existing surveys. Clearly, this is a complex discussion beyond the reach of this paper and, therefore, in this paper we follow the most common definitions found in the literature, namely the social protection definition and the productive definition (ILO, 2002; Kanbur, 2009; Maloney, 2004; Tornarolli et al., 2014). While the social protection definition stresses non-compliance with labor legislation in terms of labor protection and social security benefits, the productive definition brings up the level of productivity of jobs and the skills needed for carrying out those jobs to characterize informality.

Concretely, the social protection International Labour Organization definition (ILOD) characterizes an informal worker as a salaried or self-employed worker

² Note that although this survey was also available for the years 2008, 2009 and 2010 at the time of carrying out this research, we preferred using the 2007 version because it surveyed urban households from the whole country, while for the other years it surveyed only urban households from the metropolitan region of Santiago de Chile. Furthermore, the 2007 survey covers more households than the other surveys. For more details on the survey and its methodology, please visit the following link: <http://www.bcentral.cl/estadisticas-economicas/financiera-hogares/index.htm>

that does not contribute to a pension (or retirement) plan.³ Further, we consider as informal those salaried workers that declare not having a contract of employment, even if they do contribute to a pension plan. This definition has already been used in Chile when studying informality by Peticara and Celhay (2010) and in other studies such as ILO (2002), Garganta and Gasparini (2015), and Tornarolli et al. (2014).

In terms of the productive definition (PD), we define an informal worker as one that falls under one of these categories: a) self-employed without a post-secondary education degree, b) a salaried worker in a small private firm with five employees or less, or c) an unremunerated family member. Given that an individual could have more than one job, we apply the classification only to his/her main occupation. This definition of informal workers is similar to the one used by Maloney (2004), and Tornarolli et al. (2014).

Further, we construct a third definition of informality that combines the above definitions (COMBD definition), i.e. an informal worker is a worker that is both informal according to the ILOD definition and the PD definition. Clearly, this is a much more restrictive definition of informality.

1. Informality rates and social attributes

In figure 1, we present the informality rates for Chile in 2007 according to the three definitions of informality discussed above.⁴ According to the social protection definition (ILOD), around 36.7% of the work force is informal in Chile, which means that around 2.49 million of the almost 6.8 million workers are informal. Note that Chile had a total population of around 16.6 million people and a 41.1% employment rate in 2007.⁵ In terms of the productive definition (PD), the informality rate is 28.6%, which implies that nearly 1.94 million Chilean workers were informal. For the third definition that combined the social protection and the productive definitions (COMBD), the informality rate is 19.7%, which represents 1.33 million workers. Note, however, that as this definition is stricter, it also means that many workers that are informal according to one definition but not the other (either the ILOD or the PD) would be considered non-informal. From table 1, we see that 1.16 million workers are informal according to the ILOD definition but not according to the PD definition, 606 thousand are informal according to the PD definition but not the ILOD definition, and that 3.69 million are formal according to both definitions. Thus, the exclusion of these workers from the informal group is the reason for the rate being lower than the other two definitions of informality. If we compare pure groups, i.e. the number of formal/informal workers that comply

³ It is important to mention that self-employed workers were not obliged before 2015 to contribute to a pension plan in Chile.

⁴ We find similar results when using the data from the 2008, 2009 and 2010 surveys. The results, which are not included due to space considerations, are available upon request from the authors.

⁵ Note also that the EFH is an urban survey, but we are assuming that employment and informality figures of rural workers can be derived from urban workers. In 2007, the Chilean urban population was 14.42 million and the rural population was 2.18 million.

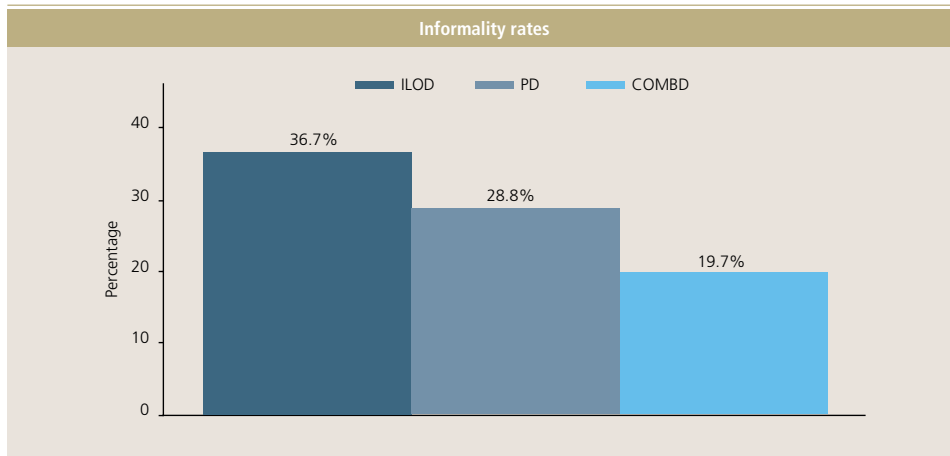
with both definitions, we would obtain an informality rate of 26.6%. Note also that in subsections II.2 and II.3 and section III, when comparing different financial attributes and behaviors of informal and formal workers using this strict definition of informality, we will use pure groups of formal and informal workers in order to avoid possible distortion of results caused by workers that are informal by one definition but not the other.

In figure 2, we present the results of informality rates when discriminating between genders, age groups, educational levels and income groups. The informality rates for women are 39.7%, 33.1%, and 23.1% according to the ILO, PD, and COMBD definitions, respectively. For men, in contrast, the informality rates are 34.6%, 25.5%, and 17.3%. Clearly, informality affects women more under all three definitions.

When classifying the different definitions of informality by income groups, a clear picture emerges. As can be seen in figure 2, informality affects more those workers with lower incomes. In stratum 1, consisting of households in the 1st to 5th income deciles, 46.6% (ILO), 41.4% (PD) and 31.1% (COMBD) of workers are informal. In stratum 2, corresponding to persons belonging to households positioned in the 6th to 8th income deciles, 33.9% (ILO), 26.2% (PD) and 16.9% (COMBD) of those employed are informal. In stratum 3, for workers belonging to households in deciles 9th and 10th of income, the informality rate falls to 28.2% (ILO), 15.8% (PD) and 9% (COMBD).

Figure 1

Informality rates in Chile



Source: 2007 Survey of Household Finances, Central Bank of Chile.

TABLE 1

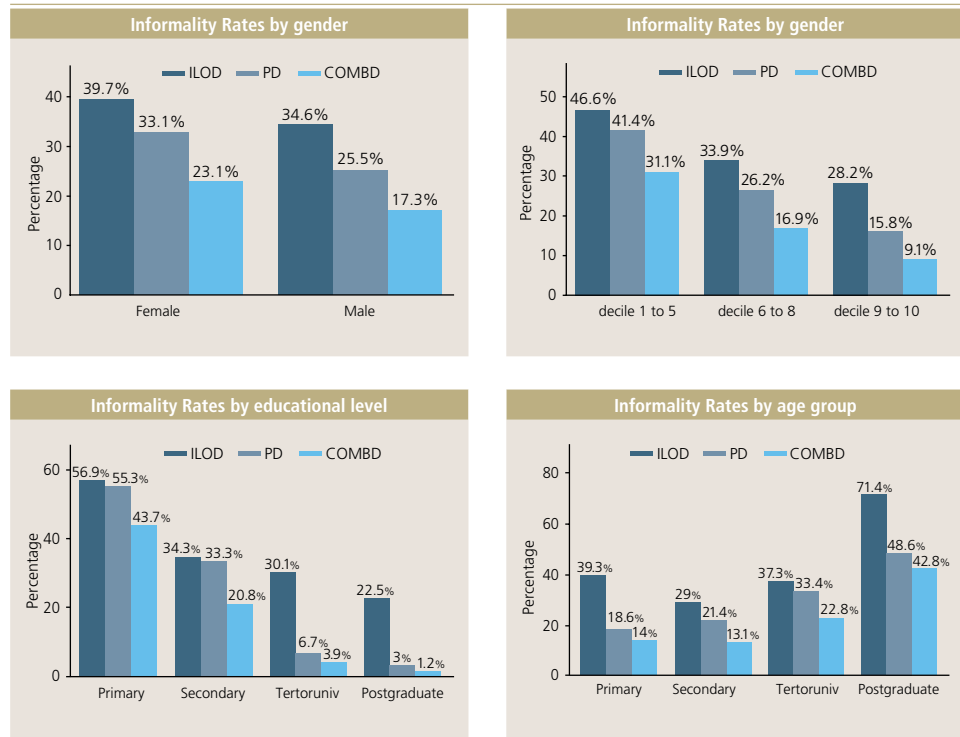
Number and share of formal and informal workers

		Productive definition											
		Female					Male						
		No		Yes		Total	No		Yes		Total		
		Persons	%	Persons	%	Persons	%	Persons	%	Persons	%		
ILO definition	No	1,416,956	0.50	283,191	0.10	1700,147	0.60	2,277,892	0.57	323,536	0.08	2,601,428	0.65
	Yes	470,100	0.17	650,326	0.23	1,120,426	0.40	687,998	0.17	689,399	0.17	1,377,397	0.35
	Total	1,887,056	0.67	933,517	0.33	2,820,573	1.00	2,965,890	0.75	1,012,935	0.25	3,978,824	1.00

Source: 2007 Survey of Household Finances, Central Bank of Chile.

Figure 2

Informality Rates by gender, income group, educational level and age group



Source: 2007 Survey of Household Finances, Central Bank of Chile.

Educational level shows also a negative relationship with the rate of informality, as can be seen in figure 2. Workers with primary or lower education levels have informality rates of 56.9%, 55.3% and 43.7% for the ILOD, PD and COMBD definitions, respectively. In addition, 34.3% (ILOD), 33.3% (PD) and 20.8%



(COMBD) of workers with secondary education are informal. Further, while workers with a tertiary or undergraduate university degree have informality rates of 30.1% (ILOD), 6.7% (PD), and 3.9% (COMBD), workers with a postgraduate university degree have rates of 22.5% (ILOD), 3% (PD), and 1.2% (COMBD). Note also that the reductions in the informality rate between lower education levels and higher education levels for the PD and COMBD definitions are very marked. This result is partially explained by the specific construction of the PD definition, where self-employed workers with lower education levels are assumed to be informal.

From figure 2, it is evident that informality is not distributed evenly by age groups. Clearly, informality affects more workers that are older than 65 years, which have rates of 71.4% (ILOD), 48.6% (PD) and 42.8% (COMBD). The second most affected group is old middle-aged workers between 40 and 65 years, with informality rates of 37.3% (ILOD), 33.4% (PD) and 22.8% (COMBD). In the third place, we have young workers aged between 15 to 25 that have informality rates of 39.3% (ILOD), 18.6% (PD) and 14% (COMBD). Finally, we observe that the age group least affected by informality is the young middle-aged between 26 and 39 years, which have informality rates of 29%, 21.4% and 13.1% for the ILOD, PD and COMB definitions, respectively. Note, however, that if we consider only the ILOD definition of informality, the second most affected group is that of young workers and in third place, old middle-aged workers.

2. Informality and access to financial services

In this subsection we compare informal and formal households in terms of access to financial services, using data for 2007.⁶ Note that we define informal (formal) households as those where all the members that are occupied are informal (formal) workers. This implies that we discard households that have no member working or where some members are informal and some are formal workers. Again, we use the three different informality definitions. For the ILOD definition, we end up having 684 informal households and 1,891 formal households. Regarding the PD definition, we have 431 informal households and 2,346 formal households. Finally, for the COMBD definition we have 284 informal households and 1,620 formal households.

We constructed four variables that proxy access to financial services, namely:

- 1) Possession of bank account
- 2) Possession of credit card
- 3) Possession of debit card
- 4) Facing credit constraints

The variable “possession of bank account” indicates households where the head reports being the holder of a bank account. In addition, the variables “possession of credit card” and “possession of debit card” indicate households where there is

⁶ The results for the 2008, 2009 and 2010 surveys do not change our main conclusions. The results, which are not presented due to space consideration, are available upon request from the authors.

at least one member that uses credit card and debit card, respectively. Further, households that face credit constraints are those that have applied to a credit in the last two years and have suffered at least one rejection. We also consider households to be credit constrained if they have been granted a credit but they have declined it because they consider the conditions to be unfavorable. Further, we consider households to be credit constrained if they have been granted a credit but the amount granted is less than what was applied for. Finally, we also consider households to be credit constrained if they do not apply for a credit because they believe that they will not be granted one or believe they will be unable to afford to repay the loan.⁷

From table 2, it is clear that there is a much higher proportion of heads of formal households that possess a bank account than heads of informal households. For the ILOD definition, while the 25.1% of heads of formal households report having a bank account, only 9.9% of heads of informal households possess a bank account. This comparison is 26.6% against 5.9% for the PD definition and 27% against 4.6% for the COMBD definition.

Regarding the use of credit cards, we see in table 2 that a higher proportion of formal households (between 17.4% and 18.6%) use credit cards than informal households (between 11.9% and 13.2%). Again, this pattern is consistent for the three definitions of informality.

Table 2

Access to financial services

(percent)

Informality definition	Household heads with bank account		Households using credit cards	
	Formal	Informal	Formal	Informal
ILOD	25.1	9.9	17.4	13.2
PD	26.6	5.9	18.6	11.9
COMBD	27.0	4.6	17.4	13.2
Informality definition	Households using debit cards		Households with credit constraints	
	Formal	Informal	Formal	Informal
ILOD	12.9	3.9	33.7	46.8
PD	13.7	1.2	36.2	44.9
COMBD	13.9	0.0	34.8	49.1

Source: 2007 Survey of Household Finances, Central Bank of Chile.

Notes: ILOD: ILO definition; PD: productive definition; COMBD: combined definition.

⁷ All these questions are available in the Survey of Household Finances conducted by the Central Bank of Chile in 2007, which allow us to construct the single variable "Face credit constrains".



With respect to the use of debit cards, from table 2 we find a similar pattern to that for the use of credit cards, ranging between 12.9% and 13.9% the proportion of formal households that use debit cards and between 0% and 3.9% of informal households. Note that the use of debit cards is much less extended than the use of credit cards. A possible explanation is that debit cards are associated with the possession of a bank account; instead, credit cards are increasingly being issued by department and retail stores without the need to own a bank account.

In table 2, we see that there is a higher proportion of informal households that suffer credit constraints in comparison with formal households. While between 44.9% and 49.1% of informal households suffer credit constraints, between 33.7% and 36.2% of formal household are credit constraints. Clearly, this pattern is consistent across all three definitions of informality. Moreover, for all the four variables that capture access to financial services, we see that informal households have less access than formal households.

3. Informality and households' assets and liabilities

The database that we use allows us to analyze the proportion of households that have assets and liabilities and compare formal and informal households. In terms of assets, we distinguish between real assets—such as motor vehicles, primary residence and other real estates—and financial assets, such as fixed-income assets (saving accounts, retirement saving plan and term deposits) and variable-income assets (shareholding, mutual and investment funds and business partnership). Further, we also have data on households' liabilities, such as bank credit card debt, personal loans by banks, other bank credit, mortgage loans, retail store credit card debt, personal loans by retail stores and other financial companies, credit by credit unions, car loans, student loans, loans by family and friends, pawnshop credit, bought on credit and other debts.

From table 3, it is clear that there is a higher proportion of formal households that possess assets in comparison to informal households for 2007 in Chile. Although there is not much difference between formal and informal households in terms of ownership of primary residence, it is clear that a higher proportion of formal households possess a motor vehicle and other real estate properties. Furthermore, the difference is blatant in terms of financial assets, especially for variable-income assets such as shareholding, mutual and investment funds and business partnership. Note also that beyond the difference between formal and informal households, possession of financial assets by households is quite low in general, especially if compared with real assets.

With respect to liabilities, it is also clear that a higher proportion of formal households are indebted in comparison to informal households. The difference is especially important for loans granted by banks, such as debt by bank-issued credit cards, personal loans by banks, other bank credit, and mortgage loans. However, when analyzing the loans granted by retail stores (credit card or personal loans) the difference between formal and informal households is less important. Another interesting pattern is that credit card penetration by retail stores is much larger than for bank-issued credit cards. Clearly, retail stores provide an important

source of financing for both formal and informal households. Finally, although a small proportion of households take loans from family and friends and from pawnshops, these are the only categories of liabilities where informal households have a higher percentage than formal households. These results may evidence that informal households have a greater tendency to search for credit through more informal channels.

Table 3

Proportion of formal and informal households with assets and liabilities
(percent)

Assets	ILO definition		Productive definition		Combined definition	
	Formal	Informal	Formal	Informal	Formal	Informal
Real assets	81.2	72.8	78.8	78.4	79.2	73.1
Motor vehicles	41.3	30.5	42.3	29.0	42.0	26.5
Primary residence	67.2	63.5	65.5	68.4	65.8	66.4
Other real states	12.2	8.1	12.4	9.0	11.6	6.1
Financial assets	14.9	7.4	15.8	6.6	15.4	5.9
Fixed-income assets	11.1	6.5	11.7	5.7	11.6	5.9
Savings account	10.1	6.5	10.5	5.5	10.2	5.8
Retirement savings plan	3.2	0.4	3.0	0.9	3.3	0.5
Term deposit	1.6	0.4	1.6	0.3	1.8	4.0
Variable-income assets	5.7	1.2	6.1	0.9	5.8	0.0
Shareholding	2.8	0.5	3.0	0.1	2.8	0.0
Mutual and investment funds	3.5	1.1	4.4	0.4	4.2	0.3
Business partnership	1.3	0.4	1.1	0.7	1.1	0.0
Liabilities	Formal	Informal	Formal	Informal	Formal	Informal
Indebted	65.8	51.0	64.5	50.8	67.1	47.4
Bank credit card debt	15.1	6.5	16.2	3.6	16.9	3.9
Personal loans by banks	15.5	9.2	15.3	8.4	15.7	7.2
Other bank credit	8.3	5.5	9.2	3.5	9.0	3.9
Mortgage loans	16.7	6.0	16.8	5.0	17.7	3.9
Retail store credit card debt	56.1	43.6	55.8	43.8	58.4	42.8
Personal loans by retail stores and other financial companies	6.0	4.8	5.7	5.2	5.9	5.0
Credit by credit unions	8.1	4.0	8.2	4.7	9.3	5.6
Car loans	1.9	1.6	2.2	1.6	1.9	1.4
Student loans	4.9	3.1	5.8	1.1	5.4	0.8
Loans by family and friends	1.0	2.9	1.6	1.5	1.3	2.2
Pawnshop credit	0.0	1.1	0.2	0.5	0.1	0.8
Bought on credit	1.4	1.1	1.5	0.7	1.3	60.0
Other debts	3.0	2.4	2.7	2.9	2.6	2.2

Source: 2007 Survey of Household Finances, Central Bank of Chile.



4. Informality and saving behavior

In this subsection we analyze the saving rates of households, discriminating by different percentiles of income and between formal and informal households for 2007. Further, we present data on the percentage of households that inform being able to save or not in the last year, discriminating by formal and informal households. Finally, we present data on households' reported reasons for saving in 2007. The first step is to conceptualize saving since there is a great diversity of definitions, with some preponderance for the standard notion of total family income, net of retirement contributions and health insurance, minus total household expenditures, including durable goods, health expenses and educational expenses. These last three items are suggested as an unconventional form of saving by, among others, Attanasio (1998) and Butelmann and Gallego (2000). Furthermore, Deaton (2005) points out, as a criticism, that usually household surveys collect revenues of young people without their contributions to the pension system and consider the income of retirees as income and not as dissaving. These two effects may contribute to underestimating saving by young people and dissaving by retirees.

Following the above mentioned literature and benefiting from the richness of the EFH, we use the following three definitions for saving:

Definition 1 (*SR1*): Saving is the difference between total household income, net of retirement contributions and health insurance, and consumption expenditures. The saving rate is given by saving divided by total household income. Total household income includes imputed rent of own property or leased property for free. Consumption expenditures include all surveyed expenses.

Definition 2 (*SR2*): Like definition 1 (*SR1*), but excluding pension incomes from total household income and considering spending on education and health as saving, i.e. consumption expenditures exclude education and health spending.

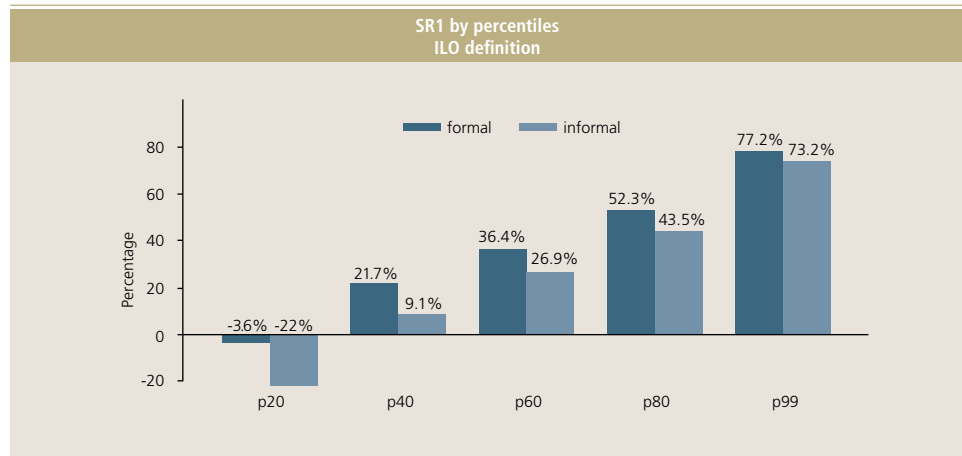
Definition 3 (*SR3*): A binary variable that assumes value 1 if the head of the household answers in the survey that the incomes of the household have been greater than the expenses in the last 12 months (i.e. they have been able to save), and 0 otherwise.

In figure 10 we present the saving rates by income percentile using definition *SR1*, discriminating between formal and informal households according to the ILOD definition for 2007. Two main conclusions are evident. Firstly, for all the different income percentiles, formal households have higher saving rates than informal households. Secondly, saving rates are increasing in the level of income for both formal and informal households, i.e. households with higher

incomes have greater saving rates. These two results are confirmed in figures 11 for the *SR2* definition of the saving rate.^{8,9}

Figure 3

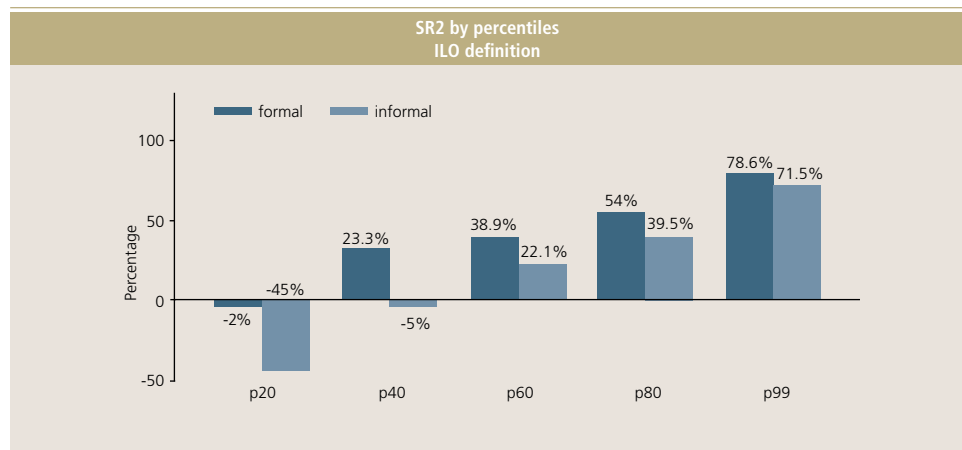
Saving rates *SR1* by income percentiles



Source: 2007 Survey of Household Finances, Central Bank of Chile.

Figure 4

Saving rates *SR2* by income percentiles



Source: 2007 Survey of Household Finances, Central Bank of Chile.

8 We reach the same conclusions using the PD and COMBD definitions of informality, which are not presented due to space considerations, but are available upon request from the authors.

9 We reach the same conclusions using two alternative saving definitions to definition 1 (*SR1*), namely as definition 1 (*SR1*), but excluding pension incomes from total household income, and as definition 1 (*SR1*), but considering spending on education and health as saving, i.e. consumption expenditures exclude education and health spending. These results are not presented due to space considerations, but are available upon request from the authors.



Table 4 presents the percentage of households that inform having or not having saved in the last year, discriminating between formal and informal households according to the different informality definitions. Note that these are the results for the *SR3* definition of saving. Similarly to the results for the other definitions of saving, we find that the percentage of formal households that save is greater than the percentage of informal households that save for all three definitions of informality. Equivalently, we find that the percentage of formal households that do not save is lower than the percentage of informal households that have not saved in the last year.

Table 4

Share of households that inform having or not having saved in the last year

(percent)

	ILOD		PD		COMBD	
	Formal	Informal	Formal	Informal	Formal	Informal
Did not save	64.2	77.7	66.3	77.0	64.8	79.7
Saved	35.8	22.3	33.7	23.0	35.2	20.3

Source: 2007 Survey of Household Finances, Central Bank of Chile.

Notes: ILOD: ILO definition; PD: productive definition; COMBD: combined definition.

In table 5, we present the reported motives for saving, distinguishing between formal and informal households according to the three different informality definitions used in this paper. Note that the data is presented as a percentage of households, including those households that do not report saving for any reason, and that households could respond one or more motive, i.e. the different motives are not mutually exclusive. The primary reason for saving for both formal and informal households is the precautionary motive. Note, however, that a larger proportion of formal households report this reason when compared with informal households. This result can be somehow counter-intuitive if we assume that informal households have less stability in their employments and a larger variation in their income streams, and would, thus, have more incentives for precautionary saving. Another interesting result is that a larger proportion of informal households than formal households state that they save for retirement. This is an intuitive result if we consider that the informality definitions imply that informal workers are less covered by contributory retirement plans than formal workers. Note also that there is no other motive for saving for which informal households report a higher percentage than formal households (an exception with mixed results is the health and education reason). Further, it is interesting to note that a larger proportion of formal households than informal households report saving in order to reduce debt. This may be an intuitive result if we consider that, in general and from table 4, a smaller proportion of informal households are indebted.

Table 5

Households' reported motive for saving

(as percentage of households)

Reported reason	ILOD		PD		COMBD	
	Formal	Informal	Formal	Informal	Formal	Informal
Precautionary	24.1	17.9	22.2	20.3	23.4	18.0
Retirement	7.7	10.4	8.2	10.8	7.7	12.7
Durable goods	9.5	8.1	9.3	9.4	9.1	8.2
Health and education	8.8	7.4	8.3	11.3	7.9	9.7
Inheritance	0.7	1.1	0.9	0.5	0.8	0.7
Save	9.4	6.8	9.2	6.7	9.4	6.3
Reduce debt	6.3	3.4	5.6	4.3	5.8	3.3
Other	8.3	5.2	8.7	5.6	8.6	6.3

Source: 2007 Survey of Household Finances, Central Bank of Chile.

Notes: ILOD: ILO definition; PD: productive definition; COMBD: combined definition.

III. DATA, METHODOLOGY AND RESULTS**1. Data and variables**

As was discussed in the last section, we use data from the Survey of Household Finances conducted by the Central Bank of Chile. In this section, we present the results for 2007, which covered 3,828 urban households from the whole country. However, we have considered only those households with at least one member employed and have included only pure households where all members are either formal or informal, i.e. we exclude households that have some members that are formal and others that are informal. Thus, our sample is reduced to between 2,533 and 1,740 urban households, depending on the saving rate variable used (*SR1*, *SR2* and *SR3*) and informality definition (ILOD, PD and COMBD).

The variables that are used in the regression analysis are the saving rates *SR1*, *SR2*, and *SR3* that were defined in the last section. In terms of the informality dummy variables, we have *infILOD*, *infPD* and *infCOMBD* that are constructed using the ILOD, PD and COMBD definitions, respectively. In addition, we have the *dprecsav* variable that is a dummy variable indicating a household that declares having saved during the last year due to precautionary motives. *dretsav* is a dummy variable capturing those households that have saved during the last year for retirement motives. The variable *ddurgoodsav* captures households that have saved in order to buy durable goods and *ddebtreducsav* capture those households that have saved to repay loans. In addition, for each of these dummy variables, we create new interaction variables with the informality variable, namely *infprec*, *Infret*, *infdurgood* and *infdebt*. For example, *infprec* captures the effect on the saving rate of those households that save for precautionary reasons and are informal.



On top of these variables, we constructed several variables that have been used in the surveyed literature in section I and whose descriptive statistics are presented in table 6.¹⁰ Note that the descriptive statistics of these variables correspond to the ILOD definition of informality. The average age of household heads (*age*) was 48 years and the average number of household members with revenues (*minc*) was 1.67 persons. The *mincsq* variable indicates the square of the number of people in the household with income. The *empspo* variable indicates the percentage of households in which the household head has a spouse or live-in partner who works. We find that 22% of households with at least one worker are in this situation. To control for dependency rates, two variables were constructed: i) *mchild* to define the number of members under 18 years of age for each household and ii) *melderly* to identify the number of family members over 65. The averages were 0.83 and 0.23 per household, respectively. With regard to the gender of the household head, 65% are males. For the educational level of the household head, four indicator variables were developed: i) *primary* to define primary education level, ii) *secondary* to define secondary education level, iii) *tertoruniv* to define college or tertiary education level, and iv) *postgraduate* to define postgraduate education level, with the following prevalence rates: 27%, 43%, 19% and 11%, respectively. Further, we constructed a categorical variable indicating households for which the household head is retired (*hhretired*), obtaining that 11% of households have a pensioner as household head.

Finally, the variable indicating that a household head owns a bank account (*bankacc*) shows an average of 18%. This variable is important because the Modigliani conceptual framework assumes that financial markets operate properly. However, in reality that is not the case and, thus, it is important to control for households that do not have proper access to credit markets and may have difficulties in smoothing their consumption over their lifetime. A binary variable was constructed to indicate whether the household owns at least one durable good, i.e. a car, a house or other property (*durgood*). This variable shows that 79% of households have at least one durable good. In addition, we have the *thincome* variable that measures the household's total monthly income expressed in millions of constant Chilean pesos of 2007 and the *thincomesq* variable that is the square of household's total monthly income, which is supposed to capture non-linearities in the relationship with the saving rate.¹¹

¹⁰ In the appendix, we include a table showing all the variable names in English as translated from their original names in Spanish as they appear in the EFH survey.

¹¹ In the regression analysis, we obtained similar results when using, instead of the level of income, an indicator variable that captures the relative position of households in the income distribution. The reason for this robustness test is that, as Sandoval-Hernandez (2010) points out, household income levels and educational levels are strongly correlated, which may imply the risk of facing multicollinearity. The results, which are not presented due to space considerations, are available upon request from the authors.

Table 6

Descriptive statistics of variables used in regression analysis

Descriptive Statistics (ILO definition)
Only formal and informal households

Variables	Labels	Mean	St. dev.	Median	10th percentile	90th percentile
<i>SR1</i>	Savings rate (definition 1)	0.11	1.04	0.26	-0.39	0.60
<i>SR2</i>	Savings rate (definition 2)	0.06	1.19	0.25	-0.57	0.61
<i>SR3</i>	Dummy variable (0 not save, 1 save)-(definition 3)	0.20	0.40	0.00	0.00	1.00
<i>inflOD</i>	Dummy variable (0 Formal household; 1 Informal household)	0.34	0.47	0.00	0.00	1.00
<i>age</i>	Age of household head	48	13	47	30	66
<i>agesq</i>	Age squared of the household head	2,443	1,343	2,209	900	4,356
<i>thincome</i>	Total household income (in millions of pesos)	0.791	3.226	0.420	0.181	1.260
<i>thincomesq</i>	Squared total household income (in millions of pesos)	11.03	348.80	0.18	0.03	1.59
<i>thincome2</i>	Total household income without pensions (in millions of pesos)	0.733	3.205	0.380	0.155	1.153
<i>thincomesq2</i>	Squared total household income w/o pensions (in millions of pesos)	10.81	342.16	0.14	0.02	1.33
<i>minc</i>	Number of household members with income	1.67	0.79	2.00	1.00	3.00
<i>mincsq</i>	Square of number of household members with income	3.40	3.46	4.00	1.00	9.00
<i>empspo</i>	Spouse or live-in partner who is employed	0.22	0.41	0.00	0.00	1.00
<i>mchild</i>	Number of children at home	0.83	0.96	1.00	0.00	2.00
<i>melderly</i>	Number of elderly at home	0.23	0.54	0.00	0.00	1.00
<i>primary</i>	Head of household with primary education	0.27	0.44	0.00	0.00	1.00
<i>secondary</i>	Head of household with secondary education	0.43	0.49	0.00	0.00	1.00
<i>tertoruniv</i>	Head of household with tertiary or university education	0.19	0.39	0.00	0.00	1.00
<i>postgraduate</i>	Head of household with postgraduate education	0.11	0.32	0.00	0.00	1.00
<i>bankacc</i>	Head of household who has current account	0.18	0.39	0.00	0.00	1.00
<i>gender</i>	Gender of household head	0.65	0.48	1.00	0.00	1.00
<i>durgood</i>	Home with durable goods	0.79	0.40	1.00	0.00	1.00
<i>hhretired</i>	Head of household is retired	0.11	0.32	0.00	0.00	1.00
<i>dprec sav</i>	Dummy precautionary saving	0.22	0.41	0.00	0.00	1.00
<i>dretsav</i>	Dummy retirement saving	0.09	0.28	0.00	0.00	0.00
<i>ddurgoodsav</i>	Dummy saving durable goods	0.09	0.29	0.00	0.00	0.00
<i>ddebtreduc sav</i>	Dummy savings debt reduction	0.05	0.22	0.00	0.00	0.00
<i>infret</i>	Interaction dummy retirement saving and informality	0.04	0.19	0.00	0.00	0.00
<i>infprec</i>	Interaction dummy precautionary saving and informality	0.06	0.24	0.00	0.00	0.00
<i>inf durgood</i>	Interaction dummy durable goods saving and informality	0.03	0.16	0.00	0.00	0.00
<i>infdebt</i>	Interaction dummy saving debt reduction and informality	0.01	0.11	0.00	0.00	0.00

Source: 2007 Survey of Household Finances, Central Bank of Chile.

2. Empirical methodology

The objective of this section is to determine whether households of informal workers have a different saving behavior relative to households with formal workers. Additionally, the relationship between saving and the main determinants commonly cited in the literature is also analyzed. The empirical strategy follows a cross-section regression analysis by using both ordinary least squares and probit estimation models. It is important to clarify that the estimates are made on a population of households containing only formal workers or only informal workers according to each of the three possible informality definitions (ILOD, PD and COMBD). In other words, in addition to discarding those households that have no member that works, we also remove those households that have some members that are formal and others that are informal. In this way, we are comparing the saving behavior of purely formal households with purely informal households.

Ordinary least squares regression analysis

We perform the regression analysis for each of the two saving rate definitions as the dependent variable (*SR1* and *SR2*). Further, for each saving rate definition, we test the three different informality dummy variables, namely the ILOD, PD and COMBD informality definitions. Further, following Butelmann and Gallego (2000), we remove the extreme percentiles 1st and 99th for each alternative definition of saving rates in order to ensure an adequate empirical treatment since household surveys typically have a high dispersion of data by the presence of outliers, which tend to strongly bias the estimation results.

Concretely, we estimate the following equation by ordinary least squares, weighting with the expansion factors given to each household in the EFH and with robust standard errors for heteroscedasticity:¹²

$$SR_i = \alpha_0 + \alpha_1 inf_i + \alpha_2 age_i + \alpha_3 agesq_i + \alpha_4 thincome_i + \alpha_5 thincomesq_i + \alpha_6 minc_i + \alpha_7 mincsq_i + \alpha_8 empspo_i + \alpha_9 mchild_i + \alpha_{10} melderly_i + \alpha_{11} secondary_i + \alpha_{12} tertoruniv_i + \alpha_{13} postgraduate_i + \alpha_{16} bankacc_i + \alpha_{17} gender_i + \alpha_{18} durgood_i + \alpha_{19} hhretired_i + \alpha_{20} Dprecsav_i + \alpha_{21} Dretsav_i + \alpha_{22} Ddurgoodsav_i + \alpha_{23} Ddebtreducsav_i + \alpha_{24} infprec_i + \alpha_{25} infret_i + \alpha_{26} infdurgood_i + \alpha_{27} infdebt_i + \epsilon_i$$

where the i subscript represents household i , SR_i is the saving rate variable, inf_i is the informality dummy variable, and the other variables are the ones described in subsection III.1. Note that this model is estimated using two different dependent variables, namely the *SR1* and the *SR2* definitions of saving rate. Further, it is estimated using the three different independent variables regarding the informality variable, namely the ILOD, the PD and the COMBD definitions of informality. Thus, the model is estimated in total for six different combinations of the saving rate variable and the informality dummy variable,

¹² See Madeira (2011) for a discussion of the computation of population weights for the EFH survey and Solon et al. (2013) for the use of population weights in studies about causal effects.

namely: $(SR1;infILOD)$, $(SR1;infPD)$, $(SR1;infCOMBD)$, $(SR2;infILOD)$, $(SR2;infPD)$, and $(SR2;infCOMBD)$.

Probit regression analysis

In the case of the probit regression we use the saving definition 3 ($SR3$) as the dependent variable. In addition, we use the three different informality dummy variables, namely the ILOD, PD and COMBD informality definitions. Note that we are not removing the extreme percentiles as in the OLS regression because the $SR3$ is a dummy variable without the variability of the $SR1$ and $SR2$ variables.

We estimate the following model using a probit regression methodology, correcting the standard errors for heteroskedasticity, but without weighting with the expansion factors given to each household in the EFH as we did in the OLS regression:

$$P[SR3_i=1] = \Phi [\beta_0 + \beta_1 inf_i + \beta_2 age_i + \beta_3 agesq_i + \beta_4 thincome_i + \beta_5 thincomesq_i + \beta_6 minc_i + \beta_7 mincsq_i + \beta_8 empspo_i + \beta_9 mchild_i + \beta_{10} melderly_i + \beta_{11} secondary_i + \beta_{12} tertoruniv_i + \beta_{13} postgraduate_i + \beta_{14} bankacc_i + \beta_{15} gender_i + \beta_{16} durgood_i + \beta_{17} hhretired_i]$$

Note that, given that there is a non-linear relationship between the dependent variable and its determinants, Φ is the standard normal cumulative distribution function corresponding to the probit methodology. Also, estimating this equation means that we are estimating the probability that the dependent variable $SR3$ is equal to 1, i.e. that a household saved in the last year, conditional on the other independent variables. Further, we have not included the dummy variables $dprecisavi$, $dretsavi$, $ddurgoodsavi$, $ddebtreducsavi$, $infpreci$, $infreti$, $infdurgoodi$, and $infdebt$ because these variables and the dependent variable $SR5$ are constructed from the same survey questions.

3. Results

OLS regression results

Tables 7, 8 and 9 present the estimation results for each of the two definitions of saving rates and for each of the three definitions of informality for the year 2007.^{13,14} It should be noted that the size of the linear coefficients of determination (R^2) varies between 0.245 and 0.284. This means that between 28.4% and 24.5% of the saving rates' variability is explained by the dependent

¹³ We reached similar conclusions when using two alternative saving definitions to definition 1 ($SR1$), namely as definition 1 ($SR1$), but excluding pension incomes from total household income, and as definition 1 ($SR1$), but considering spending on education and health as saving, i.e. consumption expenditures excludes education and health spending. The results, which are not presented due to space considerations, are available upon request from the authors.

¹⁴ We reached similar conclusions when using the 2008, 2009 and 2010 surveys. The results, which are not presented due to space considerations, are available upon request from the authors.



variables included in the regressions, so the specification of the equations appears to be correct in terms of the results.

Table 7 presents the estimation results for each of the two definitions of saving with the ILOD definition of informality. In the model using the saving rate definition 1 (*SR1*), the dummy variable indicating households that have informal workers (*infILOD*) yields a coefficient of -0.124 , which turns out to be significant at the 1% level. Further, for the *SR2* definition, the informal household coefficient (*infILOD*) is -0.201 and is significant at the 1% level. These results imply that informal households according to the ILOD definition save between 12.4 and 20.1 percentage points less than formal households.

In table 8, where the PD definition of informality is used for the two saving regressions, we find similar results. We get coefficient values of -0.167 and -0.257 for the *SR1* and *SR2* definitions of saving, respectively, all with a statistical significance level of 1%. Accordingly, for the PD definition of informality, informal households save between 16.7 and 25.7 percentage points less than formal households. Finally, in table 9 for the COMBD informality definition, the informality dummy coefficient has the values -0.228 and -0.332 for the saving definitions *SR1* and *SR2*, respectively, and in all four cases with significance levels of 1%. These results imply that for the COMBD definition of informality, informal households save between 22.8 and 33.2 percentage points less than formal households. Concluding, for all two definitions of saving rate and all three informality definitions, informal households have a lower saving rate than formal households. Note also that the negative relationship found between households' saving rates and informal workers may also be indicating, particularly in the ILOD definition of informality, that there might not be any trade-off between pension saving and voluntary household saving. This result is in contrast to the one obtained by Butelmann and Gallego (2000).

Regarding the dummy variables for the stated motives of saving and the corresponding interaction variables, we find that the variable *infprec* is positive and significant, meaning that informal households that indicate that they save for precautionary motives save between 13.6 and 32.4 percentage points more than other households. When comparing the size of these coefficients with the size of the coefficients of the informality variables *infILOD*, *infPD* and *infCOMBD*, we reach the conclusion that informal households that save for precautionary motives have saving rates similar to formal households. Note also that the *dprecsav* is insignificant in all specifications, meaning that formal households that state that they save for precautionary motives do not have significantly different saving rates than the rest of formal households. In addition, there is some evidence that the *infdurgood* variable is positive and significant, meaning that informal households that save for buying a durable good have between 19.8 and 23.6 percentage points higher saving rates than the rest of informal households. Further, there is also some evidence that informal households that save for retirement (*infret*) have between 16.3 and 20 percentage points higher saving rates than the other informal households. Note that both the *ddurgoodsav* and *dretsav* variables are insignificant, which implies that formal households that save for either of these motives do not have

significantly different saving rates than the other formal households. Finally, the other indicator variables capturing the motive of reducing debt (*ddebtreducsav* and *infdebtred*) are insignificant.

With regard to the variables that relate to the life cycle hypothesis, age and age squared of the household head, for most of the different specifications we get negative —and significant— values for the age variable, and positive and significant values for the age squared variable. Note that these results imply that the age profile of saving rates presents a U shape, instead of the expected inverted-U shape. This result is in line with the findings of Butelmann and Gallego (2000) for the Chilean economy, as well as of Sandoval-Hernandez (2010) for the Mexican economy. One possible explanation for this finding focuses on different saving preference structures between generations.

For all the regression equations, we find a positive relationship between the household's income variable *thincome* and the rates of household saving, i.e. households with higher incomes save more. These results are in line with Harris et al. (2002), Sandoval-Hernandez (2010) and Beckman et al. (2013), among others. Further, we included the square of the income variable (*thincomesq*), which is negative and significant, meaning that there is an inverted-U relationship between income and the saving rate. In addition, the number of household members with incomes (*minc*) has a significant and positive relationship in most of the estimates, also in line with the results commonly seen in the literature, such as in Sandoval-Hernandez (2010). Further, the square of the number of members with income (*mincsq*) is mostly negative and significant, meaning that there is an inverted-U relationship between *minc* and the saving rates. Further, households headed by men seem to save between 14.4 and 8.96 percentage points more than those in which the head is a woman, which is in line with Attanasio (1998) and Sandoval-Hernandez (2010). To control for household dependency ratio, two variables are introduced: *mchild* and *melderly*. *mchild* shows a negative and statistically significant relationship in all econometric specifications as in Xiao (1996), Harris et al. (2002), and Sandoval-Hernandez (2010). In terms of the *melderly* variable, a negative relationship is found, but it is only statistically significant for the saving rate definition 2 (*SR2*).

In most studies, the educational level has shown to be an important determinant of household saving, with a positive sign. In our study, education is insignificant for most specifications. This result, however, should be taken with caution because the household income level and the educational level are strongly correlated, which may imply the risk of facing multicollinearity (Sandoval-Hernandez, 2010). In order to capture the effect of households not restricted to credit, a dummy variable, *bankacc*, indicating household heads that have a bank account, is defined. In most specifications, this variable is insignificant. For the variable capturing households that own durable goods, we find highly significant results in all three specifications with the same positive sign as most studies (Attanasio, 1998; Butelmann and Gallego, 2000; Harris et al., 2002; Sandoval-Hernandez, 2010). This result may be explained by taking into account that most durable goods are bought on credit, which implies that households need to save more in order to repay the loan.

The variable *hhretired* shows a statistically significant relationship in almost all cases, but the sign of the coefficient varies depending on the definition of saving that is used. While for the saving definition 1 (*SR1*), the sign is positive, for the saving definition 2 (*SR2*), it is negative. This apparently contradictory result may be justified by taking into account that saving definition 2 (*SR2*) considers retirement income as dissaving.

Table 7
Saving rates regressions with ILOD definition of informality

Dependent variable: saving rate	(1)		(2)	
<i>inflOD</i>	-0.124***	(-3.25)	-0.201***	(-4.74)
<i>age</i>	-0.0101**	(-2.12)	-0.0156***	(-2.64)
<i>agesq</i>	0,000	(1.53)	0.000133**	(2.19)
<i>thincome</i>	0.206***	(10.58)		
<i>thincomesq</i>	-0.00719***	(-9.35)		
<i>minc</i>	0.317***	(4.59)	0.167**	(2.04)
<i>mincsq</i>	-0.0470***	(-3.55)	-0.0249	(-1.41)
<i>empspo</i>	-0,021	(-0.58)	0.0323	(0.74)
<i>mchild</i>	-0.0833***	(-4.40)	-0.056***	(-2.63)
<i>melderly</i>	-0,032	(-1.43)	-0.0976***	(-3.18)
<i>secondary</i>	-0,040	(-1.11)	-0.00283	(-0.07)
<i>tertoruniv</i>	-0,017	(-0.41)	-0.0216	(-0.39)
<i>postgraduate</i>	-0,007	(-0.15)	-0.0541	(-0.90)
<i>bankacc</i>	-0.0964***	(-2.60)	-0.0562	(-1.50)
<i>gender</i>	0.0998***	(3.11)	0.144***	(3.92)
<i>durgood</i>	0.141***	(4.01)	0.249***	(5.07)
<i>hhretired</i>	0.169***	(4.95)	-0.0849	(-1.48)
<i>dretsav</i>	-0,064	(-1.47)	0.00322	(0.08)
<i>infret</i>	0,101	(1.15)	-0.0402	(-0.44)
<i>dprecav</i>	0,007	(0.26)	-0.0506	(-1.09)
<i>infprec</i>	0.136**	(2.57)	0.193**	(2.56)
<i>ddurgoodsav</i>	-0,093	(-1.41)	-0.0738	(-1.05)
<i>infdurgood</i>	0.202**	(2.18)	0.198**	(1.98)
<i>ddebtreducsav</i>	0,032	(0.72)	0.0264	(0.64)
<i>infdebttred</i>	-0,142	(-1.40)	-0.0918	(-0.94)
<i>thincome2</i>			0.271***	(9.92)
<i>thincome2sq</i>			-0.0095***	(-8.96)
<i>constant</i>	-0,067	(-0.51)	0.0626	-0.41
No. observations	2355		2346	
Adjusted R-squared	0,245		0,251	
F	15,87		17,14	

Source: 2007 Survey of Household Finances, Central Bank of Chile.

(1) Saving rate 1 definition. (2) Saving rate 2 definition.

Econometric methodology: Ordinary least squares with robust standard errors.

t statistics in parentheses.

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table 8

Saving rates regressions with PD definition of informality

Dependent variable: saving rate	(1)		(2)	
<i>infPD</i>	-0.167***	(-3.58)	-0.257***	(-4.93)
<i>age</i>	-0.0127***	(-2.65)	-0.0145***	(-2.59)
<i>agesq</i>	0.0000952*	(1.95)	0.000125**	(2.12)
<i>thincome</i>	0.165***	(8.32)		
<i>thincomesq</i>	-0.00453***	(-5.73)		
<i>minc</i>	0.338***	(5.49)	0.148**	(2.06)
<i>mincsq</i>	-0.0551***	(-4.47)	-0.0215	(-1.43)
<i>empspo</i>	-0.034	(-1.06)	0.0303	(0.8)
<i>mchild</i>	-0.072***	(-4.31)	-0.0513***	(-2.84)
<i>melderly</i>	-0.0201	(-0.90)	-0.0971***	(-3.24)
<i>secondary</i>	-0.057*	(-1.69)	-0.0194	(-0.50)
<i>tertoruniv</i>	-0.0684*	(-1.74)	-0.033	(-0.75)
<i>postgraduate</i>	-0.0306	(-0.75)	-0.026	(-0.55)
<i>bankacc</i>	-0.0552	(-1.55)	-0.0239	(-0.68)
<i>gender</i>	0.0993***	(3.34)	0.135***	(4.34)
<i>durgood</i>	0.115***	(3.56)	0.216***	(5.34)
<i>hhretired</i>	0.160***	(4.68)	-0.0961*	(-1.75)
<i>dretsav</i>	-0.0551	(-1.02)	-0.0201	(-0.39)
<i>infret2</i>	0.163**	(2.06)	0.0338	(0.36)
<i>dprecav</i>	0.012	(0.48)	-0.015	(-0.52)
<i>infprec2</i>	0.106	(1.6)	0.198***	(2.8)
<i>ddurgoodsav</i>	0.00798	(0.24)	0.0174	(0.52)
<i>infdurgood2</i>	-0.108	(-0.62)	-0.0538	(-0.31)
<i>ddebtreducsav</i>	0.0377	(0.96)	0.0472	(1.54)
<i>infdebtreducsav</i>	-0.0401	(-0.33)	0.00266	(0.02)
<i>thincome2</i>			0.203***	(7.94)
<i>thincomesq2</i>			-0.00557***	(-5.47)
<i>_cons</i>	0.0421	(0.31)	0.113	(0.75)
No. observations	2533		2527	
Adjusted R-squared	0.25		0.254	
F	15.36		17.84	

Source: 2007 Survey of Household Finances, Central Bank of Chile.

(1) Saving rate 1 definition. (2) Saving rate 2 definition.

Econometric methodology: Ordinary least squares with robust standard errors.

t statistics in parentheses.

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.



Table 9

Saving rates regressions with COMBD definition of informality

Dependent variable: saving rate	(1)		(2)	
<i>infCOMBD</i>	-0.228***	(-3.70)	-0.332***	(-5.03)
<i>age</i>	-0.00662	(-1.09)	-0.0152**	(-2.09)
<i>agesq</i>	0.0000465	(0.77)	0.000143*	(1.87)
<i>thincome</i>	0.198***	(9.75)		
<i>thincomesq</i>	-0.00704***	(-8.86)		
<i>minc</i>	0.228***	(2.96)	0.0214	(0.23)
<i>mincsq</i>	-0.0319**	(-2.08)	0.00554	(0.26)
<i>empspo</i>	0.0393	(1.04)	0.123***	(2.84)
<i>mchild</i>	-0.0634***	(-4.35)	-0.0286*	(-1.67)
<i>melderly</i>	-0.00468	(-0.18)	-0.0755**	(-2.00)
<i>secondary</i>	-0.0732*	(-1.81)	-0.0533	(-1.17)
<i>tertoruniv</i>	-0.0527	(-1.11)	-0.0772	(-1.16)
<i>postgraduate</i>	-0.0444	(-0.91)	-0.0586	(-1.09)
<i>bankacc</i>	-0.0802*	(-1.79)	-0.0353	(-0.77)
<i>gender</i>	0.0896**	(2.57)	0.120***	(2.96)
<i>durgood</i>	0.166***	(4.15)	0.276***	(4.9)
<i>hhretired</i>	0.171***	(4.48)	-0.108	(-1.59)
<i>dretsav</i>	-0.0693	(-1.38)	0.0047	(0.11)
<i>infret3</i>	0.200**	(2.41)	0.00904	(0.09)
<i>dprec3</i>	0.00759	(0.24)	-0.0633	(-1.17)
<i>infprec3</i>	0.219***	(2.91)	0.324***	(3.31)
<i>ddurgoodsav</i>	-0.0439	(-0.94)	-0.0116	(-0.25)
<i>infdurgood3</i>	0.236**	(1.97)	0.219*	(1.7)
<i>ddebtreducsav</i>	0.0366	(0.71)	0.0345	(0.8)
<i>infdebtreducsav</i>	-0.0775	(-0.50)	-0.0493	(-0.36)
<i>thincome2</i>			0.239***	(9.64)
<i>thincomesq2</i>			-0.00858***	(-8.65)
<i>Constant</i>	-0.0766	(-0.47)	0.178	(0.93)
No. observations	1742		1740	
Adjusted R-squared	0.261		0.284	
F	15.6		17.22	

Source: 2007 Survey of Household Finances, Central Bank of Chile.

(1) Saving rate 1 definition. (2) Saving rate 2 definition.

Econometric methodology: Ordinary least squares with robust standard errors.

t statistics in parentheses.

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Probit regression results

The estimation results for the probit regression analysis for the year 2007 are presented in table 10. For the variables indicating households that belong to informal households, we find negative and significant results at the 1% level for the ILOD definition of informality (*infILOD*) and at the 10% level for the COMBD definition of informality (*infCOMBD*). In terms of the marginal effects for these estimations given the average value of the other independent variables, we obtain that households with informal workers have between 9.13% and 8.31% lower probability of saving than households with only formal workers.¹⁵ Using the PD definition, we do not find that the variable *infPD* is significant.¹⁶

Table 10

Dependent variable: saving rate (definition 3)	(1)		(2)		(3)	
<i>infILOD</i>	-0.254***	(0.0814)				
<i>age</i>	-0.0353**	(0.0173)	-0.0194	(0.0164)	-0.0252	(0.0209)
<i>agesq</i>	0.000224	(0.000184)	0.000111	(0.000175)	0.000119	(0.000224)
<i>thincome</i>	0.305***	(0.0502)	0.114***	(0.0348)	0.305***	(0.0578)
<i>thincomesq</i>	-0.0163***	(0.00298)	-0.00359**	(0.00165)	-0.0162***	(0.00332)
<i>minc</i>	0.0797	(0.176)	0.0504	(0.158)	0.238	(0.219)
<i>mincsq</i>	-0.0411	(0.0361)	-0.0333	(0.031)	-0.0754	(0.0471)
<i>empspo</i>	0.0553	(0.0892)	0.0911	(0.0834)	-0.00135	(0.107)
<i>mchild</i>	-0.235***	(0.0383)	-0.210***	(0.0358)	-0.257***	(0.0452)
<i>melderly</i>	-0.0072	(0.0842)	0.00933	(0.0787)	-0.00227	(0.104)
<i>secondary</i>	-0.137	(0.107)	-0.0625	(0.107)	-0.0572	(0.131)
<i>tertoruniv</i>	0.0198	(0.123)	0.139	(0.122)	0.0221	(0.154)
<i>postgraduate</i>	0.0473	(0.135)	0.317**	(0.131)	0.169	(0.163)
<i>bankacc</i>	0.245***	(0.0911)	0.310***	(0.0841)	0.262**	(0.105)
<i>gender</i>	-0.0757	(0.0713)	-0.0178	(0.0682)	-0.128	(0.0835)
<i>durgood</i>	0.514***	(0.0986)	0.522***	(0.0952)	0.475***	(0.114)
<i>hhretired</i>	0.0863	(0.136)	-0.00954	(0.128)	-0.0342	(0.168)
<i>infPD</i>			-0.163	(0.105)		
<i>infCOMBD</i>					-0.229*	(0.13)
Constant	0.382	(0.449)	-0.165	(0.423)	0.029	(0.549)
No. observations	1550		1672		1137	
Pseudo R-squared	0.136		0.117		0.149	

Source: 2007 Survey of Household Finances. Central Bank of Chile.
Econometric methodology: probit. Standard errors in parentheses.

¹⁵ The calculations of the marginal effects are not presented due to space considerations but are available upon request from the authors.

¹⁶ We also estimate the probit model using sample weights (the "pw" sample weight comando in Stata). The *infILOD* coefficient continues being significant but at the 5% level, the *infPD* coefficient continues being insignificant and the *infCOMBD* becomes insignificant. However, as Angrist and Pischke (2009) claim, it is not clear that using sample weights is a better estimation strategy. These results are not presented due to space considerations but are available upon request from the authors.



For the income variables *thincome* and *thincomesq*, we find similar results to the OLS results from last section. Household income and the probability of saving appear to have an inverted-U relationship, which means that the probability of saving is increasing to household income but at a decreasing rate. In the case of the age variables (*age* and *agesq*), we only find that it is significant at the 5% level for the ILOD definition but not for the other definitions, which means that we cannot establish a clear pattern for age and saving. For the household dependency ratio *mchild*, we find that households with children have a significantly lower probability of saving at the 1% level for all three informality definitions. In addition, we find a positive —and significant at the 1% level— relationship between the probability of saving and the variable *bankacc*, i.e. household heads that have a bank account. Note that this relationship was insignificant in the OLS estimations. The possession of durable goods (*durgood*) is positively and significantly related at the 1% level to the probability of saving. Note that we have not included the dummy variables that capture the motives for saving and their interaction with the informality variables because these variables and the dependent variable *SR3* are constructed from the same survey questions.

IV. CONCLUSIONS

The purpose of this paper is to compare the saving behavior of formal and informal workers. Further, we provide a socioeconomic and financial characterization for these workers. We use the Survey of Household Finances conducted by the Central Bank of Chile in 2007 and have between 2,533 and 1,740 observations of urban households from Chile, depending on the saving definition and the informality definition that is used. For the saving variable we have used three different definitions and for the informality variable we have also used three different definitions. In terms of the cross-section estimation methodology, we have used an OLS model and a probit model. Furthermore, we did robustness checks using data for 2008, 2009 and 2010.

The regression results tend to indicate that informal households save less than formal households. Although this conclusion is robust to most specifications, there are some results that indicate a non-significant difference. We also find evidence that those informal households that in the survey questionnaire declare saving for precautionary motives have higher saving rates than the other informal households. In addition, we find evidence that those informal households that declare saving in order to buy durable goods or for retirement also have higher saving rates. Further, the estimation results for the other determinants of saving are in line with the literature on saving, the most relevant determinant being the income level. In addition, the descriptive data indicate that informal workers seem to have less access to financial services and possess less financial assets and liabilities.

Putting forward some tentative explanations for the differential in the saving behavior between informal and formal households, one possible explanation is that informal households are less risk averse than formal households. Thus,

they prefer to consume a higher proportion of their income and have lower saving rates than formal households. It might be that informal households are more used to having short-term temporary jobs and a more unstable income stream than formal households. Thus, this experience of living in a more volatile environment may end up influencing their risk aversion preferences. Note that our results do not support the hypothesis that informal households have higher saving rates than formal households due to precautionary saving given the higher variability of the income stream of informal households. We find, however, support that the saving rates of informal households that state that they save for precautionary motives is higher than in the rest of informal households.

Regarding policy implications, it seems that combating informality may have positive consequences on the aggregate saving rate. However, we should be careful with this tentative conclusion as more research is needed, especially in terms of understanding why informal households save less than formal households.



REFERENCES

- Angrist, J.D. and J.S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Attanasio, O. (1998). "A Cohort Analysis of Saving Behavior by U.S. Households." *Journal of Human Resources* 33(3): 575–609.
- Attanasio, O. and M. Székely (1999). "Ahorro de los hogares y distribución del ingreso en México." *Economía Mexicana NUEVA ÉPOCA* 8(2): 267-338.
- Attanasio, O. and M. Székely (2000). "Household Saving in Developing Countries - Inequality, Demographics and All That: How Different are Latin America and South East Asia?" *Research Department Publications* No. 4221, Inter-American Development Bank.
- Beckmann, E., M. Hake and J. Urvová (2013). "Determinants of Households' Savings in Central, Eastern and Southeastern Europe." *Focus on European Economic Integration* 3: 8-29.
- Browning, M. and A. Lusardi (1996). "Household Saving: Micro Theories and Micro Facts." *Journal of Economic Literature* 34(4): 1797–855.
- Butelmann, A. and F. Gallego (2000). "Household Saving in Chile: Microeconomic Evidence." *Economía Chilena* 3(1): 5–24.
- Deaton, A. (2005). "Franco Modigliani and the Life Cycle Theory of Consumption." *Banca Nazionale del Lavoro Quarterly Review* 58(233-234): 91–107.
- Dupas P. and J. Robinson (2013). "Why Don't the Poor Save More? Evidence from Health Savings Experiments." *American Economic Review* 103(4): 1138–71.
- Garganta, S. and L. Gasparini (2015). "The Impact of a Social Program on Labor Informality: The Case of AUH in Argentina." *Journal of Development Economics* 115(July): 99–110.
- Granda, C. and F. Hamann (2015). "Informality, Savings and Wealth Inequality." Borradores de Economía No. 873, Banco de la República de Colombia.
- Gutierrez, M. (2007). "Savings in Latin America after the Mid-1990s. Determinants, Constraints and Policies." CEPAL – Serie Macroeconomía del Desarrollo No. 57.
- Harris, M., J. Loundes, and E. Webster (2002). "Determinants of Household Saving in Australia." *Economic Record* 78: 207–23.

International Labour Organization (2002). “Decent Work and the Informal Economy.” International Labour Conference 90th Session, Report IV.

Kanbur, R. (2009). “Conceptualizing Informality: Regulation and Enforcement.” IZA Discussion Papers No.4186.

Lorenzo, F. and R. Osimani (2001). “Alternativas de Política para Fortalecer el Ahorro de los Hogares de Menores Ingresos: El Caso de Uruguay.” CEPAL Serie Financiamiento del Desarrollo No. 110.

Madeira, C. (2011). “Computing Population Weights for the EFH Survey.” Mimeo, Central Bank of Chile, March.

Maloney, W.F. (2004). “Informality Revisited.” *World Development* 32(7): 1159–78.

Maurizio, R. (2012). “Labour Informality in Latin America: The Case of Argentina, Chile, Brazil and Peru.” Brooks World Poverty Institute Working Paper No. 165.

Ogbuabor, J.E., V.A. Malaolu and I.C. Mba (2013). “Informality and Domestic Savings in Nigeria: Lessons from Time Series Analysis.” *International Journal of Development and Economic Sustainability* 1(3): 24–32.

Perticara, M. and P. Celhay (2010). “Informalidad Laboral y Políticas Públicas en Chile.” ILADES-Georgetown University Working Papers No. 257, Ilades-Georgetown University, Universidad Alberto Hurtado, Chile /School of Economics and Business.

Pérez-Monteiro M., P. Cavalcante-Ferreira and L. Radusweski-Quintal (2014). “The Latin American Saving Gap.” Working Paper Anais do XL Encontro Nacional de Economia.

Reinhardt, N. (2008). “The Conundrum of Low Saving Rates in Latin America.” *Development Policy Review* 26(6): 727-744.

Sandoval-Hernandez, J. (2010). “Household Saving Behavior in México: Evidence from Cross-Sectional Analysis and Synthetic Panels.” Mimeo, University of California.

Solon, G., S. Haider, and J. Wooldridge (2013). “What Are We Weighting For?” NBER Working Papers No. 18859.

Tornarolli, L., D. Battistón, L. Gasparini, and P. Gluzmann (2014). “Exploring Trends in Labor Informality in Latin America, 1990-2010.” CEDLAS Documento de Trabajo No. 159.

Xiao, J. (1996). “Effects of Family Income and Life Cycle Stages on Financial Asset Ownership.” Association for Financial Counseling and Planning Education.



APPENDIX A

Translation of variable names in English to original names in Spanish as they appear in the EFH survey

Variables	Labels	Etiquetas
SR1	Savings rate (definition 1)	Tasa de ahorro (definición 1)
SR2	Savings rate (definition 2)	Tasa de ahorro (definición 2)
SR3	Dummy variable (0 did not save, 1 saved)- (definition 3)	Variable dicotómica (0 no ahorró, 1 ahorró) (definición 3)
<i>inflOD</i>	Dummy variable (0 Formal household; 1 Informal household)	Tasa de ahorro de hogares informales definición OIT (variable dicotómica)
<i>age</i>	Age of household head	Edad del jefe de hogar
<i>agesq</i>	Age squared of the household head	Edad al cuadrado del jefe de hogar
<i>thincome</i>	Total household income (in millions of pesos)	Ingreso total familiar (en millones de pesos)
<i>thincomesq</i>	Square Total household income (in millions of pesos)	Cuadrado del ingreso total familiar (en millones de pesos)
<i>thincome2</i>	Total household income w/o pensions (in millions of pesos)	Ingreso total del hogar sin contabilizar ingresos de pensiones (en millones de pesos)
<i>thincomesq2</i>	Square Total household income without pensions (in millions of pesos)	Cuadrado del ingreso total del hogar sin contabilizar ingresos de pensiones (en millones de pesos)
<i>minc</i>	Number of household members with income	Número de miembros del hogar con ingresos
<i>mincsq</i>	Square of number of household members with income	Cuadrado del número de miembros del hogar con ingresos
<i>empspo</i>	Spouse or live-in partner who is employed	Esposo(a) o conviviente que está empleado
<i>mchild</i>	Number of children at home	Número de niños en el hogar (menores de 18 años)
<i>melderly</i>	Number of elderly at home	Número de personas mayores en el hogar (mayores de 65 años)
<i>primary</i>	Primary education	Educación básica
<i>secondary</i>	Secondary education	Educación media
<i>tertoruniv</i>	Tertiary or university education	Educación superior o universitaria
<i>postgraduate</i>	Postgraduate education	Educación de postgrado
<i>bankacc</i>	Head of household who has current account	Jefe de hogar que posee cuenta corriente
<i>gender</i>	Gender of household head	Género del jefe de hogar
<i>durgood</i>	Home with durable goods	Hogar que posee bienes durables
<i>hhretired</i>	Head of household who is retired	Jefe de hogar jubilado o retirado
<i>dprecsav</i>	dummy precautionary saving	Variable indicadora de hogar que declara ahorro precautorio
<i>dretsav</i>	dummy retirement saving	Variable indicadora de hogar que declara ahorro para el retiro
<i>ddurgoodsav</i>	dummy saving durable goods	Variable indicadora de hogar que declara ahorro para compra de bienes durables
<i>ddebtreducsav</i>	dummy saving debt reduction	Variable indicadora de hogar que declara ahorro para reducir deuda
<i>infret</i>	interaction dummy retirement saving and informality	Variable de interacción entre dummy ahorro para el retiro y dummy hogar informal
<i>infprec</i>	interaction dummy precautionary saving and informality	Variable de interacción entre dummy ahorro precautorio y dummy hogar informal
<i>infdurgood</i>	interaction dummy durable goods saving and informality	Variable de interacción entre dummy ahorro para compra de bienes durables y dummy hogar informal
<i>infdebt</i>	interaction dummy saving debt reduction and informality	Variable de interacción entre dummy ahorro para reducción de deuda y dummy hogar informal