Working and Saving Informally: The Link between Labor Market Informality and Financial Exclusion.*

Luca Flabbi† Mauricio Tejada‡

January 15, 2025

Abstract

The high level of informality and the low level of savings observed in developing countries are fundamentally linked because informal workers have limited access to formal financial institutions. We study this link by developing and estimating a labor market model where workers can be employed both formally and informally and agents can save through both formal and informal financial institutions. We estimate the model on Colombia and use it to simulate counterfactual experiments. Results show that the interaction of working and saving informally reduces savings: If full financial access were guaranteed to informal workers, net monthly savings would increase by 17%.

Keywords: Informality, financial inclusion, savings, labor market search, structural estimation. *JEL Codes:* J3, J64, O16, O17.

*We thank for their insightful comments the seminar participants at: Catholic University of Milan (IT), Catholic University in Santiago (CL), Pennsylvania State University (PA), Pontifical Javeriana University in Bogotá (CO), University of Cambridge (UK), University of Chile (CL), University of Michigan-Ann Arbor (MI), University of Pittsburgh (PA), University of Wisconsin-Madison (WI); and the conference participants at: BSE Summer Forum (Barcelona, SP), DC Search and Matching Conference (Washington, DC), UMD-Lacea-WB Workshop on Informality (Washington, DC). Tejada gratefully acknowledges financial support from FONDECYT Iniciación, grant project No. 11196296 and ANID – MILENIO – NCS 2022_045.

†Department of Economics, University of North Carolina - Chapel Hill, USA and IZA, Bonn. E-mail: luca.flabbi@unc.edu.

‡Department of Economics, Universidad Diego Portales, Santiago Chile. Millennium Nucleus for the Study of Labor Market Mismatch: Causes and Consequences. E-mail: mauricio.tejada@udp.cl.

1 Introduction

Developing countries are characterized by high informal employment and by low saving rates. While informality may introduce some useful flexibility in overly regulated labor markets, it also lowers workers' protections and increases their employment risks.[1] A number of contributions has also shown that its presence distorts incentives and returns to productive investment decisions, hindering productivity.[2] Low saving rates make individuals more vulnerable to shocks and the overall economy less resilient[3] The literature has also shown that low saving rates in low- and middle-income economies cannot simply be explained away by the presence of many individuals "too poor to save": limited access to financial services is playing a crucial role.[4]

If both high levels of informality and low levels of saving are problems in themselves, they are also strongly linked to each other. The presence of high informality increases the need for precautionary savings because of the higher employment risk faced by informal workers. However, informal workers typically face financial exclusion that prevents them from saving enough. The interaction of informal work and financial exclusion may therefore introduce a vicious cycle leading to more informality and less savings. In addition, this deep link also makes it extremely difficult to study and address high labor market informality and low savings in isolation, without considering the important interactions and feedback loops between them. Despite this, the literature typically analyzes them as two separate issues.

This paper goes beyond this approach by developing an environment that integrates all the crucial elements giving rise to both phenomena. To this end, we develop a labor market model where workers can be employed both formally and informally and where agents can save through both formal and informal financial institutions. The reduced access to formal financial institutions captures the intensity of financial exclusion. To provide a quantitative assessment and to evaluate policy experiments, we identify and estimate the model on data from Colombia. Colombia is a good candidate to conduct the analysis because it belongs to a region – Latin America and Caribbean (LAC) – where both issues are particularly acute[5] and collects good-quality data on

[1]For an overview of the issues, see Jessen and Kluve (2021); La Porta and Shleifer (2014); World Bank (2012); Perry et al. (2007).

[2]The distortions affects both the firms side, with consequent loss of productivity (Ulyssea, 2018; de Paula and Scheinkman, 2010; La Porta and Shleifer, 2008), and the workers side, with consequent loss of human capital accumulation (Bobba et al., 2021, 2022).

[3]For an overview of the issues, see Cavallo et al. (2016); World Bank (2012); Karlan and Morduch (2010).

[4]See Dizon et al. (2020); Dupas and Robinson (2013); Batini et al. (2010) for experimental evidence; Ogbuabor et al. (2013) for macro evidence; Bond et al. (2015) for model-based empirical evidence; Karlan and Morduch (2010) for a survey.

[5]Colombia is the fourth economy in LAC. The informal sector employs between 30% and 80% of the workforce

both savings and labor market behavior.

Specifically, we develop a search model of the labor market with savings. We extend the previous literature on search and savings by allowing for different types of jobs and different types of assets to be present in the same economy. Depending on the adherence to labor market regulations, jobs can be formal or informal. Depending on the features of the Institutions in which they are held, assets can be formal or informal. Formal financial Institutions include banks, employee funds, credit unions; informal financial Institutions include informal group savings associations such as RoSCA or Chit funds.[6] Financial exclusion is represented by the higher costs that workers employed informally face in accessing formal financial institutions. The equilibrium characterizes the distributions of individuals working formally or informally, the portfolio choice between formal and informal assets, the saving rate, the unemployment rate and the total wealth level.

We estimate the model combining information from two household surveys for Colombia: The *Gran Encuesta Integrada de Hogares* (GEIH), a nationally representative monthly survey focusing on the labor market; and the *Encuesta Longitudinal Colombiana* (ELCA), a longitudinal survey focusing on consumption, saving behavior and access to financial services. The identification of the labor market parameters is relatively standard while the identification of the parameters related to financial exclusion is novel and relies on a parsimonious specification of portfolio allocation costs. The estimation results confirm the presence of a degree of financial exclusion for informal workers: they need to pay fifteen times as much as formal workers to maintain the average portfolio allocation of formal workers. The results also show that formal workers receive better and more frequent job opportunities than informal workers.

We use the estimated model to perform counterfactual policy experiments. We find that providing full financial inclusion to informal workers increases monthly net savings by 17% and the proportion of wealth in formal assets by 38%. These improvements do not increase labor market informality but do increase inequality in assets and in lifetime consumption.

Our paper contributes to numerous strands of the literature. First, we contribute to the now

in the region (Gasparini and Tornarolli, 2009) and more than 40% of the world's GDP produced by the informal sector is produced in LAC (Schneider et al., 2010). Savings in the region equal 17% of GDP compared to the 30% of other middle-income countries (World Bank's World Development Indicators). The low saving rate has persisted despite policy efforts aimed at increasing it and despite relatively good macroeconomic conditions over a number of years (Reinhardt, 2008; Cavallo et al., 2016).

[6] RoSCA stand for Rotating Savings and Credit Association and it is typically defined as an informal financial institution consisting of a group of individuals who make set contributions and withdrawals to and from a common fund. Chit funds are common in South Asia and some of them have been now formalized. In LAC, RoSCA funds are virtually all informal.

large literature on savings in developing countries.[7] The literature has shown that the low level of savings is not simply driven by the population being "too poor to save" but is significantly influenced by the institutional context. In it, the interplay of labor market informality and financial exclusion plays a prominent role but empirical results on this relation is mixed. In LAC, Lorenzo and Osimani (2001) show that informal households in Uruguay have lower saving rates than their formal counterparts. Schclarek and Caggia (2015) use a financial survey for Chile to show that, controlling for other determinants of savings, informal households save, on average, less than formal households, more so in the case of precautionary savings. In contrast, Granda and Hamann (2020) use a financial survey for Colombia to find that the saving rate of those who are informally employed is higher than that of those who are formally employed. In other developing countries, Dupas and Robinson (2013) argue that informal households save less because they are financially excluded. The experiment they run in Kenya shows that providing a safe place to save increased health care savings by more than 60%, supporting the existence of capital market segmentation between the formal and the informal sectors (see also Batini et al. (2010)). From an aggregate perspective, Ogbuabor et al. (2013) use time series data for Nigeria showing that informality potentially hinders the growth of aggregate domestic savings due to the lack of a well-working mechanism of financial intermediation. Financial segmentation also implies that informal households have lower access to formal credit and depend strongly on informal sources of financing (Gatti and Honorati, 2008; Dabla-Norris and Koeda, 2008). Our contribution provides an explanation for these mixed empirical results since we are able to estimate a joint model of both informality and financial exclusion that endogenizes both labor market and saving decisions. In addition, our counterfactual experiments provide novel estimates of the costs of financial exclusion in terms of savings, labor market outcomes, inequality and welfare.

Second, our contribution relates to the growing literature using models with frictions to explain the equilibrium effects of labor market informality.[8] Search model of the labor market, in particular, have become a popular framework to explain why different labor market contracts for similar workers can survive in equilibrium and to provide a tractable theoretical environment for empirical investigations. While this literature has been effective in producing counterfactual experiments and in assessing a wide range of issues – from the impact of enforcement to the effect

[7] See Karlan and Morduch (2010) for a general survey on saving and financial access; Karlan et al. (2014) for a survey focusing on saving by the poor; Cavallo et al. (2016) for a flagship publication focusing on LAC; Eeckhout and Munshi (2010) for saving in informal financial institutions.

[8] Important early contributions in the theoretical literature are Albrecht et al. (2009); Charlot et al. (2013); and in the macro literature is Bosch and Esteban-Pretel (2012). Contributions estimating this class of models on micro data are fewer and more recent, they include Meghir et al. (2015); Bobba et al. (2021, 2022).

on schooling decisions –, none of the previous contributions have been able to take into account saving behavior. Therefore, they cannot assess the impact of the financial exclusion associated with labor market informality.

Third, we contribute to the literature that analyzes saving behavior in the presence of idiosyncratic risk when labor markets are affected by frictions.[9] While previous contributions in this literature are the closest to our formal setting, they do not address labor and financial market informality nor focus on data from low- and middle-income countries.[10] But this focus is needed since workers in low- and middle-income countries face higher idiosyncratic risk, larger imperfection in capital markets and stronger frictions in the labor market than workers in high-income countries. The need is exacerbated by the fact that results from high-income countries do not readily transfer to the institutional context of poorer countries since informality and financial exclusion both increase the idiosyncratic risk and distort the saving rate necessary to deal with it. Indeed, when financial institutions function well in developing countries, they usefully complement social protection programs – which are usually at the center of the policy debate – in helping households to manage risk (Perry et al., 2007). The focus on financial exclusion and informality also leads us to provide a methodological contribution in this literature: the agent in our model not only jointly decide on saving and labor market outcome but also on their portfolio composition, distributing their wealth or debt between a formal riskless asset and an informal risky asset.

Finally, we contribute to the sparse literature that attempts to jointly analyze labor market informality and optimal saving decisions. Esteban-Pretel and Kitao (2021) develop and calibrate on macro data a model which includes labor market informality and savings but does not allow for financial exclusion.[11] Granda and Hamann (2020) also provide a macro contribution: they develop a somewhat more stylized labor market and allow for a link between labor market informality and some disadvantage in the financial market. However, they cannot introduce genuine financial exclusion because they assume only one type of asset. Finamor (2024) is closer to us because

[9] For papers estimating the structural parameters of the model, see Rendon (2006); Lentz (2009); Lise (2013); García-Pérez and Rendon (2020); Abrahams (2022); for seminal theoretical contributions, see Danforth (1979); Acemoglu and Shimer (1999); Browning et al. (2007); for macro-oriented contributions, see Krusell et al. (2010); Bils et al. (2011); Kaplan and Violante (2014); Ji (2021); Setty and Yedid-Levi (2021); Pizzo (2022).

[10] One recent exception is Pierri and Kawamura (2022) studying the pension system in Chile. However, Chile is the only LAC country in which informality in the labor market is *not* a first order problem, as the authors acknowledge.

[11] Flórez (2017) is a theoretical paper that provides a similar earlier contribution to the one provided in Esteban-Pretel and Kitao (2021). As in their paper, the labor market includes an informal sector but the financial market does not.

he develops and estimates a model with labor market search, saving and informality. But he focuses on the only LAC country in which informality is not prevalent (Chile) and, as the other contributions, he does not allow for any form of financial exclusion.

The paper is organized as follows. In section 2, we describe the theoretical model, define the equilibrium, and briefly discuss the solution method. In section 5, we present the data, the identification, and the estimation method and results. In section 6, we analyze a series of counterfactual experiments, assessing the importance of labor market informality and financial exclusion on saving rates and inequality. Finally, in Section 7, we summarize the main results and provide some concluding remarks.

2 Model

The model characterizes the joint dynamic of labor market and saving decisions in presence of high levels of informality. To this end, the labor market environment must generate an equilibrium where different types of contracts arise even if agents are ex-ante identical, and the saving decisions setting must include a portfolio choice between at least two types of assets, one fully accessible and the other not.

A search model is a good candidate to describe a labor market with high informality: frictions allow different job contracts to survive in equilibrium and tractability delivers a dynamic model that can be identified with data available in low- and middle-income countries.[12] A search model is not an equally good candidate to describe borrowing and saving behavior since most search models of the labor market do not allow for it and assume linear utility. But a handful of contributions exist that have been able to develop and estimate search models with saving and borrowing decisions.[13] However, they all focus on high-income countries and assume a very simple saving behavior (workers can only save in one riskless asset), ruling out the possibility of saving in formal and informal assets. Therefore, we have to significantly extend the environment proposed by the previous literature in order to capture the crucial features we are interested in.

Our model includes a rich characterization of labor market informality, allowing for the presence of formal employees, informal employees and informal self-employed. Agents can save and borrow

[12] Previous contributions have exploited exactly these features to study Brazil (Meghir et al., 2015; Bosch and Esteban-Pretel, 2012), Mexico (Bobba et al., 2021, 2022), an aggregate of large LAC countries (Albrecht et al., 2009).

[13] The full list of published contributions currently is: Rendon (2006); Lentz (2009); Lise (2013); García-Pérez and Rendon (2020). The closest to our setting is Lise (2013). For additional references, including works in progress and theory and macro-oriented contributions, see footnote 9.

through two different assets – one riskless and formal, the other risky and informal –, an extension that has never been combined before with search models of the labor market. It is the essential feature needed to capture the partial financial exclusion we observe in these economies.

The resulting model's equilibrium characterizes the distributions of individuals working as formal or informal employees, the portfolio choice between the two types assets, the saving rate, the unemployment rate and the total wealth level. We assume as exogenous the wage offers distributions, some mobility parameters and the portfolio costs.

2.1 Environment

We assume continuous time and a stationary environment. Individuals discount the future at a common Poisson rate ρ and die at a common Poisson rate θ,[14] leading to the effective discount rate $\tilde{\rho} = \rho + \theta$. Individuals are ex-ante homogeneous but they differ ex-post in their labor market histories, and hence in the total wealth they are able to accumulate. They are risk averse and derive utility from consumption. We also assume that markets are incomplete in the sense that individuals cannot fully insure against labor income risk: they can only (partially) self-insure by accumulating assets.[15]

Individuals maximize the following additive separable expected lifetime utility function:

$$E_0 \int_0^\infty e^{-\tilde{\rho}t} \left[u(c - \underline{c}) + \epsilon f \right] \tag{1}$$

where c is consumption, $\underline{c}$ is the subsistence consumption level, f is an indicator variable equal 1 if the individual is working formally, and $\epsilon > 0$ is the additional utility capturing all the benefits received by workers hired in a formal job with respect to those who are not.[16] They reflect the fact that formal workers access a bundle of different benefits ranging from the guarantees implied by labor regulations to better social protection benefits.[17] Notice that, as we will specify below, this is *not* the only difference between formal and informal workers: most of the structural parameters

[14] The presence of a death shock ensures stationarity in the model environment since we assume agents are born with zero assets.

[15] This is the standard assumption in economies with heterogeneous agents (Huggett, 1993; Aiyagari, 1994).

[16] Informal workers are increasingly receiving some benefits, too, even if they do not contribute to it (so-called "non-contributory benefits"). Since we do not have access to the value of these benefits in the empirical application, we only model how much more formal workers receive with respect to the rest, as summarized by the parameter ϵ.

[17] In this parameterization we follow Conti et al. (2018) in their study of labor market informality in Mexico and Bobba et al. (2022) and others in their view that benefits are bundled. The assumption of separability is a restriction but helps identification in a model as ours. For a similar example in an estimated search and matching model see Dey and Flinn (2008a). In their contribution, the benefit was the provision of health insurance.

of the labor market will also be allowed to differ by formality status. The instantaneous utility function satisfies $u'(\cdot) > 0$ and $u''(\cdot) < 0$.

The labor market is characterized by four states: unemployment, self-employment, employment in a formal job, and employment in an informal job. We use the expression 'job' to denote being hired in a job as an employee and 'job offer' to denote the offer to be hired as an employee. We use the expression 'self-employment opportunity' to denote the possibility to work in a given position as self-employed. The informal sector is comprised by the self-employed and the informal employees. The self-employed are all considered informal in our model to match the data we will use in the empirical application. This reflect the type of self-employment in which the population at this skill level is engaging in. These are informal activities requiring negligible skills and capital such as selling soda cans in a street corner, a labor market state frequently called in the literature "subsistence self-employment".[18] We denote the measures of workers in unemployment with v, in informal self-employment with q, and in employment with $e(f)$. Unemployed workers receive a flow income b, which captures possible unemployment benefits and other transfers and subsidies.

Individuals in all four labor market states search for jobs and receive offers. Job offers arrive at Poisson rate λ^v to the unemployed, λ^q to the informal self-employed, and $\lambda^e(f)$ to the employed. The potential differences in arrival rates reflect underlining differences in search effort (for example, unemployed searchers can afford to search full-time while employed searchers cannot) and in opportunities (for example, employee workers may meet more employee job opportunities than self-employed workers do.) A job offer is a pair $\{w, f\}$ denoting the wage w offered for a job with formality status f. Wages are drawn from a mixture distribution of informal and formal wage offers, each of them denoted by $G(w|f)$ and assumed to be continuous. The proportion of offers requiring formality status f, i.e. the mixture weight, is denoted by $p(f)$.[19] If the primitive $G(w|f)$ distributions are entirely denoted by technology and firms' behavior, the actual mixture distribution of offers received by workers is determined by policies, too. First, we introduce a payroll contribution τ on formal wages so that a formal employee receives a wage equal to $(1-\tau)w$. Second, we allow for the presence of a mandatory minimum wage w_M. Compliance with the mandatory minimum wage only occurs among formal jobs.[20] As a result, the mixture

[18]See for example, Herreño and Ocampo (2023); Donovan et al. (2023); Bobba et al. (2022); Schoar (2010).

[19]Therefore, $p(0) + p(1) = 1$.

[20]Eckstein et al. (2011) estimate a search model of the US labor market with minimum wage and imperfect compliance. We use their idea of imperfect compliance here by precisely stating what determines compliance: being a formal or an informal employee. Flinn (2006) is another relevant contribution, estimating a search-matching-bargaining model of the US labor market. While the bargaining setting is different from ours, the paper shows the importance of spillover effects of a mandatory minimum wage in a labor market with frictions. A range of these spillover effects are present in our paper and some new are allowed to take place due to the presence of

wage offers distribution that searching individuals are faced with is:

$$G_W(w|w_M) = p(0)\int_0^w g(w|0)dw + p(1)\int_{w_M}^w \frac{g(w|1)}{1-G(w_M|1)}dw \ , \ w \geq 0 \tag{2}$$

The presence of the minimum wage not only affects the wage offers distribution at formal jobs but also the overall mixture because in the range $(0, w_M)$ no formal wage offers are made.[21]

Employee jobs are terminated at the exogenous Poisson rate $\eta(f)$. The termination rate is potentially different between formal and informal jobs to allow for technological differences and to capture an important institutional feature that we do not explicitly model. Even if the presence of informal jobs denotes that the enforcement of labor law is far from perfect, it does not mean that enforcement is completely absent. As a result, informal jobs may be discovered by the authorities and terminated. Such events are captured in reduced form by differences in termination rates for formal and informal jobs.[22] When a job is terminated, agents receive a self-employment opportunity and decide if accepting it and become self-employed or rejecting it and search full time for a job as unemployed. Self-employment opportunities are represented by the net income level y they generate for the self-employed worker and they are drawn from the exogenous distribution $S(y)$. When an agent is in self-employment enjoying income y, she may receive a shock to the value of her self-employment activity, a stylized way to capture entrepreneurial risk. The shock arrives at Poisson rate π and is represented by a new draw y' from the same primitive $S(y)$. Such a draw supersedes the old one and leads the agent to either continue in self-employment at the new income level y' or to change state and transition to unemployment in order to search full time.

The financial market is characterized by two types of liquid assets. The first asset, denoted with a_1, is a standard asset with rate of return r_1. This is the asset provided by formal financial institutions and, as specified below, its accessibility depends on the labor market state of the agent. The second asset, denoted by a_2, is a risky asset and has rate of return denoted by r_2. This is the asset provided by informal financial institutions and it is equally accessible to anyone. We assume that r_2 follows an exogenous distribution $R(r_2)$. We model the dynamics of

an informal sector and to the possibility to save and borrow.

[21]The payroll and the mandatory minimum wage are the two main institutional features that characterize formal workers in Colombia, together with a bundle of benefits that we capture with the utility parameter ϵ. Other possible features such as a progressive income tax or severance payments are irrelevant for the population on which we will conduct our empirical analysis and therefore they are not introduced in the model.

[22]We prefer this highly stylized way of capturing this possibility because we have very little information actual monitoring probabilities. For an example of an estimated search model that recover a cost function of the probability of being caught evading regulations, see Meghir et al. (2015) for Brazil and Bobba et al. (2022) for Mexico.

the rate of return of the risky asset as a Poisson process with rate κ. This variation of returns represents the main feature of informal financial institutions. Since they are not registered by governments, they lack proper supervision and regulation, leading to higher proportions of risky behavior by operators and higher incidence of financial misdemeanor. Still, reputation and other market mechanisms avoid huge swings in returns, keeping the process stationary and guaranteeing a certain degree of persistence (Cavallo et al., 2016; World Bank, 2012; Adentiloye, 2006; World Bank, 1989).

We denote total wealth for a given individual as $a \equiv a_1 + a_2$ and total wealth distribution in the aggregate with $\Lambda(a)$. The share of the formal asset in the total wealth is defined as $\phi \equiv \frac{a_1}{a}$. As in Bonaparte and Cooper (2009) and Bonaparte et al. (2012), we assume that individuals are able to decide their portfolio composition but also that maintaining a given composition has a cost. Specifically, we assume a convex cost function that depends on the agent's labor market status. It is denoted by $\Psi(\phi, j)$ where $j = v, q, e(1), e(0)$, and satisfies $\frac{\partial \Psi(\phi,j)}{\partial \phi} > 0$ and $\frac{\partial^2 \Psi(\phi,j)}{\partial \phi^2} > 0$ for all j.

The differential access to the formal financial system is therefore captured by the function $\Psi(\phi, j)$. Crucially, the partial exclusion experienced by informal workers is parsimoniously captured by the condition $\Psi(\phi, e(0)) > \Psi(\phi, e(1))$ for a given ϕ. Under the parameterization, informal workers are not fully excluded from the formal financial system but that they incur a higher cost in maintaining a proportion of their assets there, an interpretation fully consistent with the empirical evidence in the LAC region and that we will test empirically in Section 5.[23] Financial institutions charge a markup $(1 + \nu)$ over the savings rate of return when they lend. Therefore, the interest rate spreads are νr_i, $i = 1, 2$ for all workers in any labor market state.

When an agent dies, her assets are *not* passed on to the next generation: the newborn agent starts life with zero assets. Following Merton (1971), the individual budget constraint in a multi-asset optimal decision problem can then be written as:

$$da = \left[(r_1 \phi + r_2(1 - \phi))(1 + \nu I_{a^-})a + \iota(j) - c - \Psi(\phi, j) \right] dt \tag{3}$$

where $j = v, q, e(f)$; $f = 0, 1$; I_{a^-} is an indicator variable that takes value 1 if $a < 0$ (debt) and 0 otherwise; $\iota(j)$ is the labor market state specific income:

$$\iota(j) = \begin{cases} b & \text{if } j = v \\ y & \text{if } j = q \\ w(1 - \tau f) & \text{if } j = e(f); \end{cases}$$

[23] For evidence on LAC see Cavallo et al. (2016); for evidence on low- and middle-income countries in other world regions, see Eeckhout and Munshi (2010); Adentiloye (2006).

and τ is the pay-roll contribution paid by workers who are employed formally.

As in Aiyagari (1994), we assume a borrowing constraint $a \geq \underline{a}$, where $\underline{a}$ is a self-imposed borrowing limit. We assume it to be equal to the self-imposed limit in the worse case scenario, i.e. the scenario where the agent is permanently unemployed, has no assets and can only borrow in the informal financial system facing a 'very high' (but finite) interest rate that we denote with $\bar{\bar{r}}_2$.[24] In this scenario, the self-imposed borrowing limit equals to the flow of income in the unemployment state, appropriately discounted:

$$\underline{a} = -\frac{b}{\bar{\bar{r}}_2(1+\nu)} \tag{4}$$

In closing the description of the environment, it is useful to summarize the differences between the formal and informal sector. The formal sector includes only employees hired formally. A formal employee has unique benefits – summarized by the utility parameter ϵ – but also unique constraints – being subject to the payroll tax τ and the mandatory minimum wage w_M. In addition, all the labor market fundamentals are different for formal jobs: the arrival rate of offers $\lambda(1)$, the job termination rate $\eta(1)$, the underlying technologically-determined wage offers distribution $G(w|1)$. Finally, access to the financial system is different for formal employees because of the different portfolio costs they face, as represented by the function $\Psi(\phi, e(1))$. The informal sector includes both formal employees and self-employed workers. While they both share the lack of benefits and costs enjoyed by formal employees, they are also different along numerous dimensions. The labor market fundamentals are different – both in terms of mobility parameters[25] and underlining income distributions[26] – and the access to the financial system is different – through potentially different portfolio costs.[27] In addition, informal employee workers experience stable income w while on the job while self-employed workers are subject to income shocks with Poisson rate π.

This setting guarantees a rich range of channels through which differences between formal and informal workers may emerge. As we show in Section 5, the tight parameterization also allows for a full identification of the model and the subsequent quantitative evaluation of the importance of each channel.

[24] This assumption has the same interpretation as the one in Lise (2013), which states $\underline{a} = -\frac{b}{r_1}$. In Lise (2013), the worst case scenario is being unemployed forever without any assets. In our case, it is being unemployed forever without any assets and being forced to borrow from the informal financial sector at the 'worst' possible rate $\bar{\bar{r}}_2$.

[25] The arrival rate of employee offers are λ^q for the self-employed and $\lambda(0)$ for informal employee. Shocks on the job are also different, with an exogenous termination rate affecting informal employee at rate $\eta(0)$ and an exogenous shock affecting self-employment income at rate π.

[26] Informal employee offers are drawn from $G(w|0)$, self-employment income is draw from $S(y)$

[27] The portfolio cost functions are allowed to be different in the two labor market states, with $\Psi(\phi, e(0))$ denoting the cost for informal employee and $\Psi(\phi, q)$ the one for the self-employed.

2.2 Value Functions

Given her stock of assets, the agent chooses the paths of consumption c and portfolio composition ϕ and decides if accepting or rejecting a job offer, in order to maximize (1) subject to (3) and (4). This problem can be represented by a set of stochastic Hamilton-Jacobi-Bellman equations for each labor market state.[28] We denote with $V(a, r_2)$ the value of being unemployed with a stock of assets of a and facing a rate of return r_2 for the informal asset. We denote with $Q(a, r_2, y)$ the value of being informal self-employed with a stock of assets of a, facing a rate of return r_2 for the informal asset, and earning y. Lastly, we denote with $W(a, r_2, w, f)$ the value of being employed in a job with formality type f, receiving a wage w with a stock of assets of a and facing a rate of return r_2.

The steady state value of unemployment $V(a, r_2)$ satisfies:

$$\begin{aligned} \tilde{\rho} V(a, r_2) &= \max_{0 \leq c \leq \bar{c}, 0 \leq \phi \leq 1} \{u(c - \underline{c}) \\ &+ \frac{\partial V(a, r_2)}{\partial a} \left[(r_1 \phi + r_2(1-\phi))(1 + \nu I_{a^-})a + b - c - \Psi(\phi, v)\right] \\ &+ \lambda^v \int_w \max\{W(a, r_2, w, f) - V(a, r_2), 0\} dG_W(w|w_M) \\ &+ \kappa \int_{r_2'} \left[V(a, r_2') - V(a, r_2)\right] dR(r_2')\} \end{aligned} \tag{5}$$

Equation (5) indicates that unemployed individuals receive a flow utility $u(c - \underline{c})$ plus the expected change in the value of unemployment. The expected change is due to three processes. First, individuals accumulate or decumulate assets by the amount $\frac{da}{dt}$, which is valued at the marginal value of assets $\frac{\partial V(a, r_2)}{\partial a}$. Second, individuals receive job offers and may move from unemployment to employment as a result. Job offers (w, f) arrive at the rate λ^v and, if acceptable, generate a value gain equal to $W(a, r_2, w, f) - V(a, r_2)$. Third, the rate of return of the informal asset is subject to change which, conditional on the current asset position a, generates a value gain (or loss) of $V(a, r_2') - V(a, r_2)$.

[28] For a detailed derivation of the Hamilton-Jacobi-Bellman equations see Appendix A.1.

Similarly, the steady state value of informal self-employment $Q(a, r_2, y)$ satisfies:

$$\begin{aligned}
\tilde{\rho}Q(a, r_2, y) &= \max_{0\le c\le \bar{c}, 0\le \phi \le 1} \{u(c - \underline{c}) \\
&+ \frac{\partial Q(a, r_2, y)}{\partial a} \left[(r_1\phi + r_2(1-\phi))(1+\nu I_{a^-})a + y - c - \Psi(\phi, q)\right] \\
&+ \lambda^q \int_w \max\{W(a, r_2, w, f) - Q(a, r_2, y), 0\} dG_W(w|w_M) \\
&+ \pi \int_{y'} \max\left\{Q(a, r_2, y') - Q(a, r_2, y), V(a, r_2) - Q(a, r_2, y)\right\} dS(y') \\
&+ \kappa \int_{r_2'} \left[Q(a, r_2', y) - Q(a, r_2, y)\right] dR(r_2')\}
\end{aligned} \tag{6}$$

Equation (6) indicates that informal self-employed individuals receive a flow utility $u(c - \underline{c})$ plus the expected change in the value of self-employment, which includes four components. As before, individuals accumulate or decumulate assets (first component), receive job offers (second component), and are subject to changes in the rate of return of the informal asset (fourth component). The additional process is the shock to the flow value of self-employment y (third component). The shock arrives at Poisson rate π. If the shock hits, the individual decides if staying in self-employment or moving to unemployment. The first action generates either a value gain or a value loss equal to $Q(a, r_2, y') - Q(a, r_2, y)$; the second action generates a value loss equal to $V(a, r_2) - Q(a, r_2, y)$.

Finally, the steady state value of employment $W(a, r_2, w, f)$ satisfies:

$$\begin{aligned}
\tilde{\rho}W(a, r_2, w, f) &= \max_{0\le c\le \bar{c}, 0\le \phi \le 1} \{u(c - \underline{c}) + \epsilon f \\
&+ \frac{\partial W(a, r_2, w, f)}{\partial a} \left[(r_1\phi + r_2(1-\phi))(1+\nu I_{a^-})a + w(1-\tau f) - c - \Psi(\phi, e(f))\right] \\
&+ \eta(f) \left[T(a, r_2) - W(a, r_2, w, f)\right] \\
&+ \lambda^e(f) \int_{w'} \max\{W(a, r_2, w', f') - W(a, r_2, w, f), 0\} dG_W(w'|w_M) \\
&+ \kappa \int \left[W(a, r_2', w, f) - W(a, r_2, w, f)\right] dR(r_2')\}
\end{aligned} \tag{7}$$

where

$$T(a, r_2) = \int \max\left\{V(a, r_2), Q(a, r_2, y)\right\} dS(y) \tag{8}$$

Equation (7) indicates that employed individuals receive a flow utility $u(c - \underline{c})$ plus the expected change in the value of employment, which includes four components. As in the previous two cases, individuals accumulate or decumulate assets (first component), receive job offers (third component), and are subject to changes in the rate of return of the informal asset (fourth component). The additional process is the termination shock, which follows a Poisson process with rate $\eta(f)$. The shock terminates the job and the agent is force to go back and search. But she can choose if searching for an employee job as an unemployed or as a self-employed. This decision is defined in equation (8). Notice that the equation implies that the agent will receive a fresh draw of flow self-employment income y.

2.3 Decision Rules

The optimal consumption and portfolio decision rules are derived from the first order conditions of equations (5)–(7):

$$u'(c-\underline{c}) = \frac{\partial V(a,r_2)}{\partial a} \tag{9}$$
$$(r_1-r_2)(1+\nu I_{a^-})a = \Psi'(\phi,v) \tag{10}$$
$$u'(c-\underline{c}) = \frac{\partial Q(a,r_2,y)}{\partial a} \tag{11}$$
$$(r_1-r_2)(1+\nu I_{a^-})a = \Psi'(\phi,q) \tag{12}$$
$$u'(c-\underline{c}) = \frac{\partial W(a,r_2,w,f)}{\partial a} \tag{13}$$
$$(r_1-r_2)(1+\nu I_{a^-})a = \Psi'(\phi,e(f)) \tag{14}$$

Equations (9), (11) and (13) are the standard inter-temporal conditions, indicating that the marginal utility of consumption is equal to the marginal value of assets. These conditions imply that the optimal rules for consumption are $c^v(a,r_2) = \underline{c} + u'^{-1}\left(\frac{\partial V(a,r_2)}{\partial a}\right)$, $c^q(a,r_2,y) = \underline{c} + u'^{-1}\left(\frac{\partial Q(a,r_2,y)}{\partial a}\right)$, and $c^e(a,r_2,w,f) = \underline{c} + u'^{-1}\left(\frac{\partial W(a,r_2,w,f)}{\partial a}\right)$. Equations (10), (12) and (14) establish that the optimal portfolio allocation is obtained by equating the marginal benefit of adjusting the portfolio to its marginal cost. Notice that the marginal benefit depends on the differential return. The resulting optimal rules for the portfolio decisions are $\phi^j(a,r_2) = \Psi(j)'^{-1}\left((r_1-r_2)(1+\nu I_{a^-})a\right)$ for $j = v, q, e(f)$. Corner solutions on 0 and 1 are possible on these rules.

The optimal labor market decision rules about accepting or rejecting job offers have the usual form. An unemployed individual with assets a and facing a rate of return r_2 for informal assets accepts a job offer $\{w,f\}$ if $W(a,r_2,w,f) \geq V(a,r_2)$ and rejects otherwise. In turn, an employed individual with assets a and facing a rate of return r_2 for informal assets accepts a job offer $\{w',f'\}$ if $W(a,r_2,w',f') \geq W(a,r_2,w,f)$. Finally, an informal self-employed earning y, with assets a and facing a rate of return r_2 for informal assets accepts a job offer $\{w,f\}$ if $W(a,r_2,w,f) \geq Q(a,r_2,y)$. We illustrate these optimal decisions further using Figure 1 in the next section.

2.4 Equilibrium and solution method

The steady state equilibrium in the model is defined as:

Definition. Given:

1. The primitive parameters $\{\rho, \theta, \lambda^v, \lambda^q, \lambda^e(1), \lambda^e(0), \eta(1), \eta(0), \psi^v, \psi^e(1), \psi^e(0), b\}$;
2. The instantaneous utility function $u(c - \underline{c})$ and ϵ;
3. The distributions of wage offers $G(w|1), p(1), G(w|0), p(0)$ and self-employment income opportunities $S(y)$;
4. The institutional parameters $\{\tau, w_M\}$;

the *steady state equilibrium* is a set of value functions $V(a, r_2)$, $Q(a, r_2, y)$ and $W(a, r_2, w, f)$ that satisfies the Hamilton-Jacobi-Bellman equations (5)–(8), together with the invariant distribution of individuals across labor market states $\{v, q, e(1), e(0)\}$ and the invariant distributions of total assets $\Lambda(a)$.

We use a two-step approach to solve the model. In the first step, we solve the Hamilton-Jacobi-Bellman equations using a value function iteration method with an upwind scheme, while in the second step we solve for the invariant distributions of labor market states and of total assets by simulation. The full derivation is in Appendix A.2.

We show the main features of the model's equilibrium using simulation generated at our estimated values. These features are not specific to our estimates and can be replicated over a wide range of parameters' values. Figure 1 shows the evolution of the value functions as a function of wage offers by labor market state. The values are obtained by simulation and are shown conditioning on zero assets and average values of informal return. For the self-employed, we show the value function at the average self-employment income. Comparing value functions clarifies the optimal decision rules in the model. In this specific example, the self-employment value function is always lower than the unemployment value function: the agent, therefore, does not work as self-employed but chooses to search full time as unemployed. When receiving employee offers, the agent will accept an informal offer if the wage is higher than w_3, i.e. the point where the informal employee value function starts to become higher than the unemployment one. Analogously, the agent will accept a formal job if the wage is higher than w_4. Figure 1 also shows how job-to-job transitions occur. If the agent is currently an informal employee at wage higher than w_5, he will switch to formal employment when receiving an offer equal to, higher than or even slightly lower than his current wage. If the agent is currently a formal employee at wage higher than w_4 but lower than w_5, he will switch to informal employment when receiving an offer equal to, higher than or even slightly lower than his current wage.

These simple optimal decision rules have meaningful empirical implications. On top of generating an equilibrium where a non-trivial proportion of agents is present on each labor market states, they imply a rich set of transitions and a specific shape for the accepted wage distributions. First, ex-ante identical agents may transition to both formal and informal employment. Second, not only a direct transition from an informal to a formal job is possible but also the opposite: a current formal employee may decide to accept an informal employee offer, given the right combination of current and offered wages. Third, the average accepted wage for formal workers is higher than the one for informal workers because their reservation wage is higher ($w_4 > w_3$). But the two distributions exhibit a significant overlap: all the wages higher than w_4 are accepted both in formal and informal jobs. These empirical implications are fully consistent with the view that labor market in LAC are not segmented into a formal and an informal sector but organized in a way that allows for contracts with different formality regimes to survive in equilibrium and for a significant number of transitions between the two regimes.[29] As it will be shown in Section 3, these are also the empirical features that characterize our data.

The effect of endogenous saving and borrowing is also reflected in the value functions reported in Figure 1. First, it explains why formal employment is more valuable at high wage levels than informal employment. Two channels are at work in determining if informal employment or formal employment are more valuable at higher wages. The first is the payroll contribution: it is proportional, and therefore increasing with the wage level, but gives access to a fixed bundle of benefits. This channel favors informal employment at high wages. The second is the access to formal financial markets: formal employees pay less to keep their accumulated savings in the formal asset, which is safer and frequently offers higher returns than the informal assets. This channel favors formal employment at high wages. At our estimated parameters, and over the relevant range of wages and assets, this second effect is stronger, generating the shape reported in the figure. Second, the effect of endogenous saving and borrowing explains why both formal and informal employment value functions have a change of slope for very low wages. In this wage range, agents want to borrow and the slope is getting steeper because the lower the wage, the closer (or higher than) the desired borrowing level is to the borrowing limit.

[29] This is the view of all the models adding search frictions in labor market with high informality. See footnote 8 for detailed references. Early influential contributions proposing this view of labor markets in LAC include Bosch and Maloney (2010); Maloney (1999). Donovan et al. (2023) is a recent description of many labor markets with high informality which supports this view for the skill group we will consider in our application: unskilled or low-skill workers.

The interplay of labor market and savings decisions is also clearly seen in the three-dimensional Figure 2 where the value function for informal employment is represented as a function of both wages and assets. The figure shows that the Figure 1 we just discussed is just one "slice" of a much more heterogeneous behavior because all the value functions are also a function of asset levels. Figure 2 also shows how the model's equilibrium is generating an endogenous correlation between wages and assets, with the value function being concave in both variables, a result already emphasized by Lise (2013).

3 Data

We use data from a relatively large LAC country where working in the informal sector and saving outside formal financial institutions is common and widespread: Colombia. While other medium and large LAC countries share these features, Colombia has the advantage of providing researchers with good quality data on both labor market and saving behavior.[30]

3.1 Data Sources

We combine information from two data sources: a standard labor market survey – the *Gran Encuesta Integrada de Hogares* (GEIH) – and a survey focusing on individuals' saving and borrowing behavior – the *Encuesta Longitudinal Colombiana* (ELCA). We focus on 2016 since it is the last year for which both surveys are available together. We need to combine both datasets because the first, while giving a good description of the labor market, does not contain information on saving and borrowing; and the second, while collecting individuals' saving and borrowing information, does not contain enough information on the labor market dynamic.

3.1.1 GEIH

The GEIH is a nationally representative survey collected monthly by the *Administrative Department of National Statistics* (DANE). The survey contains individual characteristics, such as gender, age, and schooling; and provides labor market outcomes, such as employment status, durations, monthly labor income, weekly hours worked, and occupational characteristics. We pool together all the surveys from January to December of 2016. To be consistent with the theoretical model, we extract a relatively homogeneous estimation sample, conditioning on crucial demographic characteristics. We focus on individuals who are: male, 25 to 55 years old, living in

[30] Chile has a similar data availability but experiences much milder labor market informality.

urban areas, with only secondary education completed ("unskilled"), and working full-time when employed.[31] These controls generate a sample of individuals for whom informality both in the labor and in the financial market is likely to be a relevant issue. It is also the demographic group that constitutes the main component of the Colombian labor force.[32]

We define employed workers to be *informal* when they do not contribute to social security, a description consistent with the International Labor Organization (ILO)'s definition and with the definition used by the literature on the region (Perry et al., 2007; Kanbur, 2009; Bobba et al., 2022). These workers are composed of two groups. Workers that, while informal, are in a subordinate working relationship with a well-defined employer (*informal employees*); and workers that are occupied in an activity with more independence, such as selling cheap goods in a street corner (*informal self-employed*). This second group is increasingly considered a separate labor market state defined as "necessity" self-employment (Flabbi and Tejada, 2023; Bobba et al., 2022; Narita, 2020). As seen in Section 2, we follow this recent literature by introducing two different informal labor market states: informal employees and informal self-employed.[33]

We define employed workers to be *formal* when they do contribute to social security. Potentially, this group could also be composed of employed and of self-employed workers but we ignore formal self-employed because they are essentially absent in our estimation sample.

We define individuals as *unemployed* if they are classified as not working and actively searching for a job in the labor market.

The labor income variable we use is the gross real monthly wage expressed in December 2016 US dollars. GEIH is a cross-sectional survey so we cannot rely on a panel structure, not even on a rotating panel structure as the one in the US *Current Population Survey*, to collect information on individuals' labor market dynamic. Fortunately, the survey contains unemployment and employment on-going durations information.

[31] Full-time is defined as working 40 or more hours per week (top-coded at 100 hours per week).

[32] Individuals who are 25 to 55 years old represent 68% of the labor force, while the unskilled individuals represent 76% of the labor force in that age range.

[33] Older literature sums up self-employed workers who are not professionals or technicians with salaried workers with jobs not registered with the social security system (Meghir et al., 2015; Bosch and Esteban-Pretel, 2012). However, informal employees and informal self-employed not only exhibit different labor market dynamics patterns (Flabbi and Tejada, 2023) but also, as we will document shortly, different saving behavior. For this paper's objective, it is therefore crucial to allow for the two states to be different.

3.1.2 ELCA

The ELCA is a longitudinal survey that follows a representative sample of about 10,000 households in rural and urban areas every three years; the available waves are 2010, 2013, and 2016. This survey is carried out by the *Center for Studies on Economic Development* (CEDE) of the School of Economics at the *Universidad de los Andes*. It is part of a project designed to follow individuals for 12 years in order to collect information on a wide range of issues including income, consumption, and access to financial services.[34] Crucially, the survey contains the information missing from GEIH: individuals' saving and borrowing decisions. Individuals report whether they are able to save. If they do, they report average monthly savings and the proportion of their savings that are in formal or informal financial institutions. We define the first as banks, employee funds, credit unions and similar; we define the second as cash, informal group savings (such as ROSCA funds),[35] and similar.

3.2 Descriptive Statistics

Table 1 presents descriptive statistics on labor market outcomes computed using the estimation sample extracted from the GEIH survey. It shows that informality is a very important phenomenon in the sample of unskilled men since more than 50% of them work informally. A little less than 40% are hired formally and 7.9% are unemployed. Workers earn more, on average, if they are hired formally: average monthly wages are US$317 for formal workers and US$230 for informal employees; net average labor income for the informal self-employed is US$236, comparable to informal employees. But the three earnings distributions substantially overlap: about 30% of informal workers earn more than the median wage for formal workers. With respect to durations, in our sample unemployed workers search for a job for 4 months on average, while employees stays in the same jobs on average 66 months if formal and 34 months if informal. Informal self-employed stay in the job much longer, 105 months on average. Duration information confirms the higher instability of informal jobs but only for employees. It does also confirm the significant difference in labor market dynamic between informal employees and informal self-employed.

Descriptive statistics on saving behavior are presented in Table 2 and confirm well-known facts in the literature. Formal workers save a little bit more than informal workers: 27% of them have positive saving and save on average US$60 per month (leading to a saving rate of 13.2%).

[34] Detailed information on the survey is available at: https://encuestalongitudinal.uniandes.edu.co/en/.

[35] ROSCA stands for *ROtating Savings and Credit Association*, i.e. a group of individuals acting as an informal financial institution by setting up a common fund via set contributions and withdrawals

All these values are lower for informal employees. Informal self-employed exhibit an intermediate behavior, except on the saving rate which is the highest among all states. But where the difference is most striking is in the access to formal financial institutions: Almost 50% of formal workers put most of their saving in formal financial institutions while only 17% of informal employees and 19% of informal self-employed do. Unsurprisingly, the proportion of unemployed who save is significantly lower than the one for the employed but, when they save, they put their savings in formal financial institutions in higher proportions than informal workers (33%). This is a crucial feature that our model is able to account for thanks to the forward looking behavior and the search frictions that we assume. Individuals currently unemployed may had spells in formal employment in the past and therefore had relatively convenient access to formal financial institutions.

4 Identification

We are interested in identifying the full set of model parameters as listed in the Equilibrium Definition in Section 2.4. Some of these parameters are explicitly defined by the Colombian institutional context, such as the payroll contribution rate θ. These parameters do not need an identification strategy and therefore we simply discuss the source of the institutional information in Section 4.1. The identification of all the other parameters is provided in Section 4.2.

4.1 Institutional Values

The first three values are taken directly from the Colombian institutional context in 2016: the payroll contribution paid by formal workers is set to $\tau = 0.16$ and the mandatory minimum wage for formal workers is set to $w_M = 2.29$ dollars per hour to conform to the legislation in force at the time.[36] The formal asset rate of return is set to $r_1 = 0.037$ to match the yearly rate of return of a 10-year Colombian Government Bond in 2016.

The other values require some assumptions and some aggregation. We set the markup that financial institutions charge on loans to workers relative to the rate they pay on workers' savings to $\nu = 1.14$.[37] To obtain a value for the subsistence consumption level, we assume it to be

[36] We obtain the first value from OECD (2023), the premier publication providing details about payroll taxes in OECD countries; and the second from Colombian Secretary of Labor's Circular No. 3 of 2016 that sets the minimum wage at 689,455 Colombian pesos.

[37] This value is calculated as follows: $\frac{r^l}{r^d} - 1$ where r^l is the lending interest rate and r^d is the deposit interest rate. As reported in the IMF *International Financial Statistics*, these rates where 14.6 and 6.8%, respectively, for Colombia in the year 2016

equal to the consumption amount that corresponds to the poverty line. We therefore set it to $\underline{c} = 250$ dollars per year based on the poverty line used by the World Bank.[38] Since we assume a constant death rate in the population, we can calibrate it to match Colombia's life expectancy in 2016. Using the the value reported by the World Bank *World Development Indicators*, leads to $\theta = 1.2\%$.

4.2 Parameters Identification

The first step in the identification builds on the seminal contribution by Flinn and Heckman (1982). They show that hazard rates and termination rates can be identified by duration information, based either on complete or on on-going spells. Since we observe on-going duration information for the unemployed, the self-employed and for employed workers by formality status, we can identify fours hazard rates: out of unemployment and self-employment and out of formal and informal employment. However, hazard rates are not primitive parameters in the model since transitions between states are in general a function of the exogenous arrival rates and the endogenous acceptance probability. To make progress, a parametric assumption on the wage offers and self-employment income opportunities is necessary. We follow the literature[39] by assuming they are distributed as lognormal distributions with the following parameters: $\log(w)|f \sim \mathcal{N}(\mu(f), \sigma(f))$ and $\log(y) \sim \mathcal{N}(\mu^q, \sigma^q)$. Since the lognormal is a recoverable distribution, the primitive distribution can be identified by observing its truncation at a known truncation point. Under the model, the observed wage and self-employment income distributions are exactly truncations of the primitive distributions either at reservation values or at the minimum wage w_M and therefore they identify the parameters $\{\mu(0), \sigma(0), \mu(1), \sigma(1), \mu^q, \sigma^q\}$ and the proportions $\{p(0), p(1)\}$. With their identification secured, the hazard rates can be decomposed in arrival rates and acceptance probabilities, delivering the identification of the arrival rates $\{\lambda^v, \lambda^q, \lambda^e(1), \lambda^e(0)\}$. Adding aggregate proportion in the four labor market states allows for the identification of termination rates of employee jobs by formality status, the parameters $\{\eta(0), \eta(1)\}$.[40] As in Flinn and Heckman

[38] For a more detailed explanation, see Appendix C.

[39] See the review article Eckstein and van den Berg (2007) but also some more recent works such as Bobba et al. (2022); Tejada et al. (2021); Flinn and Mullins (2015)

[40] In our setting, this identification strategy is more complicated than in the original Flinn and Heckman (1982) and in similar applications. The issue is that a given duration may terminate in more than one state: for example, an unemployment duration may end due to a transition either to formal or to informal employment; similarly, an employment duration may end due to either a termination rate or an arrival rate of a better offer. However, since the durations we observe are only on-going, we do not know in which states they terminate. Appendix B shows in more detail how we can secure identification despite this added complexity.

(1982), the effective discount rate $\tilde{\rho}$ and the unemployment flow income b are only jointly identified: we choose to calibrate the discount rate, after which b is identified through the equilibrium version of the value functions (5).[41] As shown in Section 2.1, the discount rate ρ and the death rate θ combine to create the effective discount rate $\tilde{\rho} = \rho + \theta$. We calibrate ρ at the discount rate recommended for Latin America by multilateral development banks: 12% a year (Moore et al., 2020). As shown in Section 4.1, we calibrate the death rate θ at 1.2% based on Colombia's life expectancy in 2016.

The second step in the identification focuses on the portfolio cost functions $\Psi(\phi, j)$. We choose a functional form that is parsimonious and satisfies the properties $\frac{\partial \Psi(\phi,j)}{\partial \phi} > 0$ and $\frac{\partial^2 \Psi(\phi,j)}{\partial \phi^2} > 0$:

$$\Psi(\phi, j) = \frac{\psi^j}{2}\phi^2 \ , \ j = v, q, e(1), e(0) \tag{15}$$

Hence, the parameters to identify are ψ^v, ψ^q, $\psi^e(1)$, and $\psi^e(0)$.[42] Together with labor market income and initial assets, these parameters affect the optimal consumption and portfolio decision. Observing portfolio decisions conditioning on labor market income, assets and labor market state would therefore deliver the identification of the four ϕ parameters. However, we do not observe this full set of information in our data. We observe the amount of saving and the proportion of those savings in formal assets, conditioning on labor market status and income. We do not observe the amount of borrowing and total assets. However, since the portfolio cost function is characterized by only one parameter for each labor market state, we can exploit the optimal decision rules reported in Section 2.3 – in particular equations (10), (12), and (14) – to identify the parameters. The unobserved assets distribution is essentially integrated out exploiting the steady state proportion implied by the model. Notice that equations (10), (12), and (14) assume knowledge of the rate of returns, which we discuss it below.

The third step is the identification of the rate of return in informal institutions, a more challenging task than in formal institutions because there is no credible and systematic information on how informal financial institutions behave.[43] To make progress, we follow an idea developed by Eeckhout and Munshi (2010). They find that the implied rate of return for chit funds in South

[41] This is a common procedure in the literature: on top of the mentioned Flinn and Heckman (1982), see for example Eckstein and van den Berg (2007); Flinn (2002); Bobba et al. (2022).

[42] Instead of assuming non-convex portfolio costs as in Bonaparte et al. (2012), we simplify the analysis by assuming a convex cost function. Our setting is convenient because it generates a well-defined decision rule for the portfolio choice while keeping the model tractable without losing the Bonaparte et al. (2012)'s result of the portfolio cost being the source of partial segregation in the financial market.

[43] Recall we have already discuss the fairly standard calibration of the rate of return in formal assets in Section 4.1.

India, given a stable relationships in a matching equilibrium with informal financial institutions, is around 2.1 times the rate of return of formal financial institutions. We use this information to identify the scale of the rate of return. Recall that we assume that r_2 follows an exogenous distribution $R(r_2)$. We apply the same proportion proposed by Eeckhout and Munshi (2010) to our data in order to set the support of the distribution, assuming that most of the returns are not higher than 2.1 times the formal return. Formally, we assume that that 99% of the time r_2 belongs to the interval $[0, 0.078]$.[44] To make further progress, we need a distributional assumption. We assume normality $R(r_2) = \mathcal{N}\left(\bar{r}_2, \sigma^2_{r_2}\right)$ so that, together with the support assumption, we can fully recover the parameters $\bar{r}_2 = 0.039$ and $\sigma_{r_2} = 0.015$. The last parameter left to be identified to fully describe the process is the rate at which the rate updates, κ. The rate updates at different times for different workers but each worker face the same updating process in all labor market states. As a result, information on the savings and portfolio allocations observed in our data, together with the restrictions implied by the model, is enough to identify κ as one of the structural parameters in our estimation routine.

The fourth step concerns the instantaneous utility function. In terms of functional form assumption, we follow the relevant literature[45] by assuming a Constant Relative Risk Aversion (CRRA) utility function: $u(c-\underline{c}) = \frac{(c-\underline{c})^\delta}{\delta}$. Without additional sources of information, the relative risk aversion parameter δ is notoriously difficult to identify in this class of models. For example, Flabbi and Mabli (2018) accomplish its identification by adding moments from spouses' household interaction; Dey and Flinn (2008b) by using information on employer-provided health insurance. Lacking similar additional sources of information Colombia, we use an iterative procedure that starts from a credible calibrated value, estimates the rest of the structural parameters conditioning on it and then updates the original calibrated value until convergences. Additional details are provided in the estimation section, Section 5.

To summarize, based on the model presented in Section 2 and on the additional functional forms assumptions presented in this Section, the complete set of parameters we can identify with the data at our disposal is:

$$\Xi \equiv \{b, \epsilon, \lambda^v, \psi^v, \lambda^q, \pi, \psi^q, \mu_q, \sigma_q, \kappa, p, \delta\} \cup \{\lambda^e(f), \eta(f), \mu(f), \sigma(f), \psi^e(f)\}_{f\in\{0,1\}}$$

In addition, we calibrate the parameters $\{\nu, \underline{c}, \bar{r}_2, \sigma_{r_2}, \rho, \theta\}$ and we fix at their institutional value the parameters $\{\tau, w_M, r_1\}$.

[44] 0.078 is 2.1 times the formal rate of 0.037.

[45] See for example Dey and Flinn (2005); Krusell et al. (2010); Lise (2013).

5 Estimation (Preliminary)

5.1 Estimator

We estimate the parameters of the model using the Method of Simulated Moments (MSM). This method is commonly used to estimate highly non-linear models like ours (Gouriéroux and Monfort, 2002). Denoting with Ξ the set of parameters to be estimated, with M_N^D the set of appropriately chosen statistics derived from our data sample of size N, and with $M_T(\Xi)$ the corresponding set of simulated statistics extracted from a sample of size T obtained from the steady state equilibrium implied by Ξ, then our MSM estimator $\hat{\Xi}$ satisfies:

$$\hat{\Xi}_{N,T}(W) = \text{argmin}_{\Xi} \frac{1}{2} \left[M_N^D - M_T(\Xi)\right]' W_N \left[M_N^D - M_T(\Xi)\right] \tag{16}$$

where W is a symmetric, positive-definite weighting matrix.[46]

The set of chosen moment statistics in equation (16) are: the proportion of individuals in each labor market state; the average, the standard deviation, and the 5% percentile of the observed wages/earnings distributions for formal employees, informal employees, and the self-employed; the average and standard deviation of the ongoing durations in unemployment, formal employment, informal employment, and self-employment; the proportion of individuals in each labor market state who have more than 50% of their assets in the form of formal assets and the same proportions above and below the median of the accepted wage/earnings distributions;[47] and finally, the average and the standard deviation of the distributions of observed/declared savings by labor market state, and the average observed/declared savings above and below the median of the accepted wage/earnings distributions. The procedure generates a total of 39 moments that we use to estimate 21 parameters. The complete list of moments we match is reported in Table 5. We use the Nelder-Mead simplex algorithm to minimize equation (16) and bootstrap to compute standard errors.

As mentioned in Section 4, one parameter is estimated differently from the others: the relative risk aversion parameter δ. In this case, we use the following iterative procedure. We start from

[46]The inverse of the bootstrapped variance of each moment in the sample is typically used to construct the weighting matrix. Since our moments are computed using two sources of data with considerably different sample sizes, we use the following procedure to bootstrap the vector of moments in the estimation. For the GEIH survey, we follow the standard procedure. For the ELCA survey, in turn, we first expand the number of individuals in the ELCA survey to match the number of individuals in the GEIH survey by labor market states and income quartiles. Then, we extract random samples with replacement from the expanded version of the ELCA survey to compute the moments and their variance.

[47]We use only two positions in the accepted wage/earnings distributions because of the sample size in the ELCA survey.

the δ parameter used in Bond et al. (2008), one of the few papers estimating a labor market model for Colombia. Conditioning on that value, we estimate the other model's parameters using estimator (16). At the estimated model's parameters, we simulate the model over a grid of points for δ ranging between -0.5 and 0.5 and compute the loss function in (16).[48] We then select the δ that minimizes the simulated loss function and re-estimate the model. We iterate until convergence, where the convergence criteria is no further improvements in minimizing the loss function.

5.2 Results

Table 3 reports the estimated parameters. In the top panel, we present the Poisson rates of the labor market shocks that may induce transitions between states. The point estimates for the arrival rates of job offers imply that, on average, workers receive an offer every 5.8 months when they are unemployed, and significantly less often when they are employed: about every 3.5 years when formally employed as employee or when self-employed and every 4.9 years when informally employed as employee. The offers received are almost evenly split between formal and informal employee jobs. The point estimates for the termination shocks imply that jobs are exogenously terminated every 3 years on average in informal employee jobs, and at a much lower frequency in formal employee jobs (every 7.6 years). This much higher job instability in informal jobs confirms the prior present in the literature, and leads to the question if these workers facing the higher risks, are also those facing higher costs to accumulate precautionary savings. Shocks to self-employment income (either positive or negative) arrive at a rate higher than the formal termination shock but much lower than the informal termination shock.

Comparing these estimated structural parameters with previous literature is problematic because no previous work has developed and estimated a search model of the labor market with savings on any LAC country. However, a growing literature exists that estimates simpler search models of the labor market on some large LAC countries, Colombia included. These models do not include any saving or borrowing behavior and typically do not allow for search on the job. Conditioning on these limitations, our estimated values are comparable to previous works on Colombia and confirm that the country reports lower arrival rates than those estimated on other large LAC countries, such as Mexico or Brazil.[49] Among the few papers estimating a model with

[48] The grid is chosen to include the "consensus values" for the risk aversion parameter $1-\delta$ in both macroeconomic and microeconomic literature: we using as extreme the 1.5 estimated by Evans (2005) and the 0.5 estimated by Dey and Flinn (2008b).

[49] Tejada et al. (2021) estimate a search model of the labor market on four LAC countries, including Colombia.

saving and on-the-job search on high-income countries, the only meaningful comparison that can be made is with Rendon (2006): interestingly, he estimates a ratio between arrival rates off and on the job that is similar to ours.[50]

In the second panel we collect the wage offers distributions parameters and the self-employment income opportunities parameters. The point estimate of $p(1)$ indicates that 56% of the wage offers are for formal jobs. The estimates of the location and scale parameters of the wage offers distributions imply an average offered wage of 205 US dollars per month for formal jobs and of 166 for informal jobs (see top panel of Table 4). The standard deviation of wage offers is also higher for formal jobs: 112 compared to 57 for informal jobs. These results imply that the wage offers distribution of formal jobs stochastically dominates the one of informal jobs but also that the two distributions exhibit a substantial amount of overlap. Both implications are consistent with the previous literature on informality in LAC.[51] The estimates of the location and scale parameters for the self-employment income opportunities distribution imply an average of 185 US dollars per month, a value higher than the average informal employee wage but lower than the formal employee one. The standard deviation, equal to 151, is higher than the standard deviation in both wage distributions. This relative high dispersion makes the income shock π affecting such activity a source of significant volatility.

The third panel presents the estimated parameters of the portfolio cost function and therefore captures the extent of the informal workers' exclusion from formal financial institutions. The point estimates imply a cost parameter for informal employees that is sixteen times higher than the one for formal workers; for informal self-employed workers it is even higher. Figure 3 helps in understanding what these values mean by showing the portfolio cost function across the entire range of the share of formal assets $\phi \equiv \frac{a_1}{a}$. Consider a worker holding the average portfolio allocation of formal workers: a value equal to 44%, which corresponds in the figure to the dotted

While their model is much simpler, it shares the feature of a Poisson process for job offers: the rate they estimate for Colombia on the demographic group most comparable to ours is 0.1443, quite comparable with our 0.172. The values they estimate for the other three countries (Argentina, Chile and Mexico) are all higher. Our own previous work on Colombia (Flabbi and Tejada, 2023), again using a simpler model, returns a comparable arrival rate for the unemployed of 0.145. Bobba et al. (2022) and Meghir et al. (2015) use data on, respectively, Mexico and Brazil and estimate much higher arrival rates than those we estimate for Colombia.

[50] Direct comparisons with the mobility rates estimated by the the search and saving models by Lise (2013) and Lentz (2009) cannot be made because they allow for endogenous search effort.

[51] An implication of both Meghir et al. (2015) and Bobba et al. (2022) is that the average lower level of productivity of informal jobs will be reflected in lower average job offers for informal workers. Tejada et al. (2021) provide direct estimates of wage offers distributions for formal and informal jobs in four LAC countries and confirm the ranking found here for Colombia.

red vertical line. If such a worker holds a formal job, he incurs a portfolio cost of about 8 *cents* per month; if he holds an informal job, the cost increases to 1.2 dollars per month; if he is working as self-employed, the cost reaches 1.8 dollars per month. While these values are not large when compared to average wages, they are substantial when compared to the average returns that these workers receive on their assets. Consider for example an individual holding the average amount of assets held by savers in our economy and investing 44% of them in the formal asset. Further assume that he keeps this portfolio and asset choice for a year. His total return from formal assets will be equal to 37 dollars for the year. If he is a formal employee, his portfolio costs will be only 5% of his total annul returns on his formal assets. But if he is an informal employee, his portfolio costs will be almost equal to his returns on the formal assets. These high costs are one of the reasons why informal workers, both employees and self-employed, on average, hold a lower proportion of their savings in formal assets: as shown in Figure 3 the average ϕ for informal workers is about half that of formal workers. Yet, even at this much lower value, they pay an average fee that is higher than what formal workers pay at their average portfolio allocation (compare the intersection of the blue and purple dotted and continuous lines with the intersection of the red dotted and continuous lines).

The fourth panel presents the estimated parameter for the Poisson process of the rate of return on the informal assets. The arrival rate of shocks κ implies that, for each individual agent, the return changes on average once every two and half months. Figure 4 combines this estimated value with the calibrated parameters of the distributions of the return, $R(r_2)$, to show a comparison between three simulated returns trajectories that could hit an individual worker over a period of 5 years. The trajectories are compared with the constant returns provided by the formal asset (red horizontal line). The figure highlights the considerable risk involved in investing in informal assets: the returns vary over a quite wide range, and they may jump from values well above 5% to values close to zero in a matter of months.

The fifth panel presents the parameter corresponding to the flow income from unemployment. It implies that an unemployed individual receives 269 US dollars per year, which is slightly higher than the minimum subsistence consumption level assumed in the model (250 US dollars per year).

Finally, the sixth panel presents the estimated utility function parameters. The point estimate for δ implies a coefficient of relative risk aversion equal to 0.85, a positive value relative close to 1. Our agents are therefore characterized by risk aversion, with a degree of risk aversion comparable to the one found in other empirical microeconomics papers but lower than the one assumed in most macroeconomics studies.[52] This value is very close to Bond et al. (2008), one of the few

[52]For applied micro studies recovering risk aversion in search models, see Flabbi and Mabli (2018); Dey and

papers estimating a labor market model for Colombia. The point estimate for ϵ summarizes the utility value of the bundle of benefits received when working formally, in addition to the benefits of formality explicitly parameterized in the model.[53] To illustrate the magnitude of this effect, we compare the flow utility of an average formal employee with the flow utility of an average informal worker (employee and self-employed). Without the added utility benefit of ϵ, a formal worker's utility is 1.5% higher than an informal's utility. With the added benefit estimated by ϵ, it is 1.8% higher. Therefore, while the utility from additional benefits received by formal workers is positive and significant, it does not constitute the primary advantage of working formally. That advantage is conferred by the mechanisms we explicitly model in our environment.

To provide further illustration of what our estimated parameters imply, Figure 5 shows the steady-state distributions by labor market states of total assets, portfolio allocation, consumption, and savings. As expected, the distribution of total assets for formal workers first-order stochastically dominates that of those in the informal sector (Panel (a)). Formal workers are also overrepresented among those who hold the majority of their assets in the formal financial system (Panel (b)). On the other extreme, the majority of the self-employed hold most of their assets in the informal financial system. A substantial proportion of both formal and informal workers manage to have positive savings, even if the informals exhibit a much larger spike at zero savings (Panel (d)). The unemployed, instead, are characterized by significant dissaving. Consumption levels report clear rankings, with the unemployed consuming significantly less than any employed agents and the informal employees consuming less than the self-employed and the formal employees (Panel (c)). Overall, these results confirm the significant link between working and saving informally. They also show that the relevant link is over the intensive margin of saving, not the extensive one. In other words, informal workers save less and invest less in formal financial institutions but they still contribute a significant amount of savings to the economy and they are not completely prevented access to formal financial institutions.

5.2.1 Fit

Table 5 provides an assessment of the in-sample fit of the model by comparing the sample and simulated moments included in the quadratic form (16). The model fits well across most moments related to proportions in labor market states, wage offer distributions, and durations. The main discrepancies concern the unemployment rate (which we predict too low) and the 5th

Flinn (2008b); for a non-search applied micro reference see Chetty (2006). For applied macro studies recovering risk aversion in search models, see Guler et al. (2012); for a general empirical macro reference see Evans (2005).

[53] The benefits explicitly parameterized are, as described in Sections 2 and 4, the formality-specific labor market parameters and the lower costs to access formal financial institutions.

bottom percentile for the self-employed income (too high), probably because in the data they are generated by a relatively low number of observations. Still, ranking of all these moments in the data is preserved in the simulated model.

The estimated model fits very well the portfolio choices by labor market state, as represented by the proportion of agents keeping more than 50% of their wealth in formal assets. We also match well this statistic when we condition on labor market state and level of income (above or below the median of the relevant reference group). The worst fit is among informal employees with a wage above the median where we under-predict the proportion.

The fit on the actual amount saved is worse, even if a good number of moments are matched well. Recall that we observe only how much the agents save, not how much they dissave or borrow. Therefore, we build moments (mean and standard deviation) over the product between an indicator variable equal one if the agents has positive saving and the actual amount saved. We capture well the mean of this product for formal employees (even when conditioning on values above or below the median wage). Instead, we overestimate it on informal employees and underestimate it on the self-employed, so much so that we reverse the ranking between the two. This difficulty may reflect the lower quality of the data. The sample with saving information has a much smaller sample size than the sample with labor market information (see Section 3), a problem exacerbated by the lower number of informal employees and self-employed with positive savings. As a result, the moments we fit are less credible and noisier and, accordingly, are weighed less in our quadratic form (16). This problem is so acute on the unemployed that we do not attempt to match this moment on them, considering it more noise than information.[54]

6 Counterfactual experiments (Preliminary)

Using the model and the point estimates of the parameters presented in Table 3, we perform three sets of counterfactual experiments. In the first, we evaluate the importance of the partial financial exclusion of individuals not working formally by giving them full access, i.e. by equating the portfolio costs for all individuals to those of formal workers.[55] In the second, we approximate a technological or policy change that reduces the proportion of job offers that are informal. We reduce the proportion by the amount needed to match the same increase in average wealth in the informal sector ($E[a|e(0), a > 0] + E[a|q, a > 0]$) obtained by the full access experiment. In

[54]In our sample, only 3.6% of the unemployed report positive saving. Since in our estimation sample extracted by the ELCA data set we only have 83 unemployed, any moments computed on saving conditioning on unemployment would be determined by only 3 observations.

[55]Consistently with equation (3) and table 3's estimate, we set $\psi^e(0) = \psi^q = \psi^v = \psi^e(1) = 0.007$.

the last experiment, we evaluate an the impact of a tax reform. The simulated reform reduced the payroll to induce the same increase in average wealth in the informal sector ($E[a|e(0), a > 0] + E[a|q, a > 0]$) obtained by the full access experiment. The last two experiments provide a comparison between financial market and labor market policies efficacy in increasing wealth.

In each counterfactual scenario, we evaluate the average impact on labor market outcomes and financial outcomes (both stocks and flows) after taking into account the endogenous adjustments in individuals' optimal behaviors and in the steady-state distributions. We also analyze the impact of policy experiments on welfare. Following the literature,[56] we compute the welfare measure as the average steady-state value in each labor market state, as follows:

$$\begin{aligned} \text{Welfare} &= v \int_{r_2} \int_{a} V(a, r_2) d\Lambda(a|v) dR(r_2) + q \int_{r_2} \int_{y} \int_{a} Q(a, r_2, y) d\Lambda(a|q) dG(y|y^+) dR(r_2) \\ &+ \sum_{f=0,1} e(f) \int_{r_2} \int_{w} \int_{a} W(a, r_2, w, f) d\Lambda(a|f) dF(w|f, w^+) dR(r_2) \end{aligned} \tag{17}$$

where all definitions are provided in Section 2. The only newly introduced notations are w^+ and y^+, which represent the accepted wages and self-employment income, respectively. Finally, we further assess the impact of policy experiments on cross-sectional wealth and consumption inequality, and on *lifetime* consumption inequality. We define lifetime values as the present discounted value of a 65-year long simulated labor market career. To give a more complete picture of inequality, we compute three indices in the class of the Generalized Entropy inequality indices. They are defined as:

$$GE(\alpha) = \begin{cases} \frac{1}{\alpha(\alpha-1)} \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{y_i}{\bar{y}} \right)^{\alpha} - 1 \right] & \alpha \neq 0, 1 \\ \frac{1}{N} \sum_{i=1}^{N} \frac{y_i}{\bar{y}} \ln \left(\frac{y_i}{\bar{y}} \right) & \alpha = 1 \\ -\frac{1}{N} \sum_{i=1}^{N} \ln \left(\frac{y_i}{\bar{y}} \right) & \alpha = 0 \end{cases} \tag{18}$$

where N is the number of individuals, y_i is the measure of assets or consumption for individual i and α is a parameter that weights the distance between measurement variable along the distribution. The larger the parameter α, the greater is the weight of the differences among the rich. The three values of the parameter α we report correspond to the following statistics: $GE(0)$ is the mean log deviation, $GE(1)$ is the Theil index, and $GE(2)$ is half the coefficient of variation.

Table 11 reports a set of descriptive statistics for the benchmark model and the counterfactual simulations on labor market outcomes. Table 7 presents descriptive statistics on the stocks of wealth ($a > 0$) and debt ($a < 0$), as well as on the portfolio allocations, for both the benchmark model and the counterfactual simulations. Table 9 reports descriptive statistics on flows, including

[56] See for example, Tejada (2016); Flabbi (2010); Flinn (2002).

net savings, savings, and dissavings. Finally, Tables 10 and **??** report the welfare measure and the inequality indices defined in equations (17) and (18), respectively. In all tables, the first column reports the benchmark values, while the other six columns report the policy experiments' values. For each experiment, the first column reports the value, and the second column reports the ratio between the value in the experiment and the benchmark value.

6.1 Financial inclusion

When the portfolio costs are equal for formal and informal workers overall net saving increase by 17% (Columns 2 and 3 in Table 9). Behind this large increase, there are important composition effects. As shown in Columns 2 and 3 in Table 7, both wealth and debt increase. Importantly, the proportion of total assets held by informal workers in formal financial institution increases significantly, more than doubling the benchmark value. Overall, the proportion of wealth held in formal assets increase by 38%.

Saving and borrowing behavior has also an impact on labor market decisions since it affects outside options and reservations wages. Full financial inclusion gives an advantage to workers, allowing them to smooth consumption better, to improve the optimal allocation due to precautionary motives and, ultimately, to finance labor market search more. As a result, workers are pickier in accepting a job and spend more time in unemployment: as shown in Table 11 the unemployment rate increase by 4.6%. However, the impact on informality is negligible, with the same proportion of workers in the state as in the benchmark case (there is just a tiny redistribution from self-employed to informal employee). More important is the impact on inequality, reporting a strong decrease in financial assets inequality (about 20%) and a strong increase in lifetime consumption inequality (more than 40%) (Table **??**).

6.2 Lower proportion of informal job offers

The reduction in the proportion of informal wage offers $p(0)$ needed to match the increase in average wealth in the informal sector is quite large: about seventeen percentage points, from 54.5% to 37.6%. This experiment is interesting to contrast with the previous one because it shows how to obtain the same wealth increase for informal workers but using a labor market intervention instead of a financial market intervention (Table 11). Unlike the previous policy, labor market informality now decreases, leading to a 4% increase in the proportion of formal employees. But this effect comes at the cost of the saving increase: under this new policy, net savings actually decrease by more than 15% (Table 9). As in the previous case, consumption inequality and asset inequality increase (Table **??**).

6.3 Payroll tax policy

The reduction in τ needed to match the increase in average wealth in the informal sector is less than two percentage points, from 16% at baseline to 14.5% in the experiment. Again, it is interesting to contrast this experiment with the full financial exclusion one to see how a fiscal policy could achieve the same increase in informal workers wealth. As in the first experiment, net savings increase but by a smaller amount, 10% instead of 17% (Table 9). The labor market outcomes are quite different, with an increase in the proportion of agents in all labor market states with the exception of self-employment (Table 11). As in the first and second experiment, consumption inequality and asset inequality increase, but by a slightly larger amount (Table **??**).

7 Conclusions (Preliminary)

We develop and estimate a model able to replicate the crucial features of developing countries economies: the high level of labor market informality, the low level of savings and the high proportion of assets held in informal institutions. We accomplish this result by building a model of labor market search where agents can work formally and informally and can save and borrow in formal and informal financial institutions.

We estimate the model on data for Colombia, a large LAC country with high informality and low saving. The estimates confirm the claim that working informally is linked to saving informally: informal workers face partial financial exclusion from formal financial institutions, as expressed by the higher portfolio costs they face. If full financial access were guaranteed to them, both saving and borrowing would increase, leading to a 17% increase in monthly net savings. The portfolio allocation would also significantly change, doubling the proportion of assets held in formal financial institution by informal workers. In the labor market, the policy does NOT lead to an increase in the proportion of informal workers despite removing one of the disadvantage of the state (partial financial exclusion). What happens, instead, is that workers are generally pickier in accepting a job, leading to a 4.6% increase in the unemployment rate. We contrast this policy with two other policies that use different policy levers to achieve the same increase in average wealth in the informal sector. One is a pure labor market policy (reduce the proportion of informal employee offers); the other is a simple fiscal policy (reduce the payroll contribution for formal workers). Two main results emerge. First, that the labor market policy necessary to obtain the same result is quite large: about seventeen percentage points, from 54.5% to 37.6%. Second, the impact on labor market and inequality of these two alternative policies is quite different in terms of labor market effects and inequality.

Our paper also provides two important methodological contributions to the empirical literature on labor market search. First, we extend the literature on search and saving[57] by adding a portfolio decision over a riskless and a risky asset. Second, we extend the literature on search and informality[58] by adding saving and borrowing. Both extensions are essential to study labor markets where agents works and save informally. But they are also useful in more general contexts by recognizing the important interactions and feedback loops between labor market and financial market.

[57]See Rendon (2006); Lentz (2009); Lise (2013); García-Pérez and Rendon (2020); Abrahams (2022); Danforth (1979); Acemoglu and Shimer (1999); Browning et al. (2007); Krusell et al. (2010); Bils et al. (2011); Ji (2021); Setty and Yedid-Levi (2021); Pizzo (2022).

[58]See Albrecht et al. (2009); Charlot et al. (2013); Bosch and Esteban-Pretel (2012); Meghir et al. (2015); Bobba et al. (2021, 2022).

References

Abrahams, Scott David, "Essays in Labor Economics." PhD dissertation, Duke University 2022.

Acemoglu, Daron and Robert Shimer, "Efficient unemployment insurance," *Journal of political Economy*, 1999, *107* (5), 893–928.

Achdou, Yves, Francisco J Buera, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "PDE Models in Macroeconomics," *Philosophical Transactions of The Royal Society A*, 2014, *372* (20130397), 15.

_ , Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *NBER Working Paper Series*, 2017, *23732*, 1–58.

Adentiloye, K.A., "Problems and Prospects of the Informal Financial Market in Nigeria," *Journal of Economics and Financial Studies*, 2006, *N.A.*, 81–95.

Aiyagari, S Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Albrecht, James, Lucas Navarro, and Susan Vroman, "The Effects of Labour Market Policies in an Economy with an Informal Sector," *Economic Journal*, 2009, *119* (2000), 1105–1129.

Barles, G. and P.E. Souganidis, "Convergence of Approximation Schemes for Fully Nonlinear Second Order Equations," *Asymptotic Analysis*, 1991, *4* (3), 271–283.

Batini, N., P. Levine, B. Kim Y., and E. Lotti, "Informal labour and credit markets: a survey," *IMF Working Paper*, 2010, *10/42.*

Bils, Mark, Yongsung Chang, and Sun-Bin Kim, "Worker heterogeneity and endogenous separations in a matching model of unemployment fluctuations," *American Economic Journal: Macroeconomics*, 2011, *3* (1), 128–54.

Bobba, Matteo, Luca Flabbi, and Santiago Levy, "Labor Market Search, Informality, and Schooling Investments," *International Economic Review*, 2022, *63* (1), 211–261.

_ , _ , _ , and Mauricio Tejada, "Labor Market Search, Informality, and on-the-Job Human Capital Accumulation," *Journal of Econometrics*, August 2021, *223* (2), 433–453.

Bonaparte, Yosef and Russell Cooper, "Costly Portfolio Adjustment," *NBER Working Paper Series*, 2009, *15227*, 1–25.

_ , _ , **and Guozhong Zhu**, "Consumption Smoothing and Portfolio Rebalancing: The Effects of Adjustment Costs," *Journal of Monetary Economics*, December 2012, *59* (8), 751–768.

Bond, Eric, James R. Tybout, and Hale Utar, "Credit Rationing, Risk Aversion and Industrial Evolution in Developing Countries," *NBER Working Paper Series*, 2008, *14116*, 1–53.

Bond, Eric W, James Tybout, and Hale Utar, "Credit rationing, risk aversion, and industrial evolution in developing countries," *International Economic Review*, 2015, *56* (3), 695–722.

Bosch, Mariano and Julen Esteban-Pretel, "Job Creation and Job Destruction in the Presence of Informal Markets," *Journal of Development Economics*, July 2012, *98* (2), 270–286.

_ **and William F Maloney**, "Comparative analysis of labor market dynamics using Markov processes: An application to informality," *Labour Economics*, 2010, *17* (4), 621–631.

Browning, Martin, Thomas F Crossley, and Eric Smith, "Asset accumulation and short-term employment," *Review of Economic Dynamics*, 2007, *10* (3), 400–423.

Cavallo, Eduardo, Tomás Serebrisky et al., *Saving for development: How Latin America and the Caribbean can save more and better*, Springer Nature, 2016.

Charlot, Olivier, Franck Malherbet, and Mustafa Ulus, "Efficiency in a search and matching economy with a competitive informal sector," *Economics Letters*, 2013, *118* (1), 192 – 194.

Chetty, Raj, "A new method of estimating risk aversion," *American Economic Review*, 2006, *96* (5), 1821–1834.

Conti, Gabriella, Rita Ginja, and Renata Narita, "The Value of Health Insurance: A Household Job Search Approach," *HCEO Working Paper Series*, 2018, *2018–050*, 1–66.

Dabla-Norris, E. and J. Koeda, "Informality and bank credit: evidence from firm-level data," *IMF Working Paper*, 2008, *08/94.*

Danforth, John P, "On the role of consumption and decreasing absolute risk aversion in the theory of job search," *Studies in the Economics of Search*, 1979, *123*, 109–131.

de Paula, Aureo and Jose A. Scheinkman, "Value-Added Taxes, Chain Effects, and Informality," *American Economic Journal: Macroeconomics*, October 2010, *2* (4), 195–221.

Dey, Matthew and Christopher Flinn, "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica*, 2005, *73* (2), 571–627.

— **and** — , "Household Search and Health Insurance Coverage," *Journal of Econometrics*, 2008, *145* (1-2), 43–63.

— **and** — , "Household search and health insurance coverage," *Journal of Econometrics*, 2008, *145* (1-2), 43–63.

Dizon, Felipe, Erick Gong, and Kelly Jones, "The effect of promoting savings on informal risk sharing experimental evidence from vulnerable women in kenya," *Journal of Human Resources*, 2020, *55* (3), 963–998.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman, "Labor market dynamics and development," *The Quarterly Journal of Economics*, 2023, *138* (4), 2287–2325.

Dupas, P. and J. Robinson, "Why don't the poor save more? evidence from health savings experiments," *American Economic Review*, 2013, *103* (4), 1138–71.

Eckstein, Zvi and Gerard van den Berg, "Empirical Labor Search: A Survey," *Journal of Econometrics*, February 2007, *136* (2), 531–564.

— **, Suqin Ge, and Barbara Petrongolo**, "Job and wage mobility with minimum wages and imperfect compliance," *Journal of Applied Econometrics*, 2011, *26* (4), 580–612.

Eeckhout, Jan and Kaivan Munshi, "Matching in Informal Financial Institutions," *Journal of the European Economic Association*, September 2010, *8* (5), 947–988.

Esteban-Pretel, Julen and Sagiri Kitao, "Labor Market Policies in a Dual Economy," *Labour Economics*, January 2021, *68*, 101956.

Evans, David J, "The elasticity of marginal utility of consumption: estimates for 20 OECD countries," *Fiscal studies*, 2005, *26* (2), 197–224.

Finamor, Lucas, "Labor market informality, risk, and insurance," *mimeo*, 2024, *Sao Paulo School of Economics, FGV.*

Flabbi, Luca, "Gender discrimination estimation in a search model with matching and bargaining," *International Economic Review*, 2010, *51* (3), 745–783.

_ **and James Mabli**, "Household search or individual search: Does it matter?," *Journal of Labor Economics*, 2018, *36* (1), 1–46.

_ **and Mauricio M. Tejada**, "Are informal self-employment and informal employment as employee behaviorally distinct labor force states?," *Economics Letters*, 2023, *231*, 111278.

Flinn, Christopher and Joseph Mullins, "Labor market search and schooling investment," *International Economic Review*, 2015, *56* (2), 359–398.

Flinn, Christopher J, "Labour Market Structure and Inequality: A Comparison of Italy and the US," *Review of Economic Studies*, August 2002, *69* (3), 611–645.

_ , "Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates," *Econometrica*, 2006, *74* (4), 1013–1062.

_ **and James J. Heckman**, "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, 1982, *18* (June 1981), 115–168.

Flórez, Luz A., "Informal Sector Under Saving: A Positive Analysis of Labour Market Policies," *Labour Economics*, 2017, *44*, 13–26.

García-Pérez, J Ignacio and Sílvio Rendon, "Family job search and wealth: The added worker effect revisited," *Quantitative Economics*, 2020, *11* (4), 1431–1459.

Gasparini, Leonardo and Leopoldo Tornarolli, "Labor informality in Latin America and the Caribbean: patterns and trends from household microdata survey," *Desarrollo y Sociedad*, 2009, *63*, 13–80.

Gatti, R. and M. Honorati, "Informality among formal firms: firm-level, cross- country evidence on tax compliance and access to credit," *World Bank Policy Re- search Working Paper*, 2008, *4476.*

Gouriéroux, Christian and alain Monfort, *Simulation-Based Econometric Methods* CORE Lectures, Oxford University Press, 2002.

Granda, Catalina and Franz Hamann, "On the Aggregate Implications of Removing Barriers to Formality," *Federal Reserve Bank of St. Louis Review*, 2020, *Second Quarter*, 203–20.

Guler, Bulent, Fatih Guvenen, and Giovanni L Violante, "Joint-search theory: New opportunities and new frictions," *Journal of Monetary Economics*, 2012, *59* (4), 352–369.

Herreño, Juan and Sergio Ocampo, "The macroeconomic consequences of subsistence self-employment," *Journal of Monetary Economics*, 2023, *136*, 91–106.

Huggett, Mark, "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 1993, *17*, 953–969.

Jessen, Jonas and Jochen Kluve, "The effectiveness of interventions to reduce informality in low- and middle-income countries," *World Development*, 2021, *138*, 105256.

Ji, Yan, "Job search under debt: Aggregate implications of student loans," *Journal of Monetary Economics*, 2021, *117*, 741–759.

Kanbur, Ravi, "Conceptualising Informality: Regulation and Enforcement," *IZA Discussion Papers*, May 2009, *4186*, 1–18.

Kaplan, Greg and Giovanni L Violante, "A model of the consumption response to fiscal stimulus payments," *Econometrica*, 2014, *82* (4), 1199–1239.

Karlan, Dean, Aishwarya Lakshmi Ratan, and Jonathan Zinman, "Savings by and for the Poor: A Research Review and Agenda," *Review of Income and Wealth*, 2014, *60* (1), 36–78.

_ **and Jonathan Morduch**, "Chapter 71 - Access to Finance," in Dani Rodrik and Mark Rosenzweig, eds., *Handbooks in Economics*, Vol. 5 of *Handbook of Development Economics*, Elsevier, 2010, pp. 4703–4784.

Krusell, Per, Toshihiko Mukoyama, and Aysegul Sahin, "Labour-Market Matching with Precautionary Savings and Aggregate Fluctuations," *Review of Economic Studies*, oct 2010, *77* (4), 1477–1507.

Lentz, Rasmus, "Optimal Unemployment Insurance in an Estimated Job Search Model with Savings," *Review of Economic Dynamics*, 2009, *12* (1), 37–57.

Lise, Jeremy, "On-the-Job Search and Precautionary Savings," *Review of Economic Studies*, 2013, *80* (3), 1086–1113.

Lorenzo, F. and R. Osimani, "Alternativas de política para fortalecer el ahorro de los hogares de menores ingresos: el caso de uruguay," *Serie Financiamiento del Desarrollo, Economic Commission for Latin America and the Caribbean*, 2001, *110.*

Maloney, William F, "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico," *World Bank Economic Review*, May 1999, *13* (2), 275–302.

Meghir, Costas, Renata Narita, and Jean-Marc Robin, "Wages and Informality in Developing Countries," *American Economic Review*, April 2015, *105* (4), 1509–1546.

Merton, Robert C, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 1971, *3*, 373–413.

Moore, Mark A., Anthony E. Boardman, and Aidan R. Vining, "Social Discount Rates for Seventeen Latin American Countries: Theory and Parameter Estimation," *Public Finance Review*, January 2020, *48* (1), 43–71.

Narita, Renata, "Self-employment in developing countries: A search-equilibrium approach," *Review of Economic Dynamics*, 2020, *35*, 1–34.

OECD, *Taxing Wages 2023: Indexation of Labour Taxation and Benefits in OECD Countries*, OECD, 2023.

Ogbuabor, J. E., A. Malaolu V., and I. C. Mba Mba, "Informality and domestic savings in nigeria: lessons from time series analysis," *International Journal of Development and Economic Sustainability*, 2013, *1* (3), 24–32.

Perry, Guillermo E., William F. Maloney, Omar S. Arias, Pablo Fajnzylber, Andrew D. Mason, and Jaime Saavedra-Chanduvi, *Informality : Exit and Exclusion* number 6730. In 'World Bank Publications.', The World Bank, 2007.

Pierri, Damian and Enrique Kawamura, "Life cycle, financial frictions and informal labor markets: the case of Chile," *Journal of Applied Economics*, 2022, *25* (1), 93–120.

Pizzo, Alessandra, "The welfare effects of tax progressivity in a frictional labor market," *Review of Economic Dynamics*, 2022.

Porta, Rafael La and Andrei Shleifer, "The Unofficial Economy and Economic Development," *Brookings Papers on Economic Activity*, 2008, *2008*, 275–352.

_ **and** _ , "Informality and Development," *Journal of Economic Perspectives*, September 2014, *28* (3), 109–26.

Reinhardt, Nola, "The Conundrum of Low Saving Rates in Latin America," *Development Policy Review*, 2008, *26* (6), 727–744.

Rendon, Sílvio, "Job Search and Asset Accumulation under Borrowing Constraints," *International Economic Review*, 2006, *47* (1), 233–263.

Schclarek, Alfredo and Mauricio Caggia, "Household Saving and Labor Informality : The Case of Chile," *IDB Working Papers*, 2015, *IDB-WP-581*, 1–48.

Schneider, Friedrich, Andreas Buehn, and Claudio E. Montenegro, "New estimates for the shadow economies all over the world," *International Economic Journal*, 2010, *24* (4), 443–461.

Schoar, Antoinette, "The divide between subsistence and transformational entrepreneurship," *Innovation policy and the economy*, 2010, *10* (1), 57–81.

Setty, Ofer and Yaniv Yedid-Levi, "On the provision of unemployment insurance when workers are ex-ante heterogeneous," *Journal of the European Economic Association*, 2021, *19* (1), 664–706.

Tejada, Mauricio, Claudia Piras, Luca Flabbi, and Monserrat Bustelo, "Gender Gaps in Latin American Labor Markets: Implications from an Estimated Search Model," *Annals of Economics and Statistics*, 2021, (142), 111–178.

Tejada, Mauricio M, "Lifetime inequality measures for an emerging economy: The case of Chile," *Labour Economics*, 2016, *42*, 1–15.

Ulyssea, Gabriel, "Firms, Informality, and Development: Theory and Evidence from Brazil," *American Economic Review*, August 2018, *108* (8), 2015–2047.

World Bank, *World Development Report 1989: Financial Systems and Development*, Washington, DC: World Bank, 1989.

— , *World Development Report 2013: Jobs*, The World Bank, 2012.

Table 1: Descriptive Statistics on Labor Market Outcomes [GEIH]

	Formal Employees	Informal Employees	Informal Self-Employed	Unemployed
Labor Market States (proportions)				
Proportion	0.395	0.113	0.413	0.079
Labor Income (100's US$ per month)				
Mean	3.169	2.305	2.363	–
Standard Deviation	1.138	0.644	1.081	–
Ongoing Duration (months)				
Mean	65.963	33.84	105.118	4.034
Standard Deviation	77.352	54.784	104.533	6.859
Sample				
Number Obs.	30989.0	8889.0	32469.0	6195.0

Table 2: Descriptive Statistics on Saving Behavior [ELCA]

	Formal Employees	Informal Employees	Informal Self-Employed	Unemployed
Proportion of Individuals who save				
At all	0.271	0.19	0.218	0.036
Mainly in formal institutions	0.493	0.167	0.191	0.333
Savings amount among savers (100's US$ per month)				
Mean	0.601	0.288	0.578	0.443
Standard Deviation	0.721	0.207	0.84	0.48
Saving rate among savers (savings/labor income)				
Mean	0.132	0.097	0.169	–
Standard Deviation	0.123	0.055	0.132	–
Sample				
Number Obs.	517.0	158.0	431.0	83.0

Table 3: Estimated Parameters

Definition	Parameter	Est. Value	Std. Error
Mobility Shocks			
Job offer - formal employees	$\lambda^e(1)$	0.092	(0.0013)
Job offer - informal employees	$\lambda^e(0)$	0.052	(0.0012)
Job offer - self-employed	λ^q	0.112	(0.0051)
Job offer - unemployed	λ^v	2.685	(0.0246)
Income shock - self-employed	π	0.205	(0.0058)
Job separation - formal employees	$\eta(1)$	0.162	(0.0057)
Job separation - informal employees	$\eta(0)$	0.28	(0.007)
Job Offers and Self-Employment Income			
Proportion of formal jobs	$p(1)$	0.305	(0.0025)
Logarithm of location - formal employees	$\mu(1)$	2.814	(0.0086)
Logarithm of scale - formal employees	$\sigma(1)$	0.665	(0.0077)
Logarithm of location - informal employees	$\mu(0)$	2.774	(0.0162)
Logarithm of scale - informal employees	$\sigma(0)$	0.397	(0.002)
Logarithm of location - self-employed	μ_q	3.065	(0.0025)
Logarithm of scale - self-employed	σ_q	0.516	(0.0334)
Portfolio Adjustment			
Cost - formal employees	$\psi^e(1)$	0.011	(0.0003)
Cost - informal employees	$\psi^e(0)$	0.145	(0.007)
Cost - self-employed	ψ^q	0.182	(0.0063)
Cost - unemployed	ψ^v	0.086	(0.0031)
Rate of Return Informal Assets			
Persistence	κ	1.679	(0.011)
Flow Income Unemployed			
Value	b	1.351	(0.0013)
Utility Function			
Risk aversion	δ	-0.458	(0.0008)
Value of Formal Jobs	ϵ	0.013	(0.0007)

NOTE: Asymptotic standard errors in parentheses. Estimates obtained conditioning on the following calibrated parameters: payroll tax rate $\tau = 0.16$; minimum wage $w_M = 27.48$; discount rate $\rho = 0.12$; death rate $\theta = 0.013$; return of formal assets $r_1 = 0.037$; return of informal assets distribution $R(r_2) = \mathcal{N}\left(0.039, 0.015^2\right)$; interest rate spread $\nu = 1.147$.

Table 4: Implied Estimated Values

Definition	Notation	Est. Value
Wages Offers		
Mean - formal employees	$E[w\|1]$	1.734
Std. Dev. - formal employees	$SD[w\|1]$	1.293
Mean - informal employees	$E[w\|0]$	1.445
Std. Dev. - informal employees	$SD[w\|0]$	0.596
Self-employment Income Opportunities		
Mean	$E[y]$	2.041
Std. Dev.	$SD[y]$	1.127

NOTE: Values implied by point estimates reported in Table 3. Variables are in hundreds of US dollar per month.

Table 5: Moments Fit

Statistic	Data	Model	Statistic	Data	Model
$e(1)$	0.395	0.345	$E[I_{s>0} \times s\|e(1)]$	0.163	0.177
$e(0)$	0.113	0.142	$SD[I_{s>0} \times s\|e(1)]$	0.46	0.288
q	0.413	0.456	$E[I_{s>0} \times s\|e(0)]$	0.055	0.084
v	0.079	0.057	$SD[I_{s>0} \times s\|e(0)]$	0.144	0.16
$E[w(1)]$	3.169	3.425	$E[I_{s>0} \times s\|q]$	0.126	0.142
$SD[w(1)]$	1.138	0.92	$SD[I_{s>0} \times s\|q]$	0.458	0.268
$E[w(0)]$	2.305	2.089	$\Pr[\phi > 0.5\|e(1), w \leq \tilde{w}_1]$	0.4	0.415
$SD[w(0)]$	0.644	0.588	$\Pr[\phi > 0.5\|e(1), w > \tilde{w}_1]$	0.525	0.438
$E[w_q]$	2.363	2.688	$\Pr[\phi > 0.5\|e(0), w \leq \tilde{w}_0]$	0.091	0.081
$SD[w_q]$	1.081	0.922	$\Pr[\phi > 0.5\|e(0), w > \tilde{w}_0]$	0.222	0.273
$P5[w(1)]$	2.291	2.363	$\Pr[\phi > 0.5\|q, y \leq \tilde{y}]$	0.054	0.13
$P5[w(0)]$	1.163	1.332	$\Pr[\phi > 0.5\|q, y > \tilde{y}]$	0.278	0.328
$P5[w_q]$	0.831	1.607	$\Pr[\phi > 0.5]$	0.348	0.306
$E[t/120\|e(1)]$	0.55	0.522	$E[I_{s>0} \times s]$	0.125	0.138
$E[t/120\|e(0)]$	0.282	0.315	$SD[I_{s>0} \times s]$	0.417	0.259
$E[t/120\|q]$	0.876	0.729	$\Pr[\phi > 0.5\|w \leq \tilde{w}]$	0.216	0.214
$E[t/12\|v]$	0.336	0.538	$\Pr[\phi > 0.5\|w > \tilde{w}]$	0.415	0.34
$\Pr[\phi > 0.5\|e(1)]$	0.493	0.427	$Corr(s, w)\|s > 0]$	0.585	0.624
$\Pr[\phi > 0.5\|e(0)]$	0.167	0.205			
$\Pr[\phi > 0.5\|q]$	0.191	0.236			

NOTE: $s = da/dt$ is the amount saved; E denotes average, SD standard deviation, $P5$ 5th bottom percentile, $\Pr$ probability; $I_{s>0}$ is an indicator variable equal 1 if the individual saves a positive amount, and $\tilde{x}$ denotes the median of the income variable x.

Table 6: Full Financial Inclusion Experiment - Labor Market Outcomes

	Benchmark	Counterfactual	Ratio
Labor market states (proportion)			
$e(1)$	0.346	0.345	1.000
$e(0)$	0.144	0.145	1.010
q	0.453	0.451	0.996
v	0.058	0.058	1.009
Wages (hundred of US$ per month)			
$E[w\|e(1)]$	3.416	3.418	1.001
$E[w\|e(0)]$	2.089	2.097	1.004
$E[w\|q]$	2.702	2.692	0.996
Ongoing duration (months)			
$E[t\|e(1)]$	62.618	62.408	0.997
$E[t\|e(0)]$	37.508	36.977	0.986
$E[t\|q]$	85.696	85.507	0.998
$E[t\|v]$	6.451	6.4000	0.992

Note: The experiment consists in setting $\psi^e(0) = \psi^q = \psi^v = 0.011$. Benchmark's values are: $\psi^e(0) = 0.145$; $\psi^e(1) = 0.011$; $\psi^q = 0.182$; $\psi^v = 0.086$. $s = da/dt$ is the amount saved. Results are based on simulations of 50.000 individuals. The column Ratio presents the ratio with respect to the benchmark.

Table 7: Full Financial Inclusion Experiment - Financial Outcomes (Stocks)

	Wealth ($a > 0$)			Debt ($a < 0$)		
	Benchmark	Counterfactual	Ratio	Benchmark	Counterfactual	Ratio
			Assets (hundred of US$)			
$E[a]$	18.453	18.613	1.009	-2.238	-2.299	1.027
$E[a\|e(1)]$	19.878	20.157	1.014	-1.985	-1.823	0.919
$E[a\|e(0)]$	11.302	11.814	1.045	-2.457	-2.611	1.063
$E[a\|q]$	19.211	19.220	1.000	-1.854	-1.941	1.047
$E[a\|v]$	17.427	17.531	1.006	-2.733	-2.81	1.028
			Share of Individuals (proportion)			
Total	0.904	0.896	0.990	0.095	0.103	1.093
$e(1)$	0.333	0.332	0.998	0.010	0.010	1.012
$e(0)$	0.105	0.105	1.004	0.038	0.039	1.035
q	0.422	0.413	0.978	0.033	0.041	1.229
v	0.045	0.046	1.017	0.014	0.014	0.984

Note: The experiment consists in setting $\psi^e(0) = \psi^q = \psi^v = 0.011$. Benchmark's values are: $\psi^e(0) = 0.145$; $\psi^e(1) = 0.011$; $\psi^q = 0.182$; $\psi^v = 0.086$. Results are based on simulations of 50.000 individuals. The column Ratio presents the ratio with respect to the benchmark.

Table 8: Full Financial Inclusion Experiment - Financial Outcomes (Portfolio Allocation)

	Wealth ($a > 0$)			Debt ($a < 0$)		
	Benchmark	Counterfactual	Ratio	Benchmark	Counterfactual	Ratio
Portfolio allocation os Assets (proportion of assets)						
$E[\phi$	0.313	0.417	1.334	0.223	0.467	2.092
$E[\phi\|e(1)]$	0.431	0.424	0.983	0.483	0.446	0.923
$E[\phi\|e(0)]$	0.210	0.403	1.924	0.207	0.491	2.373
$E[\phi\|q]$	0.245	0.416	1.696	0.122	0.434	3.547
$E[\phi\|v]$	0.309	0.410	1.325	0.327	0.506	1.549
Share of Individual with the Majority of Assets in Formal Institutions (proportion)						
$E[\phi > 0.5]$	0.273	0.372	1.366	0.018	0.048	2.628
$E[\phi > 0.5\|e(1)]$	0.144	0.141	0.980	0.005	0.004	0.927
$E[\phi > 0.5\|e(0)]$	0.020	0.042	2.151	0.007	0.019	2.920
$E[\phi > 0.5\|q]$	0.096	0.171	1.780	0.003	0.017	6.698
$E[\phi > 0.5\|v]$	0.013	0.019	1.386	0.004	0.007	1.594

NOTE: The experiment consists in setting $\psi^e(0) = \psi^q = \psi^v = 0.011$. Benchmark's values are: $\psi^e(0) = 0.145$; $\psi^e(1) = 0.011$; $\psi^q = 0.182$; $\psi^v = 0.086$. Results are based on simulations of 50.000 individuals. The column Ratio presents the ratio with respect to the benchmark.

Table 9: Full Financial Inclusion Experiment - Financial Outcomes (Flows)

	Savings ($s > 0$)			Disavings ($s < 0$)		
	Benchmark	Counterfactual	Ratio	Benchmark	Counterfactual	Ratio
Savings (hundred of US$ per month)						
$E[s]$	0.262	0.261	0.998	-0.261	-0.263	1.009
$E[s\|e(1)]$	0.326	0.321	0.986	-0.115	-0.119	1.027
$E[s\|e(0)]$	0.142	0.144	1.018	-0.100	-0.103	1.032
$E[s\|q]$	0.255	0.257	1.007	-0.133	-0.130	0.975
$E[s\|v]$	–	–	–	-1.207	-1.218	1.010
Share of Individuals who Save (proportion)						
$Total$	0.536	0.535	0.998	0.464	0.465	1.002
$e(1)$	0.188	0.188	1.001	0.154	0.154	0.995
$e(0)$	0.087	0.088	1.011	0.056	0.057	1.014
q	0.262	0.260	0.991	0.194	0.194	1.003
v	–	–	–	0.059	0.060	1.008

NOTE: The experiment consists in setting $\psi^e(0) = \psi^q = \psi^v = 0.011$. Benchmark's values are: $\psi^e(0) = 0.145$; $\psi^e(1) = 0.011$; $\psi^q = 0.182$; $\psi^v = 0.086$. Results are based on simulations of 50.000 individuals. The column Ratio presents the ratio with respect to the benchmark.

Table 10: Full Financial Inclusion Experiment - Welfare Measures

	Benchmark	Counterfactual	Ratio
	Total		
Aggregated Welfare	-0.221	-0.223	1.007
	By labor market state		
Formal Employees	-1.172	-1.171	0.999
Informal Employees	-0.535	-0.540	1.009
Self-employed	-1.591	-1.586	0.997
Unemployed	-0.221	-0.223	1.007
	By whether the individual accumulated wealth or debt		
Accumulated wealth	-3.483	-3.479	0.999
Accumulated debt	-3.842	-3.842	1.000

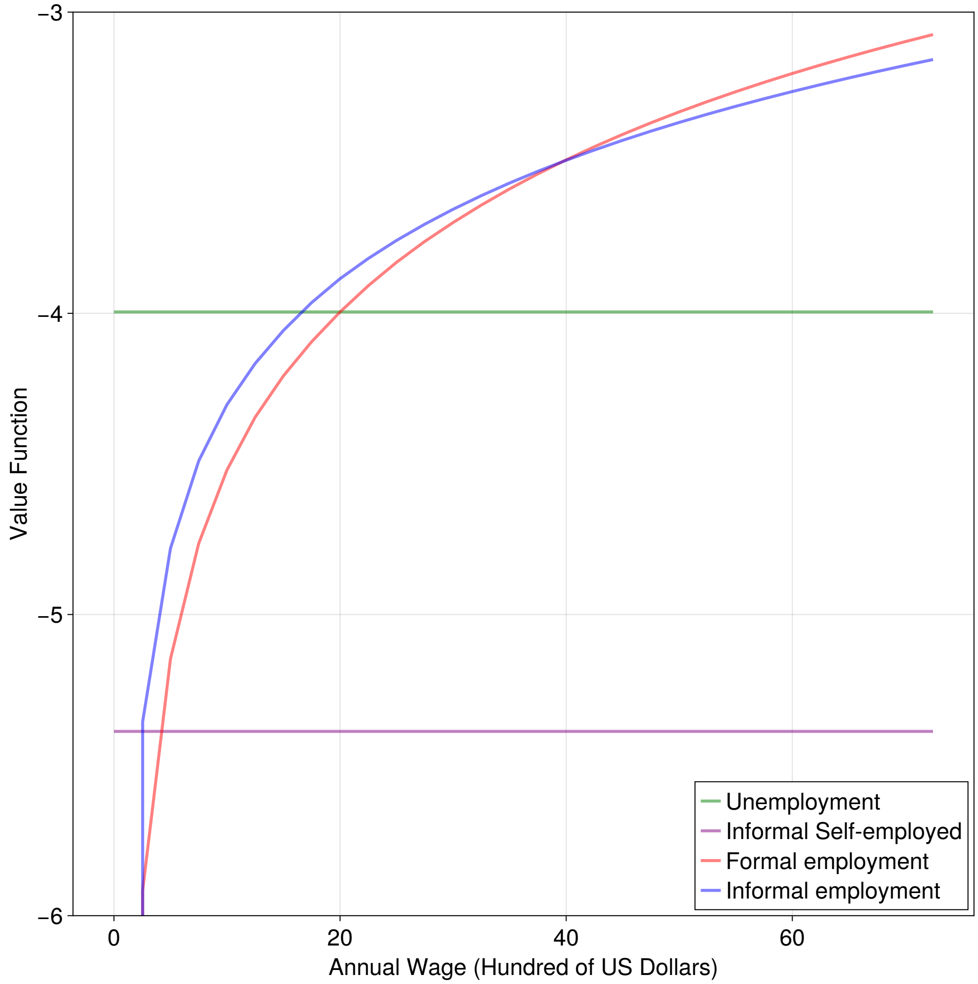
NOTE: The experiment consists in setting $\psi^e(0) = \psi^q = \psi^v = 0.011$. Benchmark's values are: $\psi^e(0) = 0.145$; $\psi^e(1) = 0.011$; $\psi^q = 0.182$; $\psi^v = 0.086$; $p(1) = 0.305$; and $\tau = 0.160$. $s = da/dt$ is the amount saved. Results are based on simulations of 50.000 individuals. The column Ratio presents the ratio with respect to the benchmark.

Table 11: Full Financial Inclusion Experiment - Inequality

	Benchmark	Counterfactual	Ratio
Total Assets			
Mean Log Deviation $GE(0)$	0.562	0.533	0.948
Theil Index $GE(1)$	0.415	0.402	0.969
Squared Coefficient of Variation $GE(2)$	0.628	0.624	0.993
Financial Assets			
Mean Log Deviation $GE(0)$	1.219	0.665	0.546
Theil Index $GE(1)$	0.597	0.427	0.715
Squared Coefficient of Variation $GE(2)$	2.517	2.086	0.829
Cross-section Consumption			
Mean Log Deviation $GE(0)$	0.058	0.058	0.997
Theil Index $GE(1)$	0.053	0.052	0.998
Squared Coefficient of Variation $GE(2)$	0.052	0.052	0.997

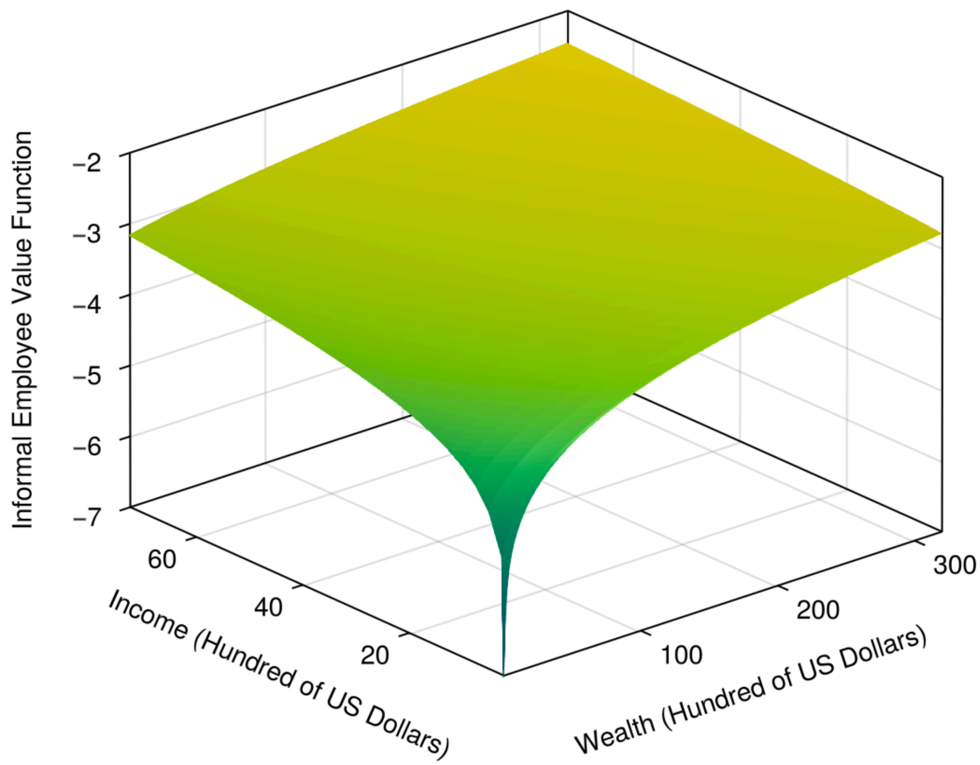
NOTE: The experiment consists in setting $\psi^e(0) = \psi^q = \psi^v = 0.011$. Benchmark's values are: $\psi^e(0) = 0.145$; $\psi^e(1) = 0.011$; $\psi^q = 0.182$; $\psi^v = 0.086$. $s = da/dt$ is the amount saved. Results are based on simulations of 50.000 individuals. The column Ratio presents the ratio with respect to the benchmark.

Figure 1: Value Functions and Wages



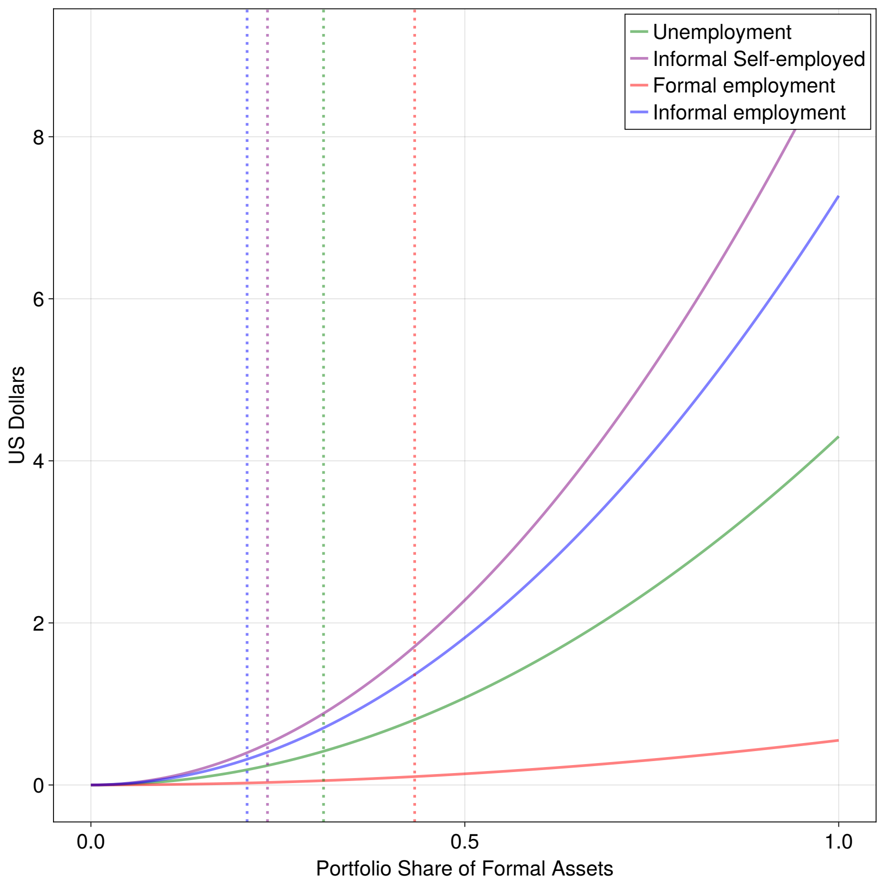
NOTE: Figure reports value functions and offered wages computed at the point estimates discussed in Section 5.2, conditioning on zero assets ($a = 0$), and average informal rate ($r_2 = E(r_2)$). The self-employed value function is computed at the average self-employed income among the self-employed ($y = E(y|i = q)$).

Figure 2: Value Functions, Wages and Assets



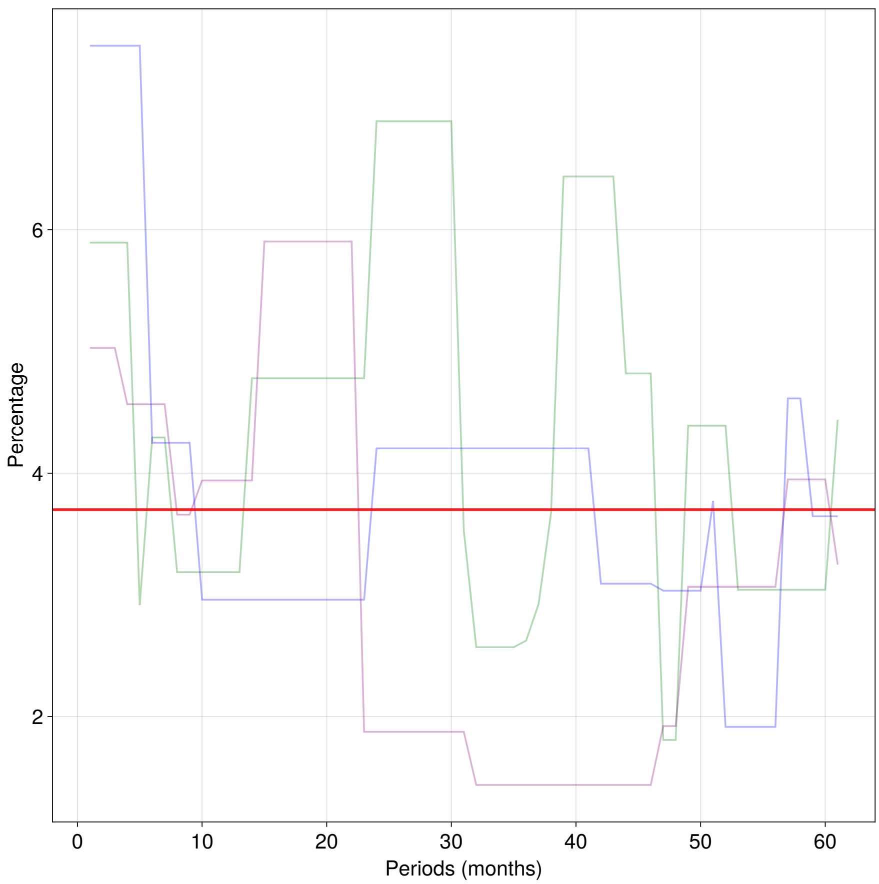
NOTE: Figure reports value functions and offered wages computed at the point estimates discussed in Section 5.2, conditioning on the average informal rate ($r_2 = E(r_2)$).

Figure 3: Portfolio Costs



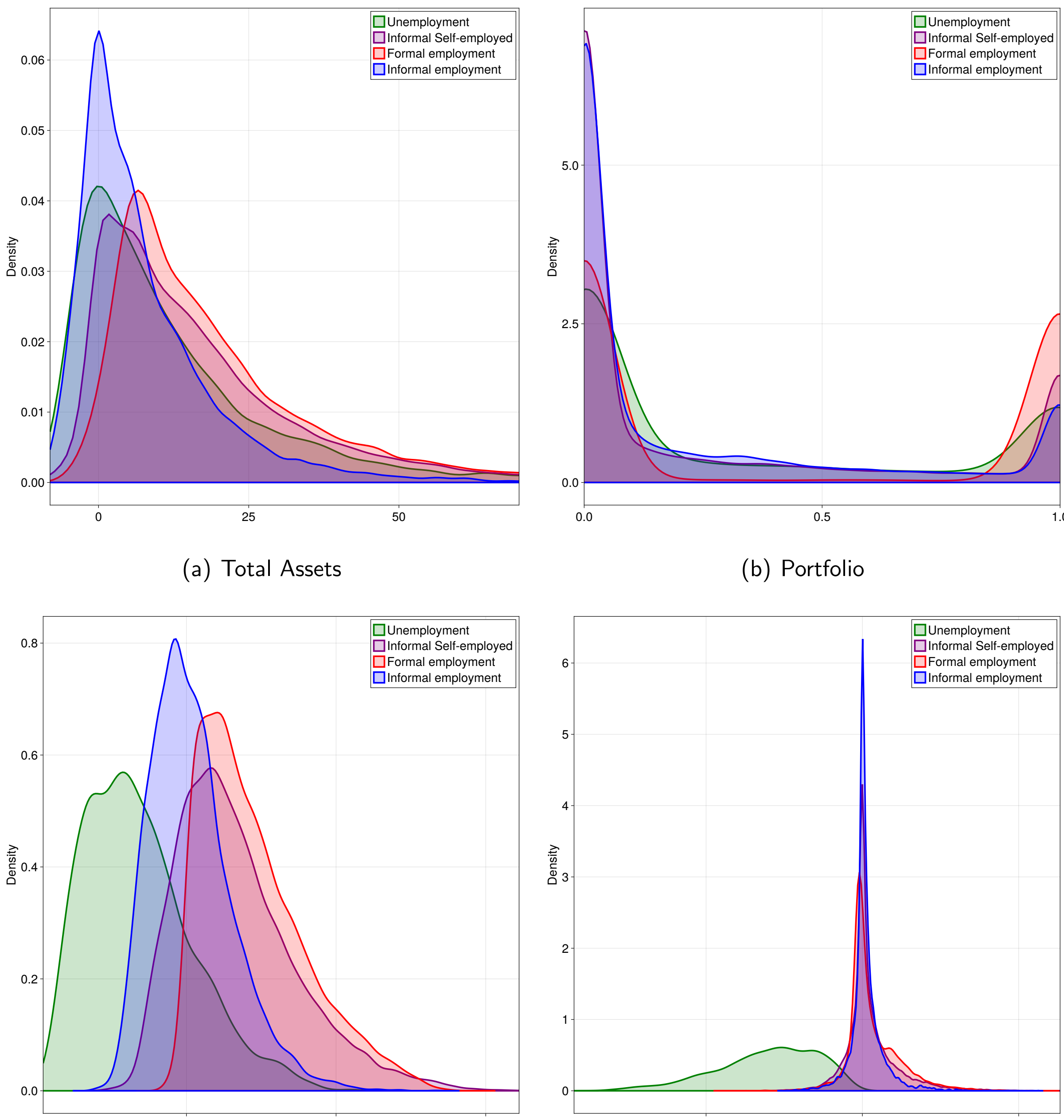
NOTE: Dotted line show the average simulated portfolio allocation by labor market state. Simulated samples of 10,000 individual-level observations based on the estimates reported in Table 3.

Figure 4: Assets Returns



NOTE: Red line show the rate of return of formal (riskless) assets. Colored lines show simulated samples of 3 individual-level rate of returns of informal assets based on the estimates reported in Table 3.

Figure 5: Steady State Distributions of Assets, Consumption and Savings

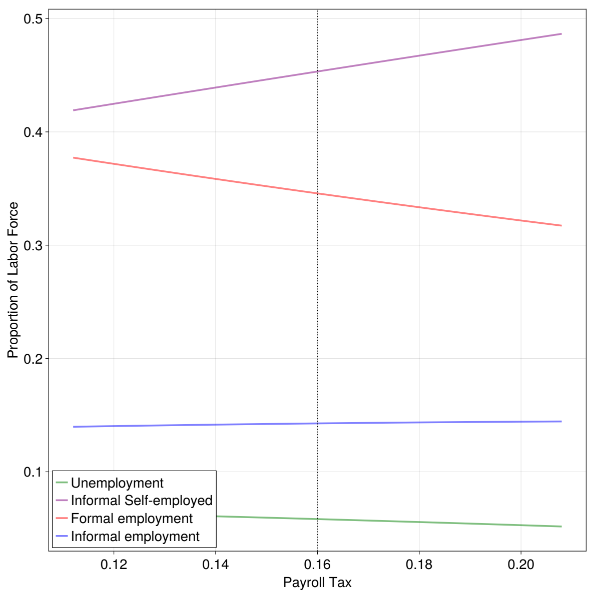


(a) Total Assets (b) Portfolio

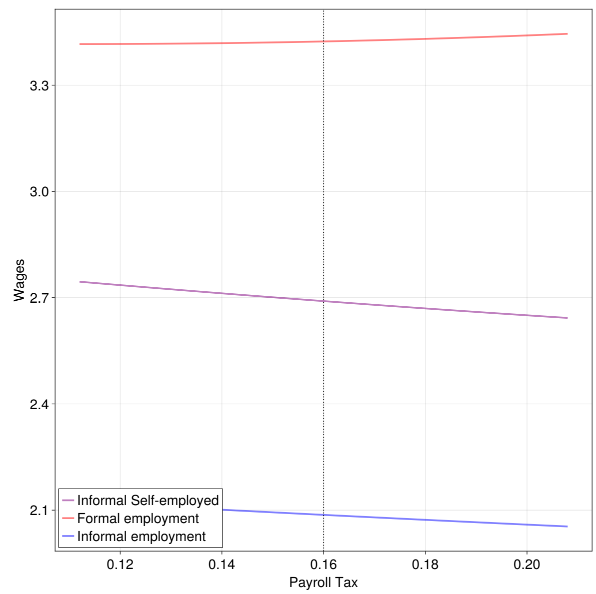
(c) Consumption (d) Savings

Note: Simulated samples of 10,000 individual-level observations based on the estimates reported in Table 3.

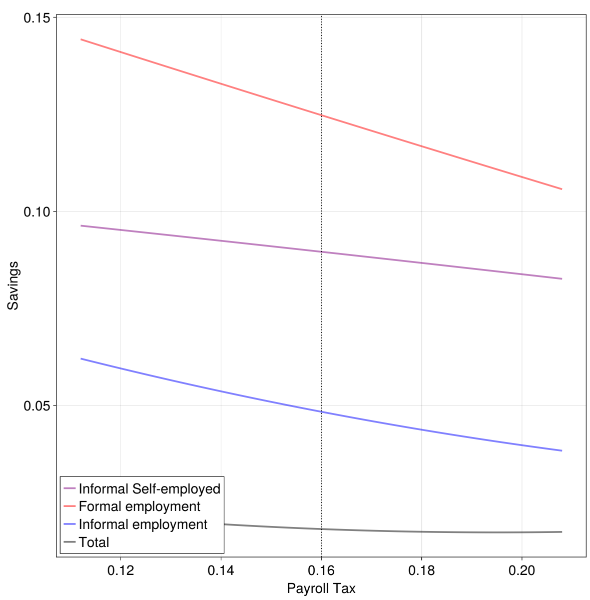
Figure 6: Effects of Payroll Tax Policy



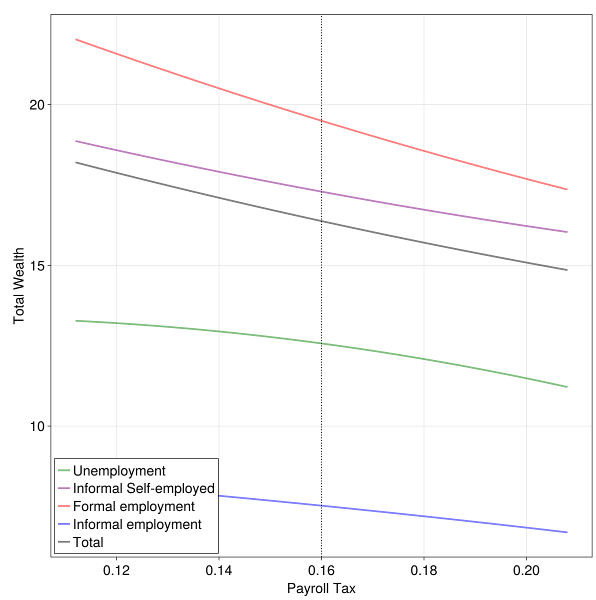
(a) Labor Market States



(b) Wages



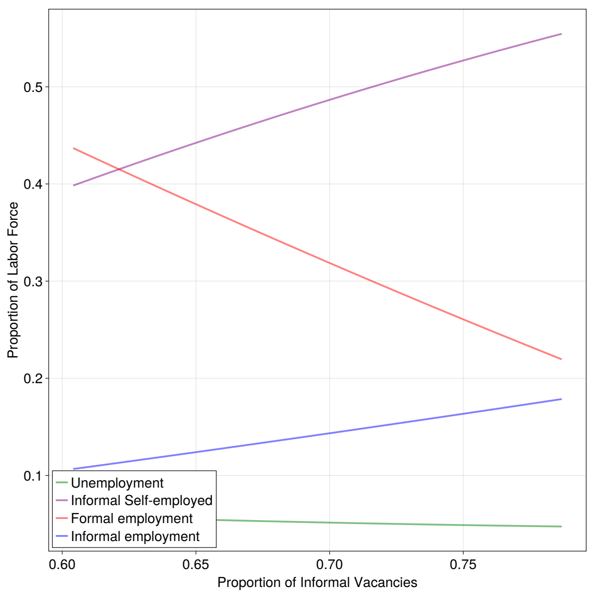
(c) Savings



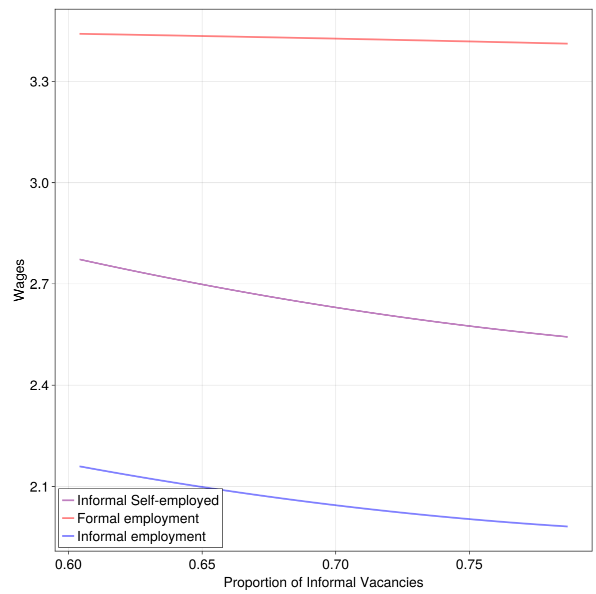
(d) Net Wealth

NOTE: Simulated samples of 10,000 individual-level observations based on the estimates reported in Table 3.

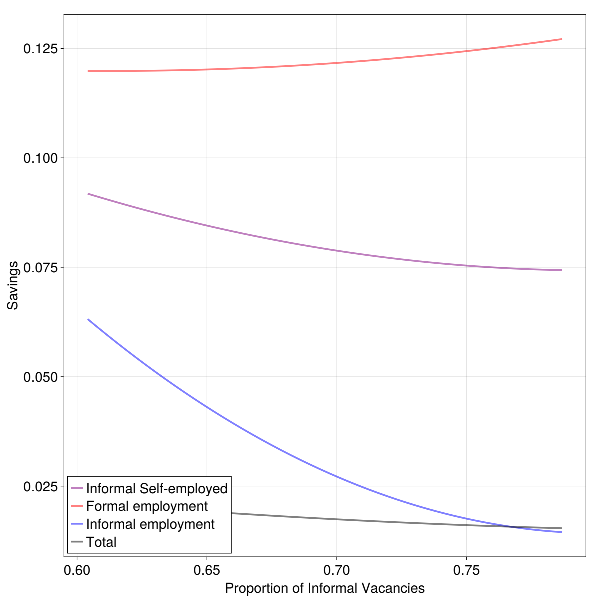
Figure 7: Effects of Restricting Informal Vacancy Creation



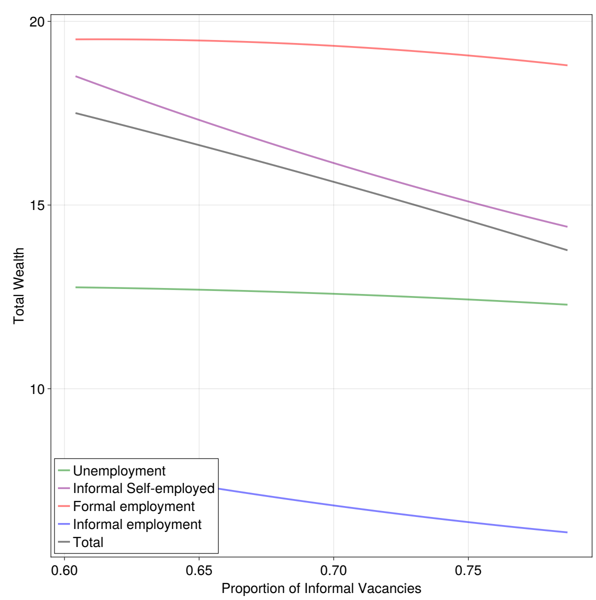
(a) Labor Market States



(b) Wages



(c) Savings



(d) Net Wealth

NOTE: Simulated samples of 10,000 individual-level observations based on the estimates reported in Table 3.

A Model Appendix

A.1 Derivation of the Hamilton-Jacobi-Bellman Equations

The individual problem is:

$$
\begin{aligned}
&\max \;\; E_0 \int_0^{\infty} e^{-\tilde{\rho} t} \left[u(c - \underline{c}) + \epsilon f \right] \\
&\quad s.to \\
&\qquad da = \left[(r_1 \phi + r_2 (1-\phi))(1 + \nu I_{a^-}) a + \iota(j) - c - \Psi(\phi, j) \right] dt \\
&\qquad a \geq \underline{a}
\end{aligned}
$$

where $j = v, q, e(f)$; $f = 0, 1$; I_{a^-} is an indicator variable that takes value 1 if $a < 0$ (debt) and 0 otherwise; $\iota(j)$ is the labor market state specific income:

$$
\iota(j) = \begin{cases} b & \text{if } j = v \\ y & \text{if } j = q \\ w(f)(1 - \tau f) & \text{if } j = e(f); \end{cases}
$$

and τ is the pay-roll contribution paid by workers who are employed formally. The borrowing limit is defined as $\underline{a} = -\frac{b}{\bar{\bar{r}}_2 (1+\nu)}$.

The approach we use to derive the Hamilton-Jacobi-Bellman equations of the model is to start with the discrete time version of model, where the length of a time period is Δt, and then take the limit where Δt goes to zero to find the continuous time counterpart. We start with the case of the employees.

The discrete time version of the value of employment $W(a, r_2, w, f)$, where the length of a time period is Δt, is:

$$
\begin{aligned}
W(a, r_2, w, f) = \max_{0 \leq c \leq \bar{c}, 0 \leq \phi \leq 1} &\Bigg\{ \left[u(c - \underline{c}) + \epsilon f \right] \Delta t + \frac{1}{1 + \widetilde{\rho} \Delta t} \left[\eta(f) \Delta t T(a_{+\Delta t}, r_2) \right. \\
&+ \lambda^e(f) \Delta t \int_w \max \left\{ W(a_{+\Delta t}, r_2, w', f'), W(a_{+\Delta t}, r_2, w, f) \right\} dG_W(w' | w_M) \\
&+ \kappa \Delta t \int_{r_2'} W(a_{+\Delta t}, r_2', w, f) dR(r_2') \\
&\left. + \left(1 - \eta(f) \Delta t - \lambda^e(f) \Delta t - \kappa \Delta t\right) W(a_{+\Delta t}, r_2, w, f) + o(\Delta t) \right] \Bigg\}
\end{aligned}
$$

where:

$$
T(a_{+\Delta t}, r_2) = \int \max \left\{ V(a_{+\Delta t}, r_2), Q(a_{+\Delta t}, r_2, y) \right\} dS(y)
$$

Multiplying by $(1+\widetilde{\rho}\Delta t)$ and rearranging we obtain:

$$\begin{aligned}
\widetilde{\rho}\Delta t W(a,r_2,w,f) = \max_{0\leq c\leq \bar{c}, 0\leq \phi\leq 1} &\{ [u(c-\underline{c})+\epsilon f]\,(1+\widetilde{\rho}\Delta t)\,\Delta t \\
&+ \eta(f)\Delta t\,[T_q(a_{+\Delta t},r_2) - W(a_{+\Delta t},r_2,w,f)] \\
&+ \lambda^e(f)\Delta t \int_w \max\{W(a_{+\Delta t},r_2,w',f') - W(a_{+\Delta t},r_2,w,f),0\}\,dG_W(w'|w_M) \\
&+ \kappa\Delta t \int_{r_2'} [W(a_{+\Delta t},r_2',w,f) - W(a_{+\Delta t},r_2,w,f)]\,dR(r_2') \\
&+ [W(a_{+\Delta t},r_2,w,f) - W(a,r_2,w,f)] + o(\Delta t)\}
\end{aligned}$$

Dividing by Δt:

$$\begin{aligned}
\widetilde{\rho} W(a,r_2,w,f) = \max_{0\leq c\leq \bar{c}, 0\leq \phi\leq 1} &\{ [u(c-\underline{c})+\epsilon f]\,(1+\widetilde{\rho}\Delta t) \\
&+ \eta(f)\,[T_q(a_{+\Delta t},r_2) - W(a_{+\Delta t},r_2,w,f)] \\
&+ \lambda^e(f) \int_w \max\{W(a_{+\Delta t},r_2,w',f') - W(a_{+\Delta t},r_2,w,f),0\}\,dG_W(w'|w_M) \\
&+ \kappa \int_{r_2'} [W(a_{+\Delta t},r_2',w,f) - W(a_{+\Delta t},r_2,w,f)]\,dR(r_2') \\
&+ \left[\frac{W(a_{+\Delta t},r_2,w,f) - W(a,r_2,w,f)}{\Delta t}\right] + o(\Delta t)\Bigg\}
\end{aligned}$$

Taking the limit as Δt goes to zero we have $\lim_{\Delta t\to 0}\frac{W(a_{+\Delta t},r_2,w,f)-W(a,r_2,w,f)}{\Delta t} = \frac{dW(a,r_2,w,f)}{dt}$ and $\lim_{\Delta t\to 0}\frac{o(\Delta t)}{\Delta t} = 0$, therefore:

$$\begin{aligned}
\widetilde{\rho} W(a,r_2,w,f) = \max_{0\leq c\leq \bar{c}, 0\leq \phi\leq 1} &\{ [u(c-\underline{c})+\epsilon f] \\
&+ \eta(f)\,[T_q(a,r_2) - W(a,r_2,w,f)] \\
&+ \lambda^e(f) \int_w \max\{W(a,r_2,w',f') - W(a,r_2,w,f),0\}\,dG_W(w'|w_M) \\
&+ \kappa \int_{r_2'} [W(a,r_2',w,f) - W(a,r_2,w,f)]\,dR(r_2') \\
&+ \frac{dW(a,r_2,w,f)}{dt}\Bigg\}
\end{aligned}$$

To find the last term of this equation we use the chain rule: $\frac{dW}{dt} = \frac{\partial W}{\partial a}\frac{da}{dt}$. Using the budget constraint we have:

$$
\begin{aligned}
\widetilde{\rho}W(a,r_2,w,f) = & \max_{0\leq c\leq\bar{c},0\leq\phi\leq 1}\{[u(c-\underline{c})+\epsilon f] \\
& + \eta(f)\left[T_q(a,r_2)-W(a,r_2,w,f)\right] \\
& + \lambda^e(f)\int_w \max\left\{W(a,r_2,w',f')-W(a,r_2,w,f),0\right\} dG_W(w'|w_M) \\
& + \kappa\int_{r_2'}\left[W(a,r_2',w,f)-W(a,r_2,w,f)\right] dR(r_2') \\
& + \frac{\partial W(a,r_2,w,f)}{\partial a}\left[(r_1\phi+r_2(1-\phi))(1+\nu I_{a^-})a+w(1-\tau f)-c-\Psi(\phi,f)\right]\Bigg\}
\end{aligned}
$$

To derive the Hamilton-Jacobi-Bellman equations for the unemployment and the self-employment states we proceed in the same fashion. The discrete time version of the value of unemployment $V(a,r_2)$, where the length of a time period is Δt, is:

$$
\begin{aligned}
V(a,r_2) = & \max_{0\leq c\leq\bar{c},0\leq\phi\leq 1}\{y(c-\underline{c})\Delta t \\
& + \frac{1}{1+\widetilde{\rho}\Delta t}\Bigg[\lambda^v\Delta t\int_w \max\left\{W(a_{+\Delta t},r_2,w,f),V(a_{+\Delta t},r_2)\right\} dG_W(w'|w_M) \\
& + \kappa\Delta t\int_{r_2'} U(a_{+\Delta t},r_2')dR(r_2') + (1-\lambda^v\Delta t-\kappa\Delta t)V(a_{+\Delta t},r_2) + o(\Delta t)\Bigg]\Bigg\}
\end{aligned}
$$

Rearranging, dividing by Δt and taking the limit as Δt goes to zero we have:

$$
\begin{aligned}
\tilde{\rho}V(a,r_2) = & \max_{0\leq c\leq\bar{c},0\leq\phi\leq 1}\{y(c-\underline{c})+ \\
& + \lambda^v\left(\int_w \max\{W(a,r_2,w,f)-V(a,r_2),0\}dG_W(w'|w_M)\right) \\
& + \kappa\int_{r_2'}\left[V(a,r_2')-V(a,r_2)\right] dR(r_2') \\
& + \frac{\partial V(a,r_2)}{\partial a}\left[(r_1\phi+r_2(1-\phi))(1+\nu I_{a^-})a+b-c-\Psi(a,v)\right]\Bigg\}
\end{aligned}
$$

Finally, the discrete time version of the value of self-employment $Q(a,r_2,y)$, where the length of a time period is Δt, is:

$$\begin{aligned}
Q(a, r_2, y) = & \max_{0 \le c \le \bar{c}, 0 \le \phi \le 1} \{ y(c - \underline{c}) \Delta t \\
& + \frac{1}{1 + \widetilde{\rho} \Delta t} \left[\lambda^q \Delta t \int_w \max \{ W(a_{+\Delta t}, r_2, w, f), Q(a_{+\Delta t}, r_2, y) \} \, dG_W(w' | w_M) \right. \\
& + \pi \Delta t \int_y \max \{ Q(a_{+\Delta t}, r_2, y'), V(a_{+\Delta t}, r_2) \} \, dS(y') \\
& + \kappa \Delta t \int_{r_2'} Q(a_{+\Delta t}, r_2', y) dR(r_2') \\
& \left. + (1 - \lambda^q \Delta t - \pi \Delta t - \kappa \Delta t) Q(a_{+\Delta t}, r_2, y) + o(\Delta t) \right] \}
\end{aligned}$$

Rearranging, dividing by Δt and taking the limit as Δt goes to zero we have:

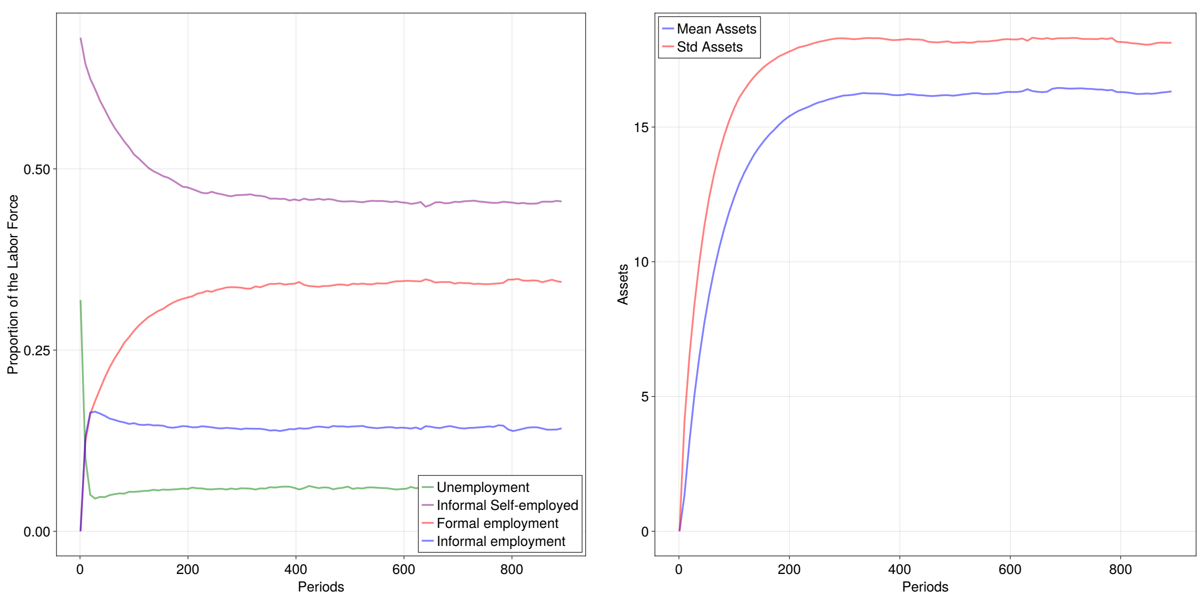
$$\begin{aligned}
\tilde{\rho} Q(a, r_2, y) = & \max_{0 \le c \le \bar{c}, 0 \le \phi \le 1} \{ y(c - \underline{c}) \\
& + \lambda^q \left(\int_w \max \{ W(a, r_2, w, f) - Q(a, r_2, y), 0 \} dG_W(w' | w_M) \right) \\
& + \pi \int_{y'} \max \{ Q(a, r_2, y') - Q(a, r_2, y), V(a, r_2) - Q(a, r_2, y) \} \, dS(y') \\
& + \kappa \int_{r_2'} \left[Q(a, r_2', y) - Q(a, r_2, y) \right] dR(r_2') \\
& \left. + \frac{\partial Q(a, r_2, y)}{\partial a} \left[(r_1 \phi + r_2 (1 - \phi))(1 + \nu I_{a^-}) a + y - c - \Psi(a, q) \right] \right\}
\end{aligned}$$

A.2 Solution Method

We use a two-step approach to solve the model. In the first step, we solve the Hamilton-Jacobi-Bellman equations using a value function iteration method. Following Achdou et al. (2014, 2017), the derivatives of the value functions are approximated using a finite difference with an upwind scheme. That is, a forward difference is used whenever the drift of the state variable (here, $da > 0$ or $dr_2 > 0$) is positive. On the contrary, the backward difference is used when a negative drift occurs. According of Barles and Souganidis (1991), if the conditions of monotonicity, stability, and consistency hold, the solution of the Hamilton-Jacobi-Bellman equations is unique. As Achdou et al. (2017) argues, stability and consistency conditions are relatively easy to hold in this type of model, however for the monotonicity condition to be satisfied the upwind scheme is crucial. An additional advantage of the upwind finite difference scheme is that it allows for the handling of optimization state constraints in a very convenient way. In particular, if the boundary condition associated with the borrowing constraint $a \ge \underline{a}$ is set for the backward difference, and not for the forward difference, and the upwind scheme is allowed to choose the right difference, according to the drift at the bottom of the assets distribution, it will be possible to guarantee that the borrowing constraint is never violated.

In the second step we solve for the invariant distributions of labor market states and of total assets. For this highly non linear model, the transition equations for the state distributions (i.e. the Kolmogorov Forward equations) do not have an obvious explicit algebraic representation and, therefore, cannot be used to compute the invariant distributions. Instead, we use a simulation approach. We simulate labor market careers for a large number of individuals – starting everyone as unemployed with zero assets – and for a large number of periods until both the distributions of labor market states and total assets stabilize. In each period of the career we use the value functions of the first step, together with the optimal decision rules, to govern individual choices on job offers, portfolio allocation and consumption/savings. In the computations we simulate 10,000 careers and obtain the invariant distributions after 800 model periods.

Figure 8: Convergence to the Steady State Distributions of Assets and Labor Market States



(a) Distribution of Labor Market States

(b) Statistics of the Assets Distribution

NOTE: Simulated samples of 10,000 individual-level observations based on the estimates reported in Table 3.

B Identification Appendix

A parametric assumption on the wage offers distribution and the observation of accepted wages and duration are typically enough to identify arrival rates and termination rates in the model (Flinn and Heckman, 1982). In our setting, this identification strategy is more complicated because a given duration may terminate in more than one state: for example, an unemployment duration may end in formal or in informal employment; an employment duration may end in unemployment, self-employment or another employee job. However, since the durations we observe are only on-going, we do not know in which states they terminate. Still, the structure of the model and the steady state equilibrium implications are enough to identify the parameters with the data at our disposal. To show it more formally, we explicitly write the hazard rates and some relevant moments.

The hazard out of unemployment is:

$$\begin{aligned} h_V(a,r_2) &= \lambda^v \int_{\mathcal{A}^V(a,r_2)} dG_W(w|w_M) \qquad (19) \\ \mathcal{A}^V(a,r_2) &\equiv \{w : W(a,r_2,w,f) > V(a,r_2), f=0,1\} \\ h_V &= \int\int h_V(a,r_2)d\Lambda(a)dR(r_2) \qquad (20) \\ E[t_V] &= \frac{1}{h_V} \qquad (21) \end{aligned}$$

where $\mathcal{A}^V(a,r_2)$ denotes the set of wage offers acceptable to an unemployed searcher with assets a and facing an informal rate r_2. In (20), we have to integrate out both the steady state distribution of assets and the stochastic return to the risky asset because they are not observable to us. But what is observable to us is the sample analog of the left hand side of equation (21), which identifies λ^v because the other components of the hazard rate – $\Lambda(a)$, $R(r_2)$ and $G_W(w|w_M)$ – are mainly identified by accepted wages and savings.

The hazard out of self-employment is:

$$\begin{aligned} h_Q(a,r_2,y) &= \lambda^q \int_{\mathcal{A}^Q(a,r_2,y)} dG_W(w|w_M) + \pi \qquad (22) \\ \mathcal{A}^Q(a,r_2,y) &\equiv \{w : W(a,r_2,w,f) > Q(a,r_2,y), f=0,1\} \\ h_Q &= \int\int\int_{\mathcal{B}^Q(a,r_2)} h_Q(a,r_2,y)dS(y')d\Lambda(a)dR(r_2) \qquad (23) \\ \mathcal{B}^Q(a,r_2) &\equiv \{y' : Q(a,r_2,y') > V(a,r_2)\} \\ E[t_Q] &= \frac{1}{h_Q} \qquad (24) \end{aligned}$$

where: t_Q denotes duration in self-employment; $\mathcal{A}^Q(a,r_2,y)$ denotes the set of wage offers acceptable to a self-employed earning y, with assets a, and facing an informal rate r_2; and, $\mathcal{B}^Q(a,r_2)$ denotes the set of the net income levels y produced by self-employment opportunities that makes an individual with $\{a,r_2\}$ choose self-employment instead of unemployment. Note that this set does not depend on the original y: the shock π updates y and therefore the individual cannot revert to the pre-shock self-employment income. Similar to the previous case, the sample analog of the left hand side of equation (24) identifies the mobility rates $\{\lambda^q, \pi\}$ because the other components of the hazard rate – $\Lambda(a)$, $R(r_2)$, $G_W(w|w_M)$, and $S(y)$ – are mainly identified by accepted wages, savings and net self-employment incomes. But in this case the parameters are only jointly identified since we have only one equation for two parameters. Below, we will discuss the steady state conditions that allow for their separate identification.

The hazard out of employment as employee is:

$$\begin{aligned} h_W(a, r_2, w, f) &= \lambda^e(f) \int_{\mathcal{A}^W(a,r_2,w,f)} dG_W(w|w_M) + \eta(f) \qquad (25) \\ \mathcal{A}^W(a, r_2, w, f) &\equiv \{w', f' : W(a, r_2, w', f') > W(a, r_2, w, f)\} \\ h_W(f) &= \int\int\int_{\mathcal{C}^W(a,r_2,f)} dG^f_W(w|w_M) d\Lambda(a) dR(r_2) \qquad (26) \\ \mathcal{C}^W(a, r_2, f) &\equiv \{w' : W(a, r_2, w', f) > T(a, r_2)\} \\ E[t_W(f)] &= \frac{1}{h_E(f)} \qquad (27) \end{aligned}$$

where: $f = 0, 1$; $\mathcal{A}^W(a, r_2, w, f)$ denotes the set of wage offers acceptable to an employee earning w in a job with formality regime f, with assets a, and facing an informal rate r_2; $\mathcal{C}^W(a, r_2, f)$ denotes the set of acceptable wage offer at formality regime f and $G^f_W(w|w_M)$ denotes the accepted wages distribution at jobs with formality regime f. As in the previous case, the sample analog of the left hand side of equation (27) jointly identifies the mobility rates $\{\lambda^e(f), \eta(f)\}$. Notice that we have two versions of equation (27): one for the formal ($f = 1$), and one for the informal ($f = 0$), so this procedure involves a total of four parameters.

At this stage, we have shown the joint identification of three couples of parameters: $\{\lambda^q, \pi\}$, $\{\lambda^e(0), \eta(0)\}$, and $\{\lambda^e(1), \eta(1)\}$. To obtain their separate identification, we exploit the following four steady state conditions:

$$\begin{aligned} \dot{v} &= 0 \qquad (28) \\ \dot{q} &= 0 \qquad (29) \\ \dot{e}(0) &= 0 \qquad (30) \\ \dot{e}(1) &= 0 \qquad (31) \end{aligned}$$

Since $v + q + e(0) + e(1) = 1$, conditions (28)–(31) only define three independent equations. By definition, they are a function of the hazard rates discussed in Equations (19)–(27). Together with the three equations determining the joint identification (equations (24) and (27) for $f = 0$ and $f = 1$), they constitute a system of six equations in six unknowns, securing the separate identification of the six mobility parameters $\{\lambda^q, \pi, \lambda^e(0), \eta(0), \lambda^e(1), \eta(1)\}$.

C Data Appendix (To be done)

We use two primary data sources from Colombia to analyze how informality interacts with labor market and asset accumulation decisions. The first dataset comes from the "Encuesta Longitudinal Colombiana" (ELCA), a household longitudinal survey conducted by the Centro de Estudios sobre Desarrollo Económico (CEDE) at Universidad de los Andes. This survey was specifically designed to track households and individuals over time to understand the socioeconomic dynamics in both urban and rural Colombia. The ELCA follows a panel design that enables researchers to observe the same households and individuals across multiple survey rounds. Data collection began in 2010 as the baseline and was followed by subsequent rounds in 2013 and 2016. To ensure comparability with labor market data, we focus exclusively on the cross-section of the 2016 wave for urban households and individuals.

The second data source is the "Gran Encuesta Integrada de Hogares" (GEIH), Colombia's primary labor force survey conducted by the Departamento Administrativo Nacional de Estadística (DANE). The GEIH is a comprehensive national household survey with a monthly frequency that covers both urban and rural areas across all Colombian departments. This survey provides detailed labor market cross-section information to characterize employment and wage patterns in the Colombian economy. We use GEIH data from 2016, stacking all 12 months of the year to avoid seasonal variations in labor market outcomes. As before, we focus only on urban areas; that is, we use both the "Cabecera" (urban centers) and "Area" (total area including urban).