Economics Bulletin

Volume 40, Issue 4

Mandatory savings, informality and liquidity constraints

Carla Moreno
Emory University

Abstract

Using a national representative survey of households from Peru, this paper characterizes workers' decisions to participate in a pension system, which indicates labor formality. Empirical findings show that a worker's income level has a positive impact on his or her likelihood to participate. To account for these findings, a three-period overlapping generations model with liquid and illiquid assets is implemented. In the model, voluntary participation in the pension system is unattractive to individuals with income under a certain threshold. The retention of illiquid assets, such as pension funds, are not optimal given income constraints. Thus, the liquidity constraint set by a pension system with a mandatory savings policy induces these workers to choose informality.

Acknowledgements: Special thanks to Kaiji Chen, David T. Jacho-Chavez and Karen Kopecky for their guidance and valuable advice. Thanks to the editor John P. Conley, associate editor Julio Davila and two anonymous referees for comments and suggestions.
Citation: Carla Moreno, (2020) "Mandatory savings, informality and liquidity constraints", *Economics Bulletin*, Volume 40, Issue 4, pages 3274-3295
Contact: Carla Moreno - carla.moreno@emory.edu.
Submitted: March 03, 2020. **Published:** December 23, 2020.

1 Introduction

Mandatory savings policies are a funded alternative to pension systems. The self-financed design is attractive to emerging economies after experiencing foreign debt shocks and domestic banking crises (Holzmann, 1997). Many countries in Latin America adopted this solution, partially or completely, setting a policy that mandates workers to save a percentage of their income for retirement. Nevertheless, workers in economies with an unregulated labor market are able to skip contributions for retirement. For purposes of this paper, a worker not contributing to the pension system is defined as an informal worker (Arabsheibani et al., 2006).

For many countries in the Latin American region, informality represents more than 50 percent of the labor market. This fact is manifested in the reduced participation and coverage from pensions that are based on mandatory savings. For example, in Peru, less than 20 percent of the workforce actively contributes to the pension system and more than 70 percent of the elderly do not receive a contributory pension (Saco et al., 2014).

There are many factors that lead workers to voluntarily choose a job in the informal market. Previous literature on informality discussed the role of moral hazard, preferences for current consumption, external savings opportunities (Packard, 2007), flexibility in hours (Maloney, 2004), and liquidity preferences (Holzmann et al., 2000), amongst many others. Most of the previous research displays the importance of these factors based on empirical evidence; however, models exposing the mechanisms behind them are limited. This study contributes to the literature providing a theoretical model of income and liquidity factors on informality decision for liquidity constraint individuals.

Using the 2011-2017 waves of the National Household Survey from Peru (ENAHO), this paper accounts for the impact of income level on the decision to work informally. The two-assets approach of Kaplan et al. (2014) is used to rationalize this relation and describe how the liquidity constraint imposed by mandatory savings changes according to the household's income level. Households with lower income level are affected by an income constraint and, at the same time, the liquidity constraint binds with more frequency. The model explore the effect of this mechanism in the informality decision and is able to show that workers will choose informality if their optimal illiquid savings is below the minimum set by the savings policy. As a result, one can find an income threshold under which the probability of informality is higher.

The rest of the paper is organized as follows: Section 2 introduces an outline of previous literature, follow by the empirical evidence and characteristics of the informal labor using the case study of Peru in Section 3. Section 4 develops the theoretical model and Section 5 summarizes the results of a numerical exercise showing the effects of the liquidity constraint. Section 6 concludes.

2 Theoretical Background

In the current literature, informal work is many times understood as an optimal decision made by the worker. Maloney (2004), Packard (2007) and Bosch and Maloney (2010) evidenced a voluntary entry to the informal labor market and García and Badillo (2018) argues is present in at least a subset of informal workers. Maloney (2004) introduces the informal labor in Latin America as a sector closer to the "*voluntary entreprenurial small firm*" found in advance economies rather than the traditional interpretation as a residual and inferior sector. His definition does not imply that all workers in the informal sector are not in poverty but that they are better off choosing the informal job over the formal which they are qualified for.

A frequent variable to explain the decision to work in the informal market is the over-regulation of the formal market, which makes the informal unregulated sector more attractive. Williams (2017) evaluates this explanation based on 3 possible factors: high tax rates, high level of corruption, high level of state interference in the free market. Results using 36 countries show evidence of slight significance of income tax and none of the other measures. Williams finds that high levels of informality have a stronger relation with under-regulated and underdeveloped economies as well as lack of state protection of workers from poverty. These findings suggest that informality decisions are not only determined by a response to tax burden, other factors linked to the worker's poverty risk and income should also be considered. Moreover, there is not a unique way how an individual variable affects the sector decision. For example, a job with higher wages in one sector would not automatically imply a better sector for all workers (Maloney, 2004). There are different effects on informality from the interaction of factors such as preferences and moral hazard (Packard, 2007), hours flexibility, stability, recognition, non-cash benefits (Maloney, 2004), pension subsidies (Attanasio et al., 2011), taxes (Ulyssea, 2010), interest rates, amongst others. This paper contributes to the literature introducing a theoretical model to analyze a specific factor affecting the formality decision: liquidity preferences over income.

Holzmann et al. (2000) identifies liquidity as one of the factors that limits contributions to the pension system. Poorer households are income constrained and have higher rate of discount in future consumption; consequently, longevity risk is not of primary concern for them. Savings in an illiquid assets might set an intolerable constraint in poor household that want to smooth consumption. It also impose constraints on households investing in traditional or informal retirement arrangements which can extend from building a house to education of their children. Using a survey and behavioral experiment conducted in Peru, Barr and Packard (2005) found that having more children and more housing, as a share of accumulated asset, makes workers less likely to contribute to the pension system.

To explore the mechanism acting over retirement savings this study will borrow from models with asset accumulation to smooth consumption with a distinction between liquid or illiquid assets. Kaplan et al. (2014) contribution showed the importance on including illiquid savings when modeling individuals' behavior, introducing the *Wealthy hand-to-mouth* households. These households have zero (or very small) liquid assets and consume all of their disposable income each period, both characteristics of a hand-to-mouth household. How-

ever, they keep a positive level of wealth in illiquid assets creating a liquidity constraint that affects the consumption-savings decision. The distinction between wealthy hand-to-mouth and poor hand-to-mouth explains how come households with significant wealth choose to consume all their income each period instead of smooth consumption as expected by the life-cycle permanent income hypothesis.

This work adds to the literature on variables affecting the informality decisions providing a model to characterize the role of income and the liquidity constraint. An economy with mandatory savings for retirement imposes a liquidity constraint to every formal worker on it. Following similar behavior as the wealthy hand-to-mouth, workers with lower income level are affected by an income constraint. At the same time, the liquidity constraint created by the mandatory savings binds with more frequency for this group. This mechanism drive these workers to prefer the informal labor market where they do not need to save a minimum level on an illiquid asset for retirement.

3 Data and Empirical Findings

This section documents the characteristics of an economy with mandatory savings and a big informal labor market using the Peruvian National Household Survey (INEI, 2017) with quarterly data from 2011 to 2017. The survey has national representation across 28 quarterly periods. Estimations and analyses of the workforce will focus on the subsample of workers aged 18 to 65 years whose statistics about savings behavior have been available since 2015 (12 quarters). Formal workers are defined as those employed workers contributing to the pension system in the same or previous month as when the survey was administered. In this context, workers contributing to the pension system are also holders of an illiquid asset: their future pensions.

Workers differ in some key characteristics according to their formality status. Formal workers have in average higher education, higher income and are working more hours per week. Having higher education increases the likelihood of getting higher paid jobs; thus, higher income[1].

Table 1 summarizes the average real income of formal and informal workers by education and wether or not they have liquid savings. Workers with liquid savings have higher income even within education and labor market status groups. Following the wealthy hand-to-mouth theory outlined in Kaplan et al. (2014), lower-wage workers fall under a binding income constraint where liquid savings is not always desirable. Notice that even though formal workers conditional on education, have higher average income than informal ones, not everyone holds liquid savings. Given their mandatory retirement savings, formal workers face an additional liquidity constraint that affects their desire to save in liquid assets.

Similar savings behavior is observed when organizing results by income quintiles, seventy-five percent of the workers in the lowest income quintile do not save at all. The other 25

[1] A further description of the income distribution by sector is found in appendix A.

percent prefer to save only in liquid assets. In contrast, sixty percent of workers in the top income quintile have illiquid retirement savings and 20 percent have both, liquid and illiquid. Further analysis by quintile is presented in appendix A.

Table 1: Average Real Income by Education and Assets

	Informal		Formal	
Liquid Savings	Yes	No	Yes	No
Less than High School				
Average Real Income	268.99	194.37	435.04	398.96
	(6.58)	*(1.73)*	*(11.92)*	*(6.71)*
High School Education				
Average Real Income	354.17	296.20	518.79	482.62
	(5.46)	*(5.29)*	*(12.62)*	*(6.29)*
More than High School				
Average Real Income	440.35	418.80	807.82	777.74
	(8.79)	*(6.51)*	*(15.69)*	*(9.26)*

Note: Bootstrap Standard errors in parentheses. Average real monthly income (2011 USD) for 2015 to 2017 by education as reported in the ENAHO survey. Includes main and secondary jobs, before taxes and deductions.

These facts support the known importance of income level in savings decisions. Less income-constrained individuals will save at higher rates. Furthermore, when individuals are not income constrained, the facts illustrate the effects of having liquidity constraints on savings decisions. At higher income levels, being liquidity-constrained might be optimal even if it results in individuals holding no liquid assets. The data indicates that these wealthy but liquidity- constrained workers prevail in a system with mandatory illiquid pension savings.

Income level may influence an individual's decision over formality. At the same time, data shows that formal jobs pay higher salaries. Thus, a worker's income and formality status might be simultaneously determined. An instrumental variables approach, following Escanciano et al. (2016), is used here to identify the effect of income on formality status. First, an exogenous instrument is created using the residuals from the estimation of real income on the other explanatory variables of formality. Instead of incorporating an exclusion restriction in the first-stage regression, the non-linearity between real income and age is exploited to create an instrument that is not collinear with the other explanatory variables in the second-stage regression. In the second-stage the formality probability is regressed on the explanatory variables and the correction term from the first-stage.[2]

[2]Econometric model and results of first-step regressions are shown in Appendix B

Table 2: Formality Probability

Determinants	Logit	Cloglog
Log of real Income	2.297***	1.658***
	(0.059)	(0.047)
Log of Weekly Hours Worked	-0.467***	-0.406***
	(0.042)	(0.033)
Female	0.589***	0.463***
	(0.032)	(0.025)
Age	0.007***	0.000
	(0.001)	(0.001)
Self-employed	-1.964***	-1.431***
	(0.029)	(0.023)
Highschool Education	0.521***	0.452***
	(0.039)	(0.032)
More Than Highschool	0.755***	0.524***
	(0.060)	(0.049)
Correction term	-1.007***	-0.925***
	(0.059)	(0.048)
Constant	-15.101***	-10.920***
	(0.233)	(0.182)

Note: Survey-weighted estimation of formality probability for employed workers in the non-agricultural sector from the ENAHO survey. Bootstrap Standard errors in parentheses. *** p<0.01, ** p<0.05, and * p<0.1

Once the impact of income is isolated, its effect on the probability of formality is still positive and significant even after controlling for education, hours worked, age, self-employment, and gender. Table 2 presents the results for non-agricultural workers in Peru. Additionally, estimations using discrete income quintile variables show an increasing probability of formality as the income quintile is increased. Estimations by quintile are available in Appendix C. The relation of real income and the probability of being formal shows that workers' decision to contribute to the pension system depend positively on their income level.

4 Model

To recreate the observed impact of income on the formality decision this section propose an adaptation of Kaplan et al.'s (2014) portfolio decision model in the context of mandatory savings.

Consider an economy with J number of households that live during three life stages: young ($t = 1$), adult ($t = 2$), and retiree ($t = 3$). For each household j, preferences over

consumption are given by:

$$u(c_1^j) + \beta u(c_2^j) + \beta^2 u(c_3^j),$$

where $u' > 0$, $u'' < 0$ and the time discount factor $\beta \in (0, 1)$. Household j receives labor income y_t^j and consumes c_t^j during $t = 1, 2$, whereas in $t = 3$ the household only consumes.

In this economy, households choose between utility in the formal sector(v_j^f) or in the informal (v_j^i). Households will select the sector that gives them a higher lifetime utility, such that $U_j = \max\{v_j^f, v_j^i\}$.

Formal households are subject to a mandatory savings policy for retirement that sets a minimum percentage of the worker's labor income to be saved in a retirement account. The retirement account is an illiquid asset that can only be consumed in the last period, $t = 3$. Informal households can only save with liquid assets.

4.1 The Formal Household Problem

During the first and second life stage, household j works in the formal labor market, receives an income, y_t^j, and pays a variable income tax, τ. There are two savings instruments for transferring income to the future: a retirement account that is only available at $t = 3$, understood as an illiquid asset, and a savings account, m_t^j, which transfers resources into the next life stage as a liquid asset. The formal household follows a mandatory retirement savings policy contributing a percentage, x_t^j, of their labor income to their illiquid asset. The policy sets a minimum percentage at $\bar{x}$. Thus, for every working period the contribution to the illiquid asset is $x_t^j y_t^j$.

The formal household maximization problem is described in Equation (1), where j subscripts have been dropped for simplicity:

$$\begin{aligned} v^f = \max_{m_1, x_1, m_2, x_2} & \; u(c_1) + \beta u(c_2) + \beta^2 u(c_3) \qquad (1) \\ s.t. & \\ c_1 + m_1 &= (1 - x_1 - \tau) y_1, \\ c_2 + m_2 &= (1 - x_2 - \tau) y_2 + r^f m_1, \\ c_3 &= r^f m_2 + \zeta(x_1 y_1, x_2 y_2), \\ x_1 \geq \bar{x},\ & x_2 \geq \bar{x}, \\ m_1 \geq 0,\ & m_2 \geq 0. \end{aligned}$$

The formal sector has a one-period rate of return r^f for the liquid asset, m, and a return function $\zeta(.)$ for the illiquid asset. Households know their income path (y_1^j, y_2^j) and the return functions for assets. There is no uncertainty in the model.

Portfolio allocations and consumption decisions are made during the $t = 1$ and $t = 2$ life stages, while in period $t = 3$, the retiree has access to the retirement savings $\zeta(x_1^j y_1^j, x_2^j y_2^j)$ and possible liquid savings from a past period, m_2^j.

4.2 The Informal Household Problem

During the first and second period, individuals work in an informal job and receive an income y_t^j. The informal savings instruments are limited to liquid savings, m_t^j, with a return rate r^i. In the informal labor market, households avoid the mandatory savings policy and income taxes, τ. Nevertheless, households are subject to receive a non-contributory minimum pension, $\underline{c}$, accessible to retirees in $t = 3$ and financed with taxes, τ, paid by the formal sector. As in the formal sector, there is no uncertainty and no borrowing.

Therefore, dropping subscript j for simplicity, the life-cycle consumption maximization problem for the informal household can be written as:

$$\begin{aligned} v^i &= \max_{m_1, m_2} u(c_1) + \beta u(c_2) + \beta^2 u(c_3) \qquad (2) \\ &s.t. \\ c_1 + m_1 &= y_1, \\ c_2 + m_2 &= y_2 + r^i m_1, \\ c_3 &= \max\,[r^i m_2, \underline{c}], \\ m_1 \geq 0,\ m_2 &\geq 0. \end{aligned}$$

The minimum pension guarantees a consumption floor, such that:

$$c_3 = \max\,[r^i m_2, \underline{c}] \qquad (3)$$

Given there is no incentive in postponing consumption for period three household will rather consume the minimum pension and do not save for last period; thus, $m_2 = 0$. For the case when $c_3 = m_2 r^i$, the Euler equations will be same as in the formal sector problem.

Because of formal sector benefits, such as access to financial markets and "proof of income", the return rate in the formal sector is higher as the one available for workers in the informal sector [3].

$$r^f > r^i \qquad (4)$$

4.3 The Government Problem

The government collects income tax, τ, from formal labor workers and pays minimum pension benefits, $\underline{c}$. The government budget is in equilibrium and the constraint is always binding, such that:

$$G + n^{MP}\underline{c} = \tau \sum_{j=1}^{J} \mathbb{I}(y_1^j + y_2^j), \qquad (5)$$

where $\mathbb{I}$ is an indicator function that equals 1 if household j is formal; n^{MP} is the number of households that receive a minimum pension benefit, $\underline{c}$, and other government expenses are

[3] Empirical facts using Peruvian data showed that about 70 % of workers in informality do not have access to a bank account while this fraction is reduced to approximated 8 % for workers under formality.

defined as G. The minimum pension benefit $\underline{c}$ is set exogenously while the income tax, τ, is endogenously determined according to the number of formal households and n^{MP}.

The model captures the effect of real income on the decision to be formal and contribute to the pension system. Higher income-level individuals will prefer to be liquidity-constrained in order to gain higher returns in the future. They will save a higher portion of their income in illiquid assets, through higher contributions to their retirement fund. Because the mandatory saving policy for retirement adds a minimum bound to the illiquid savings, individuals with low income who have optimal illiquid savings values under this minimum bound are better off outside the mandatory policy and will choose to work informally.

5 Numerical Exercise

To illustrate the intuition, a numerical example is provided for an economy where workers ex-ante are heterogeneous in their endowment profiles (y_1^j, y_2^j) with $y_2^j > y_1^j$ for all j. The exercise uses $J = 500$ observations drawn from a log-normal distribution. Income tax, τ, is endogenously determined satisfying government budget in Equation (5). Parameters values are provided in appendix D.

Figure summarizes the results of this economy with a mandatory savings policy for retirement, $\bar{x}_{min}$, that sets the minimum contribution at 28 percent of the workers income. As a result, the formality rate for the economy is 34.60 percent, the informality rate 65.45 percent and 22.40 percent of the workers receive a minimum pension when retired. A worker in the formal sector saves for retirement a percentage of her income equal or bigger to $\bar{x}_{min}$ and pays income taxes of $\tau = 5.82\%$ every working periods. A proportion of informal workers save in liquid assets for their retirement period while the rest of them receive the minimum pension benefit, $\bar{c}$.

The size of the responses depend on the parameters and income distribution of the economy, which determines the number of people over the income threshold that will optimally choose to save equal or higher rates than the minimum $\bar{x}$. The model recreates the liquidity constraint mechanism over which mandatory savings affects the formality decision: higher mandatory savings rates induce more lower-income workers to choose informality. With lower number of workers choosing formality the income tax should increase to pay the non-contributory minimum pension, impacting formal worker's liquidity constraint a second time.

In this exercise, a policy that sets a minimum contribution bigger than 28% will increase the number of individuals with binding liquidity constraints, reducing the proportion of workers finding the formal sector optimal. As a consequence, the level of income tax τ will be higher to finance the minimum pensions. This will increase the number of workers opting out of the formal sector; finally, pushing the formality probability to zero.

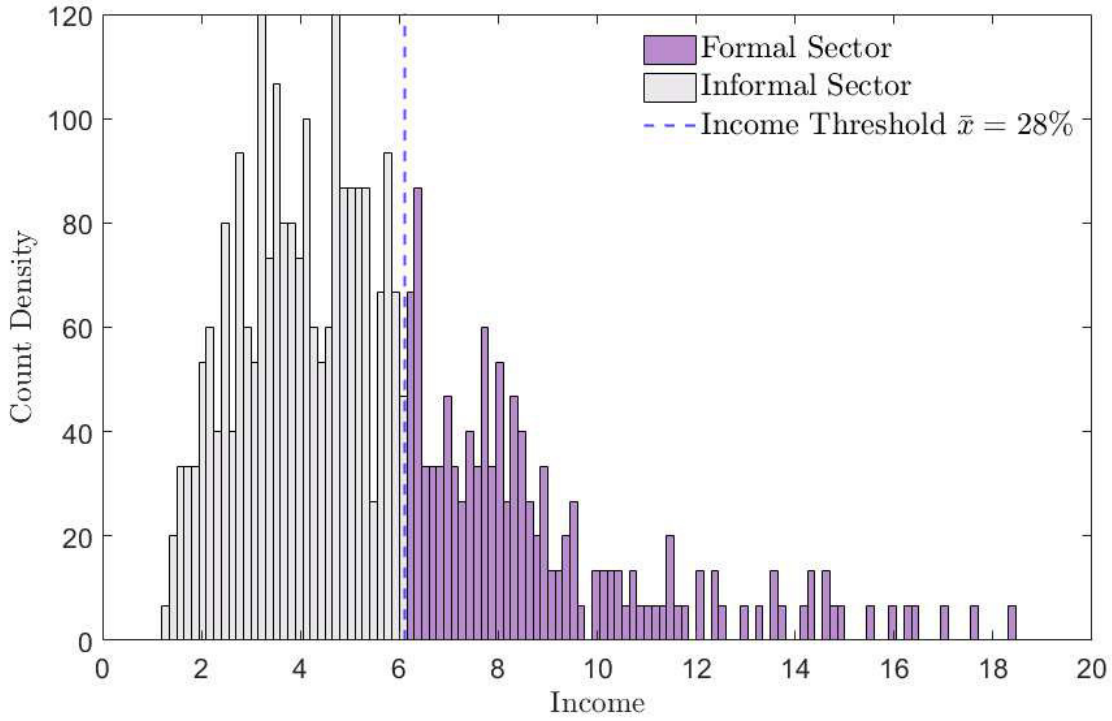

Figure 1: The figure shows the income distribution of the economy under a mandatory savings policies $\bar{x} = 28\%$. The dark shadowed bars represent the income levels in the formal sector and the light shadowed bars represent the ones in the informal sector. The dotted line shows the income threshold over which workers find optimal to work formally contributing $\bar{x}$ for retirement.

A robustness check is performed where the effects of taxes and minimum pension over the sector decision are eliminated by setting the minimum pension at zero, $\bar{c} = 0$. In this set up, the minimum mandatory savings rate, $\bar{x} = 28\%$, is the only government policy affecting the liquidity constraint. Similar to the previous exercise, the results in appendix D.3 indicate that the informal sector is preferred for very low-income workers. These workers still present a binding liquidity constraint imposed by a high savings policy for retirement in the formal sector. The income threshold is also lower because the trade-off of choosing formality (no access to the social pension, $\bar{c}$, and paying taxes, τ) are both zero. Therefore, this economy has a more attractive formal sector yet, maintains an informality level of 29.60%.

6 Conclusion

Mandatory savings are introduced as a way to finance retirement pensions; however, they also force a liquidity restriction on households. Using as framework the wealthy hand-to-mouth model (Kaplan et al., 2014) the paper develops a two-assets model that accounts for the savings decision when introducing mandatory saving policy and informal labor market to the economy. For lower-income households, the liquidity constraint is frequently binding. In economies with a large informal labor market, workers in this scenario have a possibility to opt-out of the policy choosing to work informally. The numerical results show that for workers under a certain income threshold, it is not optimal to hold illiquid retirement savings that are equal or higher than the ones impose by the mandatory policy. Therefore, such households will optimally choose the informal labor market. The model in this paper is used to assess the mechanism by which the liquidity constraint affects workers' sector decisions.

These results provide a theoretical framework for future analysis and debates on how pension systems are implemented in countries with informal labor markets. The minimum required contributions in the mandatory savings systems are set with the aim of accumulating enough savings during the working period to afford a pension for the retirement years. As shown in this paper, the setting of this parameters in countries with informality have additional effects over the number of workers opting for informal jobs. A policy where the minimum contribution is high might secure adequate pensions to formal workers but will increase the number of individuals outside the pension system working informally. Thus, leaving a higher number of workers unprotected in their old age.

An extension of this study might include the analysis of other components affecting the formality decision such as interest rates, the return function on savings, risk aversion, the social pension level, discounting and taxes (Attanasio et al., 2011), amongst others.

References

Arabsheibani, G. R., Carneiro, F. G., and Henley, A. (2006). *On Defining And Measuring The Informal Sector*. On Defining And Measuring The Informal Sector. The World Bank.

Attanasio, O., Meghir, C., and Otero, A. (2011). Pensions, work, and informality: The impact of the 2008 chilean pension reform. *Unpublished paper, University College London*.

Barr, A. and Packard, T. G. (2005). Seeking solutions to vulnerability in old age: preferences, constraints, and alternatives for coverage under peru's pension system. *CSAE WPS/2005-05*.

Bosch, M. and Maloney, W. F. (2010). Comparative analysis of labor market dynamics using markov processes: An application to informality. *Labour Economics*, 17(4):621–631.

Escanciano, J. C., Jacho-Chávez, D., and Lewbel, A. (2016). Identification and estimation of semiparametric two-step models. *Quantitative Economics*, 7(2):561–589.

García, G. A. and Badillo, E. R. (2018). Rationing of formal sector jobs and informality: The colombian case. *Journal of International Development*, 30(5):760–789.

Holzmann, R. (1997). Pension reform, financial market development, and economic growth: Preliminary evidence from chile. *IMF Econ Rev*, 44(2):149–178.

Holzmann, R., Packard, T., and Cuesta, J. (2000). Extending coverage in multi-pillar pension systems: Constraints and hypotheses, preliminary evidence and future research agenda. *World Bank Social Protection Discussion Paper*, 2.

INEI (2017). Encuesta nacional de hogares,enaho. https://webinei.inei.gob.pe. Accessed on 06/04/2018.

Kaplan, G., Violante, G. L., and Weidner, J. (2014). The wealthy hand-to-mouth. Report, National Bureau of Economic Research.

Maloney, W. F. (2004). Informality revisited. *World Development*, 32(7):1159–1178.

Packard, T. G. (2007). *Do Workers in Chile Choose Informal Employment? A Dynamic Analysis of Sector Choice*. The World Bank.

Saco, C., Amparo, M., Seminario, B., Leiva, F., Moreno, C., and Zegarra, M. A. (2014). La desestructuración del sistema peruano de pensiones.

Ulyssea, G. (2010). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics*, 91(1):87–99.

Williams, C. C. (2017). Tackling employment in the informal economy: A critical evaluation of the neoliberal policy approach. *Economic and Industrial Democracy*, 38(1):145–169.

Appendix A Savings and Income

A.1 Income Distribution

Figure 2 plots the formal and informal workers' income distributions. Because the mandatory savings policy is enforced only to salaried workers, the minimum income of these workers is also affected by the minimum wage regulations of the country.

Nevertheless, there exists a significant overlap in the sector choice by income level exposing the workers choices to contribute or not to the pension system beyond the minimum wage. 75% of workers in the non-agricultural sector that are contributing to the pension system have a log real income level equal or greater to 6.91, while 75% of the informal workers have a log real income level lower than 6.91. These facts support previous findings in the literature, where wages are not the only factor affecting formality and helps understand the sector choice as a voluntary decision.

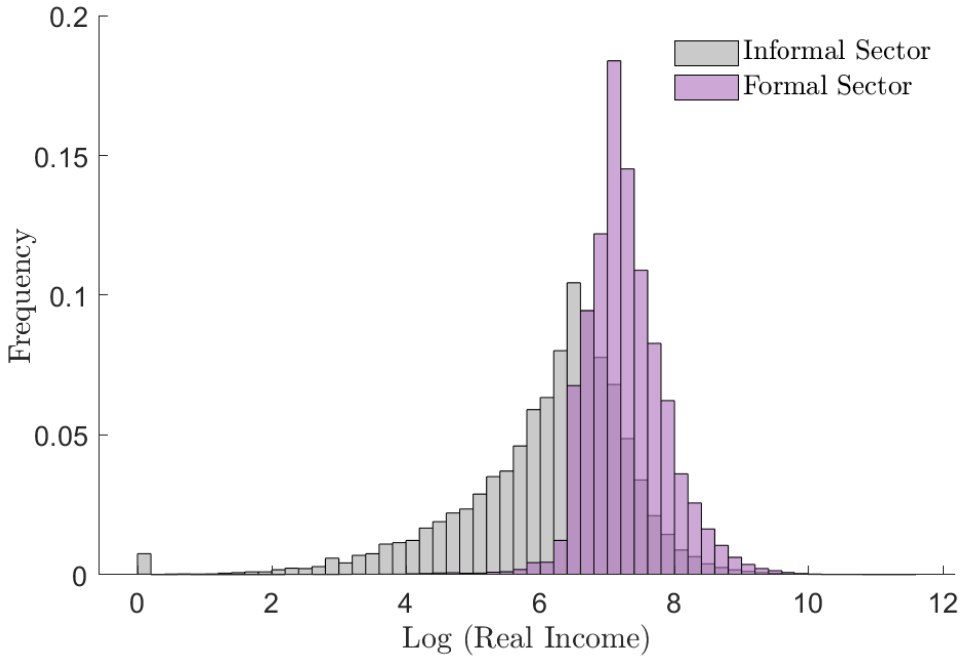

Figure 2: Distribution of average real monthly income in 2011 USD for non-agricultural workers between 18 to 65 years old as reported in the ENAHO survey. Includes main and secondary jobs, before taxes and deductions. Workers in the formal sector are defined as workers contributing to the pension system

A.2 Savings Decision

In Figure A.2, workers are grouped into five income levels. If a worker contributes to the pension system, $a = 1$; otherwise, $a = 0$. If a worker holds liquid savings, $m = 1$; otherwise, $m = 0$. Empirical facts for Peru are in line with evidence found in the literature, higher income households have higher levels of illiquid assets, while lower income households tend to save only liquid assets or not at all.

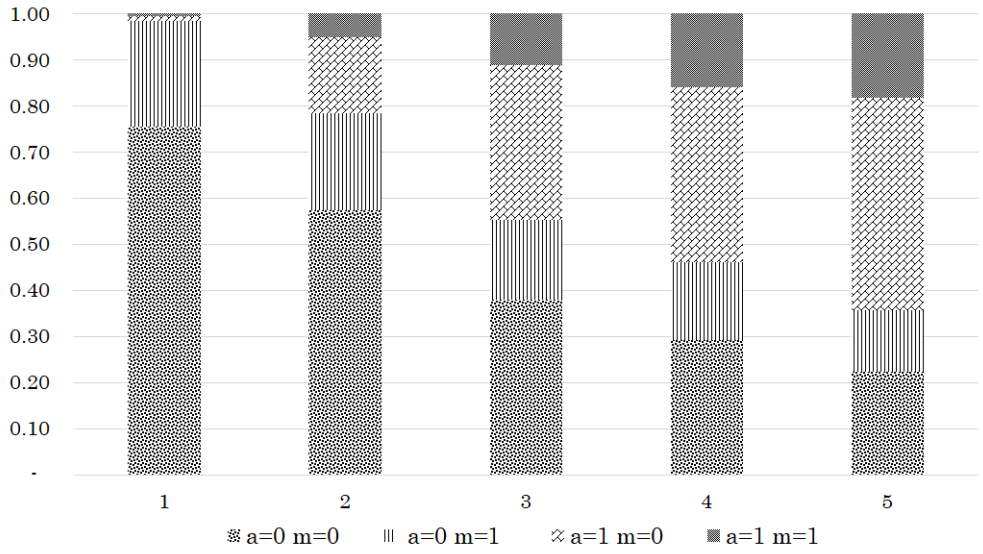

Figure 3: Calculations using survey answers from workers between 18 to 25 years old in the ENAHO survey's, quarterly data from 2015-2017. Liquid savings are represented by a and illiquid savings refers to contributions to the pension system, represented by m. Income quintile is based on real monthly income in USD 2011 and includes the main and secondary job, before taxes and deductions.

Appendix B Econometric Model

The model to estimate is specified as:

$$Formality = \mathbb{I}(\alpha_0 + \alpha_1 Education + \alpha_2 Hours + \alpha_3 Gender + \alpha_4 Age + \alpha_5 SelfEmployed + \alpha_6 Agriculture + \delta_0 LogIncome - e \geq 0)$$

Given the simultaneous determination of income and formality, unobservables that affect the formality decision are likely to be related to unobservables affecting labor income. The problem of endogenity is addressed following the methodology in Escanciano et al. (2016). Using the non-linear relation of income and age, the methodology allows to estimate a correction term based on the residuals that will capture the external factors affecting both income and formality probability. The semi parametric methods have proven to strongly identify parameters with the advantage of not relying on finding an exogenous instrument, which can be a difficult quest.

Solutions with semi parametric regression.
Log of real income is the endogenous regressor X^e that can be identified as:

$$X^e = g_0(X_1) + u \tag{6}$$

with endogeneity taking the form of:

$$e = \gamma u + \epsilon$$

In this estimation, X_1 includes the variables education, hours worked, gender, age, self-employment and agricultural labor indicators. Equation (6) can be restated as:

$$\text{Log Real Income} = g_0(Education, Gender, Age, Hours, Agriculture, SelfEmployed) + u$$

Let X_1 non parametrically define $g_0(X_1) := E[X^e/X_1]$ then $E[u/X_1] = 0$.
Having the function $g_0(.)$ non linearly explaining X^e provides sufficient source of identification to disallow perfect colinearity between the estimation residuals, $\hat{\mathbf{u}}$, and X_1. Thus, in the first step it is possible to recover the residuals from Equation (6) and use them in the second stage to capture the endogenous effects and successfully estimate:

$$Formality = \mathbb{I}(\alpha_0 + \alpha_1 Education + \alpha_2 Hours + \alpha_3 Gender + \alpha_4 Age + \alpha_5 SelfEmployed + \alpha_6 Agriculture + \delta_0 LogIncome + \gamma \hat{\mathbf{u}} - e \geq 0)$$

The methodology applied in this paper uses the variable **Age** as the non-linear regressor of the endogenous logged real income. This variable fits both assumptions needed according to Escanciano et al. (2016):

1. *Is a continuous regressor that is linear in Y, with complete support.*
   Age is a continuous and linear regressor for Formality probability. Non-linear relation between age and formality probability were rejected by Wald test using age^2 as an explanatory variable for formality.

2. *$g_0(X_1)$ is non linear in age.*
   Wald test using Age^2 as a explanatory variable for real income is significant and widely evidenced in the literature.

Results of first step estimation are reported in Table 3.[4] All coefficients are significant and later used to predict income. Finally, residuals values $\hat{\mathbf{u}}$ are given the name of "correction term" which are introduced in the second step of the estimation. The final results of the estimation are in Table 2.

[4] Additionally, other sources of non linearity are introduced in $g_0(.)$ with interaction terms between gender and education, age and education; and agricultural labor and age.

Table 3: Determinants of Income

Variables	Log of Income
Highschool Education	0.206***
	(0.012)
More Than Highschool	0.513***
	(0.021)
Male x Education[1]	-0.188***
	(0.005)
Education[1] x Age	0.008***
	(0.000)
Self-employed	-0.321***
	(0.006)
Log of Weekly Hours Worked	0.878***
	(0.032)
Log of Weekly Hours Worked squared	-0.036***
	(0.005)
Female	-0.542***
	(0.005)
Age	0.063***
	(0.001)
Age squared	-0.001***
	(0.000)
Agricultural sector	-0.533***
	(0.022)
Agricultural sector x Age	-0.012***
	(0.001)
Constant	2.374***
	(0.058)
R-squared	0.38

[1] Deviation from mean education level

Note: Bootstrap Standard errors in parentheses.

*** p<0.01, ** p<0.05, and * p<0.1

Appendix C Robustness Check by Quintile

Using discrete variables for each quintile income level, a regression is performed with results shown in Table 4.

Table 4: Regression by Real Income Quintile

Variables	Logit
Real Income	
Second Quintile	0.689***
	(0.088)
Third Quintile	2.701***
	(0.081)
Fourth Quintile	3.653***
	(0.081)
Fifth Quintile	4.116***
	(0.082)
Log of Weekly Hours Worked	0.090***
	(0.026)
Female	0.244***
	(0.020)
Age	0.022***
	(0.001)
Self-employed	-2.199***
	(0.020)
High School Education	0.914***
	(0.029)
More than High School	1.649***
	(0.028)
Constant	-5.644***
	(0.131)
Observations	219,170

Note: Survey-weighted Logit estimation of the formality probability for employed workers in non-agricultural sector from the ENAHO survey. Bootstrap Standard errors in parentheses. *** p<0.01, ** p<0.05, and * p<0.1

As expected, having a higher level of income increases the odds of participation in the formal labor market. Using the predicted marginal proportions to compare the effects between income quintiles, while controlling for other factors, the probability of formal participation increases as real income increases.

The same intuition is found in the marginal probabilities. Marginal probability of formality is 42.8 percent for workers in the top income quintile, while for workers in the first quintile it is only 3.8 percent.

Table 5: Predictive Margins by Real Income Quintile

	Margin	Std. Err.	z	P > z
Quintile				
1	0.0387	0.0032	12.14	0
2	0.0565	0.0026	21.64	0
3	0.2378	0.0025	95.52	0
4	0.3754	0.0028	134.5	0
5	0.4286	0.0054	79.03	0

Note: The margin value shows the marginal probability to be formal for a non-agricultural worker according to the income quintile from the ENAHO survey.

Appendix D Numerical Results

The model is solve in two steps. The first stage, each household learns their endowment (y_1, y_2) and given the set of exogenous parameters, they make a decision over their optimal levels of asset holdings by solving both problems in Equations (1) and (2): Maximizing lifetime utility from working in the formal sector, v^f, and maximizing life utility from informal sector, v^i. In the second stage, the household compares utilities v^f and v^i, choosing the optimal sector.

The quantitative exercise was preformed for 500 observations drawn from a log-normal income distribution with $\mu = log(300)$ and $\sigma = 10$. The illiquid savings return function has the following specification:

$$\zeta(.) = z^2(x_1 y_1^f) + z(x_2 y_2^f).$$

D.1 Initial parameters

Households display a Constant Relative Risk Aversion (CRRA) utility function and the model is solved using the following exogenous parameters:

β	γ	r^f	r^i	z	$\underline{c}$	$\bar{x}_1$	G
0.95	2	1.35	1	1.45	1.3Min$\{y_1\}$	0.28	0

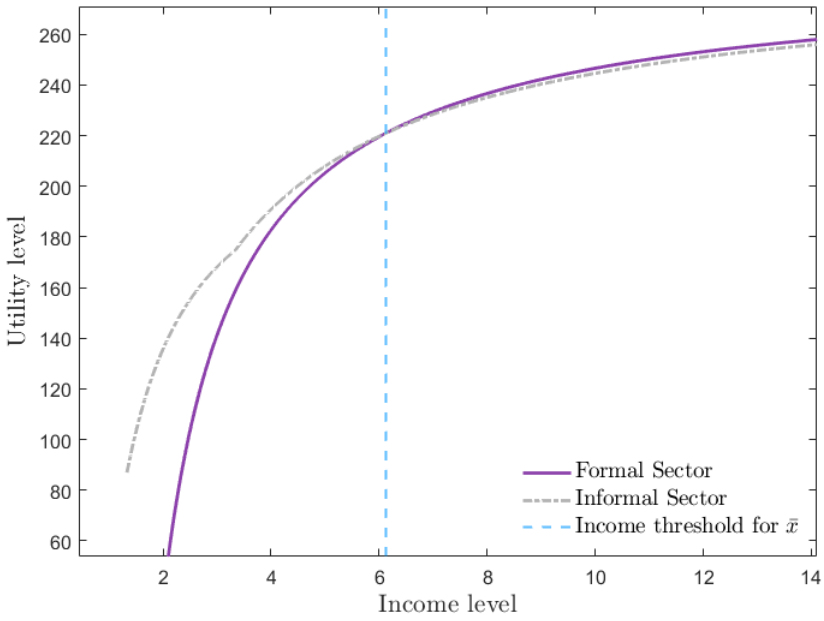

Figure 4: Compared utility level for each sector according to income endowment in the first period. The dotted gray line indicates utility levels achieved by solving the Informal Household's Problem in Equation (2). Income threshold indicates the income level under which the informal sector utility is higher than in the formal sector.

D.2 Robustness check: Utility specification

A robustness check is preformed for different utility specifications following the quantitative exercise for the same 500 observations drawn from a log-normal income distribution with $\mu = log(300)$ and $\sigma = 10$.

Constant Absolute Risk Aversion (CARA)

In this problem the households display a Constant Absolute Risk Aversion utility function:

$$U(C) = -\frac{1}{\alpha}e^{-\alpha C} \quad \alpha > 0$$

The exercise can be solve with the following exogenous parameters:

β	α	r^f	r^i	z	$\underline{c}$	$\bar{x}_1$	G
0.95	2	1.35	1	1.45	1.3Min$\{y_1\}$	0.16	0

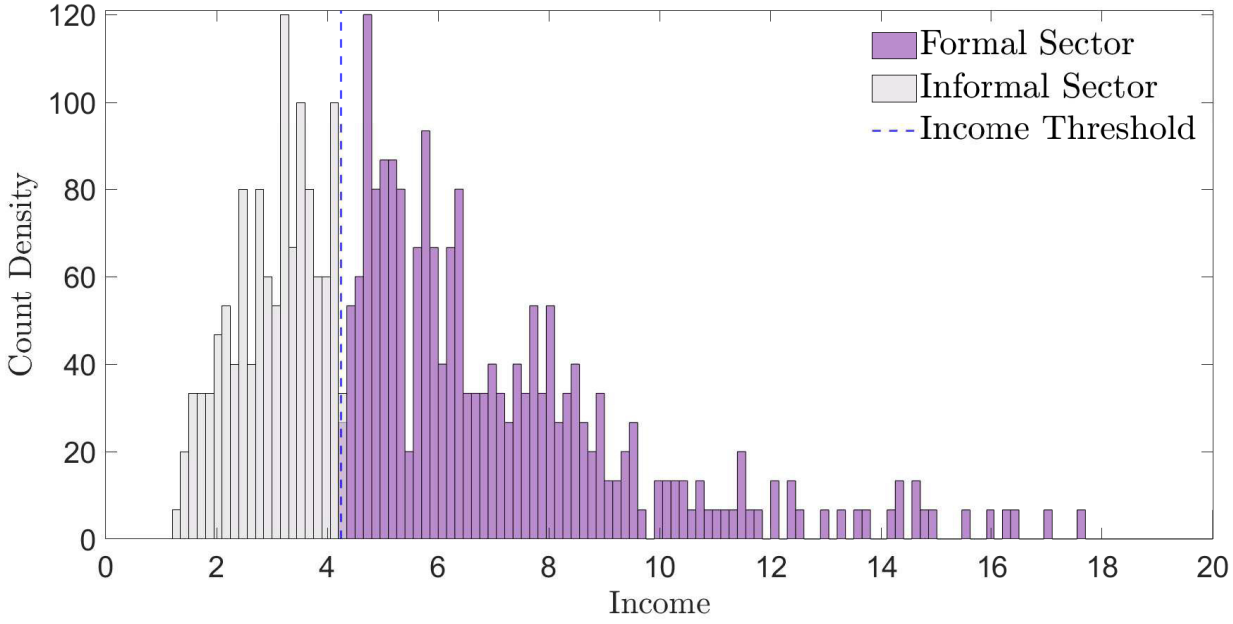

Figure 5: The figure shows the income distribution of the economy with households with a CARA utility function. The dark shadowed bars represent the income levels in the formal sector and the light shadowed bars represent the ones in the informal sector. Households with income level over the Income threshold find optimal to work formally given $\bar{x}$.

Quadratic Utility

In this problem the households display a quadratic utility function:

$$U = C - \frac{a}{2}C^2, \quad a > 0$$

The exercise can be solve with the following exogenous parameters:

β	a	r^f	r^i	z	$\underline{c}$	$\bar{x}_1$	G
0.95	0.2	1.35	1	1.45	1.3Min$\{y_1\}$	0.1	0

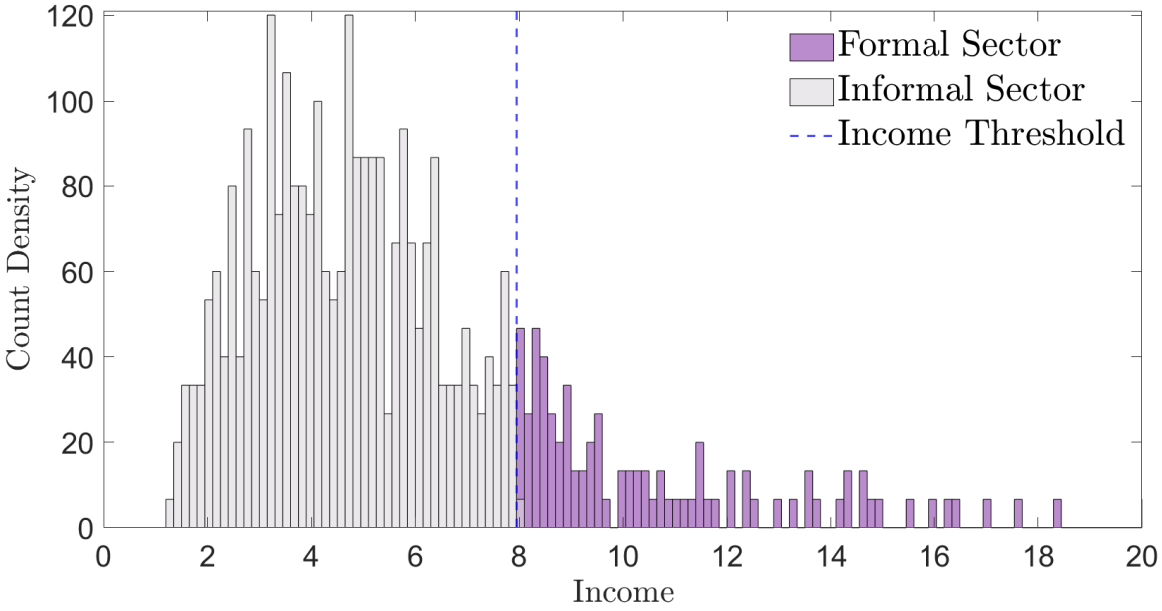

Figure 6: The figure shows the income distribution of the economy with households with a Quadratic utility function. The dark shadowed bars represent the income levels in the formal sector and the light shadowed bars represent the ones in the informal sector. Households with income level over the Income threshold find optimal to work formally given $\bar{x}$.

D.3 Robustness check: Minimum social pension and tax scheme

In this section we can observed the results of the simulation in an economy as described in the quantitative exercise in section D.1 but with no minimum social pension, $\underline{c} = 0$. The main mechanism behind the worker's decisions is the liquidity constraint. Results show a lower level of informality in this type of economy but still a positive percentage of the workers that are affected by the constraint will choose informality as their preferred sector.

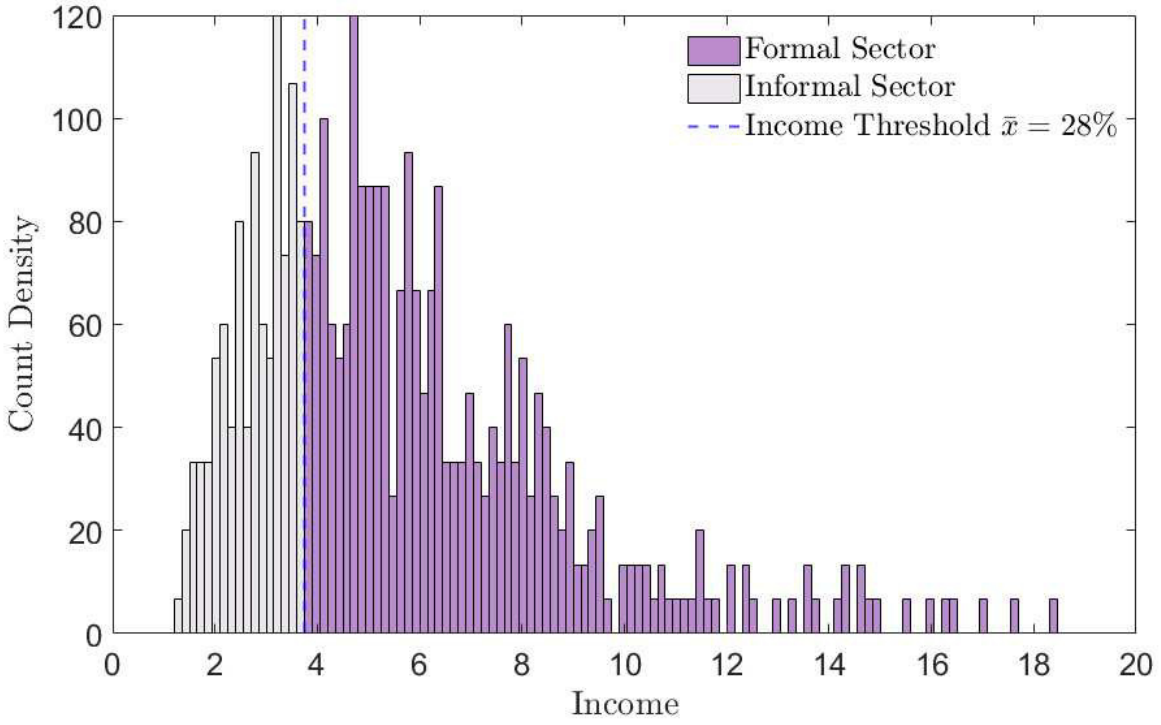

Figure 7: The figure shows the income distribution of the economy with households with a CRRA utility function. The dark shadowed bars represent the income levels in the formal sector and the light shadowed bars represent the ones in the informal sector. Households with income level over the Income threshold find optimal to work formally given $\bar{x}$.