USFQ School of Economics
Working Paper Series

Age-dependent taxation in emerging markets: A quantitative assessment

Carlos Uribe-Teran[1], Ivan Gachet[2], Diego F. Grijalva[3],[4]

[1] School of Economics, Universidad San Francisco de Quito, cauribe@usfq.edu.ec
[2] World Bank, igachet@worldbank.org
[3] School of Economics, Universidad San Francisco de Quito
[4] Business School, Universidad San Francisco de Quito, dgrijalva@usfq.edu.ec

Working Paper No. 0002 Issue Date: July 8, 2021

Abstract

We study the design, short- and long-run implications of a major reform that substitutes the pay-as-you-go social security system and the personal income tax scheme, with a simplified age-dependent tax schedule, in the context of an emerging market. Our model features overlapping generations, rich heterogeneity among households and informal labor markets, where informality and unemployment are modeled as exclusion mechanisms. We show that the optimal reform triggers a welfare gain of 6.1% measured in compensated equivalent variations, and that most of such gain is explained by a significant wealth effect triggered by the reform. In the short run, however, the reform produces a 12.8% welfare loss mainly explained by the migration from a pay-as-you-go social security schedule to a system of private accounts, in which older cohorts during the transition are the ones more affected. The presence of informality in the labor market, moreover, is key to explain the redistributive consequences of the reform.

Keywords: Consumption and saving, fiscal policy, taxation, social security, demographic trends and macroeconomic effects, occupational choice, informal labor markets, Latin America, Ecuador.
JEL Codes: E21, E62, H20, H55, J11, J24, J46, O54.

UNIVERSIDAD SAN FRANCISCO DE QUITO

SCHOOL OF ECONOMICS USFQ

Age-Dependent Taxation in Emerging Markets: A Quantitative Assessment[*]

Carlos Uribe-Terán **Ivan Gachet** **Diego F. Grijalva**

First Version: September 27, 2019
This version: July 8, 2021

Abstract

We study the design, short- and long-run implications of a major reform that substitutes the pay-as-you-go social security system and the personal income tax scheme, with a simplified age-dependent tax schedule, in the context of an emerging market. Our model features overlapping generations, rich heterogeneity among households and informal labor markets, where informality and unemployment are modeled as exclusion mechanisms. We show that the optimal reform triggers a welfare gain of 6.1% measured in compensated equivalent variations, and that most of such gain is explained by a significant wealth effect triggered by the reform. In the short run, however, the reform produces a 12.8% welfare loss mainly explained by the migration from a pay-as-you-go social security schedule to a system of private accounts, in which older cohorts during the transition are the ones more affected. The presence of informality in the labor market, moreover, is key to explain the redistributive consequences of the reform.

Keywords: Consumption and saving, fiscal policy, taxation, social security, demographic trends and macroeconomic effects, occupational choice, informal labor markets, Latin America, Ecuador.
JEL Codes: E21, E62, H20, H55, J11, J24, J46, O54.

[*]Uribe-Teran: School of Economics, Universidad San Francisco de Quito (corresponding author cauribe@usfq.edu.ec); Gachet: World Bank; Grijalva: Business School and School of Economics, Universidad San Francisco de Quito. We are grateful to Danilo Trupkin, Lian Allub and Guillermo Alves for helpful and detailed comments on an earlier version of this paper. We also like to thank Cristina Arellano, Fernando Álvarez-Parra, Manuel Toledo, Carlos Díaz, Serafín Frache, Juan Dubra, Sergio Parra and participants at CAF Research Seminar in Buenos Aires, Seminars at USFQ, Universidad de Montevideo, Universidad Católica de Montevideo, the LACEA invited session organized by CAF and the 2nd World Bank Tax Conference: Taxing Personal Income and Wealth in Developing Countries for helpful debates, suggestions, and comments. Rafaela Andrade, Priscila Beltran and María Emilia Villacreses provided outstanding research assistance. We also like to thank The Development Bank of Latin America (CAF) for generous financial support under the call for proposals "Demographic and Technological Trends, Labor Markets, and Health and Pension Systems in Latin America". Any opinions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect CAF's or any other organization's views.

1. Introduction

Demographic change and high levels of informality in emerging markets put high and increasing pressure on health and pension systems (Altamirano-Montoya et al., 2018). In the Latin American context, moreover, inequality is high and state capacity is lacking, meaning that governments must grapple with the design of tax policy that is both efficient and redistributive. Following the mirrless tradition (Mirrlees, 1971), optimal tax reforms should aim at minimizing market distortions, especially those related to labor supply. Related to this, a general finding in the literature is that *tagging* (i.e. conditioning tax rates on individuals' observable –and immutable– characteristics, originally proposed in Akerlof, 1978) allows to capture most of the welfare gain of a Mirrlessian reform when the optimal tax scheme depends on age (Weinzierl, 2011). With this finding in mind, can an age-dependent tax scheme replace a pay-as-you-go social security system and a more standard personal income tax scheme, and generate welfare gains, even in the presence of informal labor markets? How would such reform affect inequality?

Based on the largely documented power of age-dependent taxation (Erosa and Gervais, 2002; Karabarbounis, 2016; Weinzierl, 2011; Heathcote et al., 2020), in this paper we design a major reform to the social security system and the personal income tax (PIT) scheme in the context of labor markets characterized by high levels of informality. Our proposal consists two main ingredients. First, to guarantee the sustainability of the social security system in the presence of demographic change and high levels of informality among young workers, we move from a pay-as-you-go social security schedule to a system of private accounts without minimum savings requirement. Second, we simplify the PIT tax scheme replacing progressive labor income and capital income taxation with an age-dependent tax scheme levied over labor income. With this part of the reform, we aim at tackling directly the households' needs of self-insurance triggered by the reform in the social security system.

To study the optimality of the reform quantitatively, we build a large scale macroeconomic model that features overlapping generations, households that face uninsured idiosyncratic risk and partially insured occupational risk, age- and sector-specific productivity and, stochastic retirement and life spans. In the status-quo, working-age population face progressive labor income taxation, flat tax rates over capital income and consumption, and a payroll tax to finance social security transfers. During unemployment spells, workers previously hired in the formal sector have access to unemployment insurance, and there is universal coverage by the social security system to all retirees. Firms hire formal and informal labor to maximize their profits. We assume that firms are not penalized for hiring informal workers.

We model informality and unemployment as exclusion mechanisms in the labor market. This implies that workers cannot control their occupation status, as these are modelled as unexpected shocks that directly affect the workers' productivity and access to social security transfers. In particular, if an informal worker is hit by unemployment, she will need to get through the spell using her own savings, thus without access to unemployment insurance. Besides facing lower productivity compared to a formal job, informal workers do not pay neither of the current labor income taxes (Perry et al., 2007).

The optimal reform implies a change to the structure of the model. In the post-reform economy, the payroll tax, the labor income progressive tax scheme and the flat tax rate for capital income are entirely replaced by a labor income age-dependent tax scheme. Here, tax rates can be positive or negative (targeted transfers to formal workers and the unemployed), and the primary role of the revenue collected under the age-dependent tax scheme is to finance these transfers. The rest of the revenue is used to finance government expenditure, which is then redistributed among all living households through a lump-sum transfer.

To compute the optimal reform we consider a planner that measures social welfare through a utilitarian welfare function and seeks at maximizing welfare in the steady state. From here, the optimization proceeds in steps. First, we check that moving from the pay-as-you-go social security scheme to the private account system produces a welfare gain in the long run. The fact that this type of reform produces such gains has been well documented in previous literature (e.g. Conesa and Garriga, 2000; McKiernan, 2020), so there are no surprises in this step.

In the second step, we allow the planner to set optimally age-dependent tax rates (positive or negative) that affect formal workers only, but we also put in the menu the possibility of a progressive tax scheme for labor income, and a flat tax rate for capital income. For this, we assume that whatever is positively taxed with age-dependent tax rates is deductible for progressive taxation, while age-dependent proportional transfers are not taxable. By doing this, we do not replace the PIT system by assumption, but as an optimal response to social welfare maximization.

We calibrate this model for Ecuador, an interesting case that can shed light on social security reforms for other emerging markets. Since 1970, its population has almost tripled and employed informal population represented more than 60% for the 2008-2017 period.[1] In addition, laws approved in recent years have increased the financial pressure on pension systems due to additional coverage combined with lower revenue,[2] making a social security reform inevitable.

The optimal tax scheme produces a welfare of 6.1% measured in compensated equivalent units, a reduction in consumption and wealth inequality, and an increase disposable income inequality. The drivers or the welfare gain are the enforcement of self-insurance triggered by the inverted U-shape age-dependent tax scheme, and a large wealth effect

[1]This share is higher than the 40% average officially reported. The reason is that we define informal workers as those that do not contribute to social security.

[2]In 2015, the Ecuadorian government stopped its 40% contribution to the pension system (in case of financing gaps). Likewise, it reduced the share allocated to the pension system from 9.74% to 6.06% of workers' contribution. Coverage was also extended to housemaids.

explained by the major simplification of the tax code. The former allows households to accumulate assets faster early in life, and produces a significant increase in the stock of assets in the long run. the latter increases households' disposable income, allowing them to consume more and to enjoy more leisure. The drivers for the results in inequality are related to better insurance because of the increased capital stock (for consumption and wealth) and the elimination of the progressive labor income tax scheme (in the case of disposable income).

The presence of occupational risk, however, reduces the elasticity of labor supply, since households need to take advantage of formality spells that are characterized by high productivity and, during early ages, a proportional transfer implied by the age-dependent tax scheme. In the long-run equilibrium, this produces a shift in the structure of labor supply towards a higher proportion of formal working hours. On average, moreover, leisure increases for young cohorts, and slightly decreases for cohorts that are close to retirement.

Because of the nature of the reform, we also study its effects in the short run. For this, we compute the transition dynamics between the steady state in the status-quo and the post-reform stationary equilibrium. In this exercise we find that the reform produces a 12.8% welfare loss on impact, which is mostly concentrated on older cohorts at the time of the reform.

We organize this paper as follows. In section 2 we present a short literature review to pin down our contribution. Section 3 describes the model and Section 4.1 gives an overview of the estimation and calibration procedures taken to bring the model to the data. In section 5 we provide a detailed description of the optimal reform and explain in detail the changes implied by the reform over the structure of the model. Section 6 presents the results, where we show how the reform affects the economy both in long- and in the short-run. The analysis of the transition dynamics, moreover, allows us to pin down population groups that need to be compensated in the short run. In Section 7 we study how our main results would change have we had assumed a setting with full-employment. Section 8 concludes.

2. Contribution

We contribute to the literature on optimal taxation and social security reforms in two important ways. First, we show how one can rely on the power of age-dependent taxation to design policy reforms that significantly simplify the tax code, triggering important welfare and efficiency gains, and even improved redistribution.[3] Second, we discuss how

[3] We thus contribute to the extensive literature on optimal income taxation; see, for example, Conesa and Krueger (2006); Erosa and Koreshkova (2007); Conesa et al. (2009); Krueger and Ludwig (2013); Kindermann and Krueger (2014); Krueger et al. (2015); Krueger and Ludwig (2016a); McGrattan and Prescott (2017); Heathcote et al. (2020); Uribe-Teran (2020).

the emerging markets context in general, and informal labor markets in particular, affect the design and implications of the optimal age-dependent tax reform something that, to the best of our knowledge, has not been documented in previous literature.

In particular, our paper thus fits in the literature on optimal taxation and social security reforms in emerging markets and how labor market informality affects the design of the optimal tax reform. Regarding the first group of papers, Golosov et al. (2003) show that dynamic optimal fiscal policy is very difficult to implement since the preferred tax scheme depends on households' detailed income history. Weinzierl (2011), however, shows that *tagging* allows to capture most of the welfare gain of the full reform when average marginal rates of the labor income tax depend on age. This constitutes a partial reform since it does not make use of income history. We build upon this finding to show how the totality of the personal income tax system could be replaced by an age-dependent tax scheme, even in the presence of informal labor markets.

In this line, the growing literature on age-dependent taxation focuses on studying optimal tax schedules on labor income for developed economies. One of the most salient result in this research is related to the inverted U-shape of the optimal age-dependent average tax rates. Erosa and Gervais (2002) find that age-dependent taxes arise as a natural implication of life-cycle behavior. The same result is obtained by Karabarbounis (2016), who analyzes optimal labor income tax designs in a very complex setting where the tax schedule is tagged to households' assets, age and filing status (one or two earners), Heathcote et al. (2020) who also allow for age-varying progressivity and Weinzierl (2011).

Ndiaye (2017), moreover, studies the effects of flexible retirement on optimal tax policy design. He finds that when retirement is exogenous (and fixed), the labor tax scheme is increasing in age (as in Farhi and Werning, 2013), but becomes concave once retirement is endogenous. The mechanism behind this result is that labor supply elasticity increases significantly as the individual ages when retirement is endogenous. In this paper, we emulate the effect of this mechanism by imposing stochastic retirement.

We also study how social security reforms affect labor market informality in the short- and long-run. In this regard, Antón (2014); Fernández and Villar (2017); Morales and Medina (2017) find that a tax cut in payroll taxes reduced labor market informality in Colombia, while Gruber (1997) finds that payroll tax reforms applied in Chile had no effects on employment, and most of the variation was concentrated on wages. Other research that follows this line include Bosch and Esteban-Pretel (2015, 2012); Ulyssea (2010); Margolis et al. (2014).

With an approach that is more embedded in the quantitative macro literature, McKiernan (2020) studies the welfare effects of the 1980's social security reform in Chile, where the PAYG contribution scheme was replaced by a fully funded system. In this case the author finds that moving to the fully funded social security scheme increases informal labor supply in the long-run up to a level that is 5 times higher than the pre-reform equilibrium. The key feature that McKiernan (2020) and our paper share in common is that both studies analyze tax reforms that aim to reduce market distortions triggered by the social security contributions scheme, while the literature in the previous paragraph studies the labor supply responses of particular tax reforms in reduced form. Our findings, however, support previous empirical findings that moving towards a private accounts system reduces informality in the long run.

We depart from the literature in the way we model informality. In general, informality is modeled as the result of searching frictions (Zenou, 2008; Amaral and Quintin, 2006; Meghir et al., 2015; Alvarez-Parra and Toledo, 2016), as optimal responses of firms and entrepreneurs (Maloney, 2004; Granda and Hamann, 2015; Ulyssea, 2018; Fortin et al., 1997; De Paula and Scheinkman, 2010; Haanwinckel and Soares, 2020), as an alternative use of time within the household's labor choices (McKiernan, 2020), or as an intensive margin that allows households to reduce the size of taxable labor income (Doligalski and Rojas, 2020).

Instead, we model informality and unemployment as an exogenous source of partially insured risk to reflect that, especially among wage earners, informality is a result of exclusion from the labor market rather than a consequence of individual endogenous choices. Our contribution to the literature is, thus, to study how informality by exclusion affects optimal policy reform and households' choices over the life cycle.[4]

3. The Model

We consider an overlapping generations economy where households work for a limited number of years and retire for the rest of their lives under a pay-as-you-go social security system (in the status quo). During their productive years, workers can move between formality, informality and unemployment. We assume that transitions across occupation status are exogenous, representing an additional source of risk for workers. Workers also face stochastic early retirement.

The typical period in our model is as follows. First, households receive an occupation shock, which determines the type of contract under which they will provide labor. In the case of formal and informal workers, they receive an idiosyncratic productivity shock and decide how much labor to provide, and how much to consume and save. Unemployed and retired households decide how much to consume or save. Formal workers have access to unemployment insurance if the shock occurs in the future, while access to the retirement fund in granted for all retirees.[5]

[4]See, for example, (Perry et al., 2007; Bosch and Maloney, 2007; Kucera and Roncolato, 2008; Mondragon-Velez et al., 2010; Alloush et al., 2013; Williams and Youssef, 2015; Medvedev and Oviedo, 2016; Canelas, 2018).

[5]In principle, this might seem like a strong assumption since, in reality, access to the retirement fund depends on workers' labor history. We want to keep the model as tractable as possible, however, and, in the absence of endogenous decisions on occupational status, conditional access to retirement transfers does not affect our main results.

The representative firm hires capital, formal and informal labor to produce the consumption good in the economy (numeraire) and maximize profits. We assume a perfect foresight equilibrium, so the economy faces no aggregate uncertainty.

3.1. Households

The economy is populated by one-member households that enter the labor market at model age 1 under a random occupation status (formal, informal or unemployed). Households face stochastic life spans, age- and occupation-specific productivity and uninsured productivity shocks. They also face *partially* insured occupational status in the sense that only formal workers have access to unemployment insurance. All retirees have access to the pensions fund. Each period workers are endowed with one unit of time that has to be distributed between leisure and work.

3.1.1. Demographics

There are J overlapping generations. Each period, a continuum of agents enters the labor market and working population grows at a constant rate n. Agents face stochastic life spans, where ψ_j is the conditional probability of being alive next year; all agents die with certainty at age J. Assets left by the dead are confiscated by the government, taxed at the current tax rate for capital income, and distributed among the living via a lump-sum transfer. This structure implies a stationary age distribution with ν_j denoting the proportion of households of age j at any point in time.

Workers face occupational risk, so they move exogenously between states $S = \{f, i, u\}$ where f is a formal contract, i an informal contract and u unemployment. We understand a formal contract as a labor relation where the worker contributes to the social security fund and pays taxes. Moreover, a formal contract guarantees access to unemployment insurance.

Transitions among occupation states are governed by a transition probability matrix $\mathbf{P}$, where each entry $p(s_{j+1}, s_j) \equiv \mathrm{P}(s_{j+1}|s_j)$ is the conditional probability of moving from state s_j to state s_{j+1} the following year. We also assume that $p(s_{j+1}, s_j) > 0$ for all possible transitions, so that the Markov chain generated by $\mathbf{P}$ and S converges to a unique stationary distribution.

Workers also face stochastic early retirement, where the probability of early retirement is conditional on age p_j^r. All agents retire with certainty at age j^r, i.e. $p_j^r = 1$ for all $j \geq j^r$. During working age, agents face age-specific productivity levels that depend on occupational status ε_j^s where $\varepsilon_j^s = 0$ for $j > j^r$, $\varepsilon_j^u = 0$ for all $j \geq 1$ and $\varepsilon_j^f \geq \varepsilon_j^i$ for all $j \geq 1$. The latter reflects that, on average, formal contracts are at least as productive as informal contracts within every age group.

3.1.2. Income Process

Households' income comes from wages w_t^s for $s = \{f, i\}$ that are determined in equilibrium. The income process is affected by idiosyncratic productivity shocks η_t, age-specific and occupation-dependent productivity level ε_j^s with $\varepsilon_j^u = \varepsilon_j^r = 0$ for all $j \in \{1, \ldots, J\}$ and all $t \geq 0$, and the amount of hours supplied by workers in each occupation status ℓ_t^s. Formally, households' labor income during period t at age j in occupation status s is given by

$$\log y_{jt}^s = \log \varepsilon_j^s + \log \eta_t + \log w_t^s + \log \ell_{jt}^s,$$

where we do not include a household-specific subscript to mantain notation tractable. The idiosyncratic productivity shock follows an $AR(1)$ process such that

$$\log \eta_{t+1} = \rho_\eta \log \eta_t + \epsilon_{t+1},$$

where ρ_η is the coefficient of autocorrelation and ϵ_t is assumed to be normally distributed with mean zero and variance σ_η^2. Let $F_\eta(\eta_{t+1}|\eta_t)$ be the cumulative distribution function for η_{t+1} conditional on the value of η_t which we assume stationary.

3.1.3. Recursive Formulation

We begin by defining the state-space for the households' problem. As before, we avoid using household-specific subscripts for ease of notation, but maintain the time subscript, since we are interested in both the stationary equilibrium and the transition path. With this in mind, the state space can be written as

$$x_t = \{a_t, \eta_t, s_t, \tilde{s}_t, j\},$$

where $s_t \in \{f, i, u, r\}$, $\tilde{s}_t = s_{t-v}$ and v measures the duration of unemployment, so $\tilde{s}_t$ records the last contract type under which the household was hired before being hit by the unemployment shock. For reason that will become clear in what follows, we find it helpful to partition this steady state defining the relevant portion for each occupation status. With this reasoning, we define the state space for formal and informal workers as $x_t^f = \{a_t, \eta_t, f, j\}$ and $x_t^i = \{a_t, \eta_t, i, j\}$ respectively, for unemployed workers $x_t^u = \{a_t, u, \tilde{s}_t, j\}$ and for retirees $x_t^r = \{a_t, r, j\}$.

With these considerations, we first work the case of a household that is hired under a formal contract. In this case, she decides how much to consume and work and pays for consumption, income taxes and social security contributions. Let p_j^r denote the probability of retirement at age j keeping in mind that $p_j^r = 1$ for all $j > j^r$. The problem for the formal worker can be written as,

$$
\begin{aligned}
& V_t(x_t^f) = \max\left\{U(c_t,\ell_t) + \beta\psi_j \,\mathrm{E}_t\, \tilde{V}_{t+1}^f(x_{t+1})\right\}, \\
\text{s. t.} \quad & a_{t+1} = (1+R_t)a_t + (1-\tau_t^{ss})y_{jt}^f - T((1-\tau_t^{ss})y_{jt}^f) + Tr_t + G_t - (1+\tau_t^c)c_t, \\
& a_{t+1} \geq 0, \quad c_t > 0, \quad 0 \leq \ell_t \leq 1,
\end{aligned} \tag{1}
$$

where

$$
\begin{aligned}
\mathrm{E}_t\, \tilde{V}_{t+1}^f(x_{t+1}) &= (1-p_j^r)\left[\int \sum_{s'=i,f} p(s',f)V_{t+1}(x_{t+1}^{s'})F_\eta(d\eta_{t+1}|\eta_t)\right. \\
&+ \left. p(u,f)V_{t+1}(x_{t+1}^u)\right] + p_j^r V_{t+1}(x_{t+1}^r)
\end{aligned}
$$

is the expectation taken over occupational transitions, possible early retirement and future realizations of the idiosyncratic shock, β is the discount rate, $R_t = (1-\tau^k)r_t$ is the capital return net of taxes, τ_t^{ss} is the payroll tax, $T(\cdot)$ is the progressive tax scheme for labor income, Tr_t is the lump-sum transfer originated in the assets left by the deceased, G_t is a government lump-sum transfer and τ_t^c is the tax rate for consumption that balances the government's budget constraint.

When the worker is hired in the informal sector, she does not pay labor taxes and decides how much to consume, save and work, given the tax rates for consumption and capital income, and equilibrium prices. For this group, the problem is

$$
\begin{aligned}
& V_t(x_t^i) = \max\left\{U(c_t,\ell_t) + \beta\psi_j \,\mathrm{E}_t\, \tilde{V}_{t+1}^i(x_{t+1})\right\}, \\
\text{s. t.} \quad & a_{t+1} = (1+R_t)a_t + y_{jt}^i + Tr_t + G_t - (1+\tau_t^c)c_t, \\
& a_{t+1} \geq 0, \quad c_t > 0, \quad 0 \leq \ell_t \leq 1,
\end{aligned} \tag{2}
$$

where

$$
\begin{aligned}
\mathrm{E}_t\, \tilde{V}_{t+1}^i(x_{t+1}) &= (1-p_j^r)\left[\int \sum_{s'=i,f} p(s',i)V_{t+1}(x_{t+1}^{s'})F_\eta(d\eta_{t+1}|\eta_t)\right. \\
&+ \left. p(u,i)V_{t+1}(x_{t+1}^u)\right] + p_j^r V_{t+1}(x_{t+1}^r).
\end{aligned}
$$

If the worker is unemployed she only decides how much to consume and save. The occupation status during the last period active in the labor market determines whether or not she has access to unemployment insurance. We assume that if the worker receives a job offer (formal or informal) in the next period, then she leaves unemployment with the average idiosyncratic shock $\overline{\eta}$. The problem of the unemployed worker can thus be written as,

$$
\begin{aligned}
& V_t(x_t^u) = \max\left\{u(c_t) + \beta\psi_j \operatorname{E}_t \tilde{V}^u_{t+1}(x_{t+1})\right\}, \\
\text{s.t.} \quad & a_{t+1} = (1+R_t)a_t + b_t^u(\tilde{s}_t) + Tr_t + G_t - (1+\tau_t^c)c_t, \\
& a_{t+1} \geq 0, \quad c_t > 0,
\end{aligned} \tag{3}
$$

where

$$
\begin{aligned}
\operatorname{E}_t \tilde{V}^u_{t+1}(x_{t+1}) &= (1-p_j^r)\left[\int \sum_{s'=i,f} p(s',u)V_{t+1}(x_{t+1}^{s'})F_\eta(d\eta_{t+1}|\eta_t) \right. \\
&+ \left. p(u,u)V_{t+1}(x_{t+1}^u)\right] + p_j^r V_{t+1}(x_{t+1}^r),
\end{aligned}
$$

and $b_t^u(\tilde{s}_t)$ is the unemployment insurance transfer with $b_t^u(i) = 0$ and $b_t^u(f) = \lambda_t^u$. This implies that only workers with formal contracts have unemployment insurance.

Regarding retirees we assume that, contrary to occupational status shocks, the retirement shock is irreversible and all workers have access to the pension fund in the form of lump-sum transfers to retirees. The program for retirees is

$$
\begin{aligned}
& V_t(x_t^r) = \max\left\{u(c_t) + \beta\psi_j V_{t+1}(x_{t+1}^r)\right\}, \\
\text{s.t.} \quad & a_{t+1} = (1+R_t)a_t + b_t^r + Tr_t + G_t - (1+\tau_t^c)c_t, \\
& a_{t+1} \geq 0, \quad c_t > 0.
\end{aligned} \tag{4}
$$

3.2. Firm

We consider a representative firm that hires capital and labor to produce the consumption good in the Economy; labor can be hired in the formal or the informal sector. We assume a Cobb-Douglas production function with respect to capital and labor, and a CES aggregator for both types of labor. Formally,

$$
Q(K_t, L_t^f, L_t^i) = AK_t^\alpha \left[\alpha^f \left(L_t^f\right)^\gamma + (1-\alpha^f)\left(L_t^i\right)^\gamma\right]^{\frac{1-\alpha}{\gamma}},
$$

where α represents the share of output paid to capital, α_f is the proportion of formal labor required for production and γ is related to the elasticity of substitution between formal and informal labor, K_t is the available stock of capital in the economy, L_t^f is formal labor and L_t^i is informal labor.

The firm seeks to maximize profit by hiring labor and capital in competitive markets. We assume that households are the owners of capital, so the firm pays rent net of depreciation for each unit used for production, and wages for each unit of formal and informal labor that is hired. We also assume that there are not adjustment costs for any input. The firm, thus, maximizes profit every period, but decisions in period t do not affect the state that the firm faces in $t+1$, so the maximization problem can effectively be described as an static problem. Formally, each period the firm solves,

$$\max \qquad \Pi_t(K_t, L_t^f, L_t^i) = Q(K_t, L_t^f, L_t^i) - w_t^f L_t^f - w_t^i L_t^i - (r_t + \delta)K_t,$$

where δ is the capital's depreciation rate. We assume that the firm can change the type of contract (formal or informal) or terminate it without restrictions (exclusion mechanism). The first order conditions for the firm's maximization problem determine prices in the economy and can be written as

$$r_t = \alpha \frac{q_t}{k_t} - \delta, \tag{5}$$

$$w_t^f = (1-\alpha)\alpha^f \left(\frac{h_t^f}{h_t}\right)^{\gamma} \frac{q_t}{h_t^f}, \tag{6}$$

$$w_f^i = (1-\alpha)(1-\alpha^f) \left(\frac{h_t^i}{h_t}\right)^{\gamma} \frac{q_t}{h_i}, \tag{7}$$

where $q_t = Ak_t^{\alpha}\left[\alpha^f \left(h_t^f\right)^{\gamma} + (1-\alpha^f)\left(h_t^i\right)^{\gamma}\right]^{\frac{1-\alpha}{\gamma}}$ and k_t are output and capital per unit of effective labor, respectively, and h_t^f and h_t^i are the ratios of formal and informal labor with respect to total labor hired by the firm, respectively.

Similar to Krueger and Ludwig (2016b), this configuration allows us to derive an explicit expression for the formality wage premium which can be written as

$$\frac{w_t^f}{w_t^i} = \frac{\alpha^f}{1-\alpha^f}\left(\frac{h_t^f}{h_t^i}\right)^{\gamma-1}. \tag{8}$$

Since (8) depends on the supplies of formal and informal labor, the wage premium is endogenous in the model and its determined on equilibrium. We can, moreover, use this expression as foundation to calibrate the value of the elasticity of substitution between formal and informal labor γ to match the wage premium that we observe in the data.

3.3. Social Security

In the status-quo, the social security system manages a fund that is distributed between unemployment and retirement insurance. Formal workers contribute by means of a flat payroll tax τ_t^{ss} under a pay-as-you-go contribution scheme. In the benchmark, we replicate the current contribution scheme in Ecuador.

A proportion θ_t of the total revenue available in the fund is used to finance the retirement insurance transfer, and the remaining $1-\theta_t$ of the fund finances unemployment insurance for previously formal workers. We require the social security system to always run a balanced budget, so short- of long-run deficits are not possible. Thus, transfers λ_t^u and b_t^r are determined in equilibrium.

3.4. Government

The role of government is limited to collecting taxes and then redistributing its revenue among all the living households throught a lump-sum transfer. Capital income and consumption are taxed under flat schemes with rates τ^k and τ_t^c. Labor income, on the other hand, is taxed using a progressive tax scheme following Kindermann and Krueger (2014). Taxable income corresponds to labor income from formal workers net of the payroll tax:

$$\widehat{y}_{jt}^f = (1 - \tau_t^{ss}) y_{jt}^f.$$

Under this tax scheme, the marginal tax rate for labor income can be computed as

$$T'(\widehat{y}_{jt}^f) = \begin{cases} 0 & \text{if } \widehat{y}_{jt}^f < \underline{y}_t, \\ \underline{\tau}_t + \tau_t^m(\widehat{y}_{jt}^f - \underline{y}_t) & \text{if } \underline{y}_t \leq \widehat{y}_{jt}^f < \overline{y}_t, \\ \overline{\tau}_t & \text{if } \widehat{y}_{jt}^f \geq \overline{y}_t, \end{cases}$$

where $\overline{\tau} \geq \underline{\tau}$ are the maximum and minimum marginal tax rates respectively, $\underline{y}_t$ is the amount of labor income that is exempt, $\overline{y}_t$ is an income threshold after which the marginal tax rate becomes flat, and $\tau_t^m = (\overline{\tau} - \underline{\tau})/(\overline{y}_t - \underline{y}_t)$ corresponds to the linear increase in the marginal tax rate for income levels that lie within $\underline{y}_t$ and $\overline{y}_t$.

In the status-quo, these are the only sources of tax revenue for the government to finance a lump-sum transfer to all the living households in the economy. We assume that this expenditure is a fixed proportion of output, so the lump-sum transfer is $G_t = gQ_t$ with $0 \leq g \leq 1$.

3.5. Equilibrium

In this section we define the non-stationary and stationary competitive equilibrium of the economy in the status quo. This is necessary because, in spite of the long-run welfare gain triggered by the optimal reform, some of its components might generate welfare losses in the short-term. We define the competitive equilibrium over the state space of the economy, but keep in mind that conditional on the occupational status, there might be sections of the state space with zero mass. Nonetheless, the unconditional distribution has non-zero mass over the entire state-space.

Remember, thus, that the full state variables in the economy are given by

$$x_t = \{a_t, \eta_t, s_t, \tilde{s}_t, j\},$$

where $a_t \in \mathbb{R}_+$, $\eta_t \in \mathbb{R}_+$, $s_t \in \mathcal{S} = \{f, i, u, r\}$, $\tilde{s}_t \in \tilde{\mathcal{S}} = \mathcal{S} - \{r\}$ and $j \in \mathcal{J} = \{1, \ldots, J\}$, so the state-space for the economy is

$$\mathcal{X} = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{S} \times \tilde{\mathcal{S}} \times \mathcal{J}.$$

Let $\mathscr{B}(\mathbb{R}_+)$ be the Borel σ-algebra of $\mathbb{R}_+$ and $\mathscr{P}(\cdot)$ the power set for $\mathcal{S}$, $\tilde{\mathcal{S}}$ and $\mathcal{J}$. Let, moreover, $\mathcal{M}$ be the set of all finite measures defined over

$$\left(\mathcal{X}, \mathscr{B}(\mathbb{R}_+) \times \mathscr{B}(\mathbb{R}_+) \times \mathscr{P}(\mathcal{S}) \times \mathscr{P}(\tilde{\mathcal{S}}) \times \mathscr{P}(\mathcal{J})\right).$$

With this preamble, Definition 3.1 presents the non-stationary competitive equilibrium conditions and Definition 3.2 deals with the stationary equilibrium.

Definition 3.1 (Competitive Equilibrium)**.** *Given sequences of capital and consumption tax rates* $\{\tau_t^c, \tau_t^k\}_{t=1}^{\infty}$*, the progressive labor income tax scheme* $\{T(\underline{\tau}_t, \overline{\tau}_t, \underline{y}_t, \overline{y}_t)\}_{t=1}^{\infty}$*, payroll tax rates* $\{\tau_t^{ss}\}_{t=1}^{\infty}$*, proportion of the pension fund that is paid to retirees* $\{\theta_t\}_{t=1}^{\infty}$ *and initial conditions* K_1, Φ_1*, a competitive equilibrium is a sequence of functions for households* $\{V_t, c_t, a_t, \ell_t\}_{t=1}^{\infty}$*, production plans for the firm* $\{K_t, L_t^f, L_t^i\}_{t=1}^{\infty}$*, pension and unemployment benefits* $\{\lambda_t^u, b_t^r\}_{t=1}^{\infty}$*, prices* $\{r_t, w_t^f, w_t^i\}_{t=1}^{\infty}$*, transfers derived from accidental bequests* $\{Tr_t\}_{t=1}^{\infty}$ *and measures* $\{\Phi_{t+1}\}_{t=1}^{\infty}$ *with* $\Phi_t \in \mathcal{M}$ *for all* $t \geq 1$ *such that:*

(i) Households maximize their life-time expected utility, so $\{V_t^s, c_t, a_t, \ell_t\}_{t=1}^{\infty}$ *for all* $s \in \mathcal{S}$ *solve problems (1) to (4),*

(ii) Production plans $\{K_t, L_t^f, L_t^i\}_{t=1}^{\infty}$ *maximize profits of the firm, so prices* $\{r_t, w_t^f, w_t^i\}_{t=1}^{\infty}$ *satisfy equations (5) to (7),*

(iii) The social security system's budget constraint are satisfied, so

$$b_t^r \int \Phi_t(dx_t^r) = \theta_t \int \tau_t^{ss} \varepsilon_j^f \eta_t w_t^f \ell_t(x_t^f) \Phi_t(dx_t^f), \tag{9}$$

$$\lambda_t^u \int \Phi_t(da_t, \cdot, u, f, dj) = (1 - \theta_t) \int \tau_t^{ss} \varepsilon_j^f \eta_t w_t^f \ell_t(x_t^f) \Phi_t(dx_t^f), \tag{10}$$

(iv) The accidental bequest lump-sum transfer is given by

$$Tr_{t+1} = \int (1 - \psi_j) a_t(x_t) \Phi_t(dx_t), \tag{11}$$

(v) Markets clear:

$$C_t = \int c_t(x_t)\Phi_t(dx_t), \tag{12}$$

$$K_{t+1} = \int a_{t+1}(x_t)\Phi_t(dx_t), \tag{13}$$

$$L_t^f = \int \varepsilon_j^f \eta_t \ell_t(x_t^f)\Phi_t(dx_t^f), \tag{14}$$

$$L_t^i = \int \varepsilon_j^i \eta_t \ell_t(x_t^i)\Phi_t(dx_t^i), \tag{15}$$

$$K_{t+1} + C_t + G_t = Q(K_t, L_t^f, L_t^i) + (1-\delta)K_t, \tag{16}$$

(vi) The government runs a balanced budget,

$$G_t = \tau_t^c C_t + \tau_t^k K_t + \int T((1-\tau^s s_t)\varepsilon_j^f \eta_t w_t^f \ell_t(x_t^f)\Phi_t(dx_t^f), \tag{17}$$

(vii) Law of Motion: Let $f_\eta(\eta_t)$ denote the unconditional probability distribution function of η_t, $f\eta(\eta_{t+1}|\eta_t = F'(\eta_{t+1}|\eta_t)$, $p_0(s)$ the unconditional probability of being born into an occupational state $s \in \mathcal{S}$ and $p(\tilde{s}_t)$ the probability of ending an employment spell in occupational status $s_t \in \{f, i\}$. Then

$$\Phi_{t+1} = H_t(\Phi_t), \tag{18}$$

where $H_t : \mathcal{M} \to \mathcal{M}$ is:

For all $j \in \mathcal{J} - 1$,

$$\Phi_{t+1}(\mathcal{X}) = \int P_t(x_t; \mathcal{X})\Phi_t dx_t,$$

with

$$P_t(x_t;\mathcal{X}) = \begin{cases} (1-p_j^r)\psi_j p(s_{t+1}, s_t) f\eta(\eta_{t+1}|\eta_t) & \textit{if } a_t \in \mathbb{R}_+, s_{t+1} \in \{f,i\}, j+1 < j^r \in \mathcal{J}, \\ (1-p_j^r)\psi_j p(s_{t+1}, s_t) p(\tilde{s}_t) & \textit{if } a_t \in \mathbb{R}_+, s_{t+1} \in \{u\}, j+1 < j^r \in \mathcal{J}, \\ p_j^r \psi_j & \textit{if } a_t \in \mathbb{R}_+, j \in \mathcal{J} - 1, \\ 0 & \textit{otherwise,} \end{cases}$$

and, for $1 \in \mathcal{J}$

$$\Phi_{t+1}(\mathcal{X}) = (1+n)^t \begin{cases} p_0(s) f(\eta) & \textit{if } a_1 \in \mathbb{R}_+, \\ 0 & \textit{otherwise.} \end{cases}$$

Definition 3.2 (Stationary Equilibrium)**.** *A stationary equilibrium is a competitive equilibrium in which per capita variables, policy functions, prices and policies, are constant; distributions are invariant, and aggregate variables grow at the constant growth rate of the population.*

4. Matching the Model to the Data

We estimate most of the required parameters to discipline the model, and we leave for calibration those that cannot be fully identified with our data. We use official population forecasts and household surveys from the National Institute of Statistics (INEC), a panel of formal firms' balance sheets from the Superintendency of Companies (SuperCias), and time series of macroeconomic aggregates from the Federal Reserve Bank of St. Louis economic dataset (FRED Economic Data). We provide detailed information about each source in the following sections.

4.1. Calibration

We assume a utility function that is additively separable between consumption and leisure, and over time. Formally,

$$U(c, 1-\ell) = \frac{c^{1-\sigma_1} - 1}{1-\sigma_1} + \chi \frac{(1-\ell)^{1-\sigma_2}}{1-\sigma_2},$$

where σ_1 and σ_2 correspond to the coefficients of risk aversion associated with each good. We calibrate σ_1 and σ_2 following previous literature and χ to reproduce the average hours worked that we observe in the data. To compute this target, we use annualized hours worked by wage-earners obtained from Ecuador's labor surveys for the period 2001-2017.[6]

We calibrate the discount factor β to replicate the capital-output ratio in steady state. We obtain Ecuador's series of capital stock and output from the Federal Reserve of St. Louis dataset (FRED Economic Data) for the period 1965-2014.

The calibration of the idiosyncratic shock allows the model to replicate income inequality along the life-cycle. We assume that households are log-normally distributed over possible productivity shocks at age 20 with mean zero and standard deviation σ_μ. The standard deviation of the error term in the $AR(1)$ process of the productivity idiosyncratic shock σ_η is calibrated to generate the variance of the logarithm of labor income at age 50. The autocorrelation coefficient of the $AR(1)$ process is set to replicate the linear increase in the variance of the logarithm of labor income from age 20 to 50 that we observe in the data. This calibration is summarized in Table 1.

[6] In particular, throughout this paper we use the Tax Dataset for Ecuador covering the period 2001-2017 constructed by Gachet et al. (2018). This dataset combines tax and social security legislation with official labor surveys published by The National Institute of Statistics and Census (INEC).

Parameter	Target	Value
σ_1	Literature	2.0000
σ_2	Literature	3.0000
χ	Average hours worked	0.7572
β	Steady-state capital-output ratio	0.9744
σ_μ	Variance of log income at age 20	0.5903
σ_η	Variance of log income at age 50	0.1152
ρ_η	Linear increase in variance of log income	0.9900

Table 1: **Calibrated Parameters for Households.** Parameters are computed minimizing a weighted quadratic loss function except for σ_1, σ_2 and ρ_η.

We calibrate the depreciation rate δ so that the model matches the investment-capital ratio in steady-state. The elasticity of substitution between formal and informal labor γ is calibrated to hit the steady-state formality wage premium that we observe in households' labor survey data from INEC. Table 2 presents the values for these parameters.

Parameter	Target	Value
δ	Steady-state investment-capital ratio	0.0774
γ	Steady-state formality wage premium	1.0103

Table 2: **Parameters for Representative Firm.**

On the revenue side of the fiscal budget constraint, the government runs a progressive tax scheme on labor income, together with flat tax rates on capital income and consumption. The social security system, moreover, operates under a fully balanced pay-as-you-go scheme in the status-quo, where formal workers contribute to the social security fund by means of a flat payroll tax rate.

Parameter	Target	Value
$\underline{\tau}_t$	Lowest labor income marginal tax rate	0.0500
$\overline{\tau}_t$	Highest labor income marginal tax rate	0.3500
τ_t^k	Capital income tax rate	0.2500
τ_t^{ss}	Payroll tax rate	0.0945

Table 3: **Parameters for Government and Social Security.** Tax rates correspond to Ecuador's current legislation, except for the capital income tax rate which corresponds to a simple average of nominal marginal tax rates.

The lowest and highest marginal tax rates in the labor income progressive tax scheme $\underline{\tau}_t$ and $\overline{\tau}_t$, and the social security payroll tax rate τ_t^{ss} are set according to current legislation. The capital income tax rate τ_t^k corresponds to a simple average of the applicable marginal tax rates according to the Personal Income Tax scheme for 2016. The consumption tax rate τ_t^c ensures that the government runs a balanced budget at all times, so it is determined in equilibrium. We present these tax rates in Table 3.

4.2. Estimation

In this section we describe the strategy used to estimate the parameters that we can identify with our data. This is a particularly challenging task, since it is the first time that this type of model is built for Ecuador. To our surprise, however, household surveys and publicly available data provide very rich information to estimate the rest of parameters required to discipline the model.

4.2.1. Population Dynamics

By population dynamics we refer to the growth rate of population and the conditional survival probability. In this regard, INEC publishes official population forecasts for Ecuador in a very detailed fashion, including population growth by age and gender.[7]

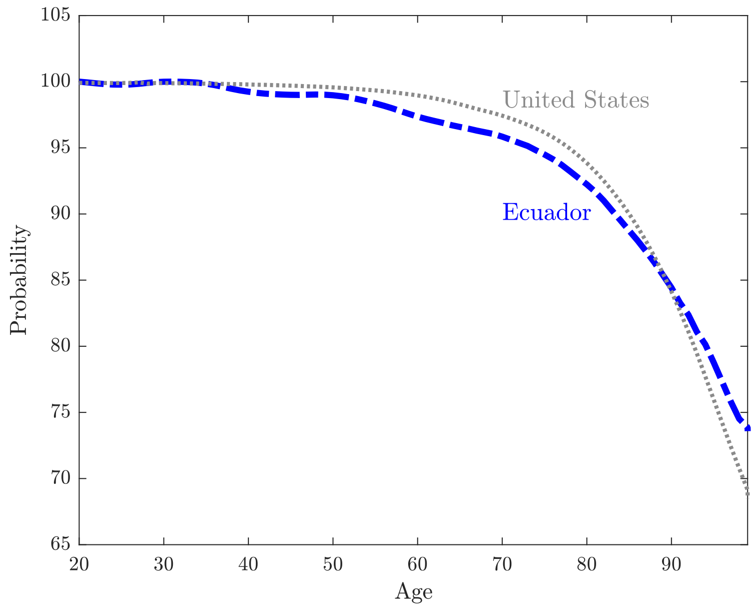


Figure 1: **Conditional survival probability from age 1 to 99 (ψ_j).** Survival probabilities for Ecuador are computed using official population forecasts published by the National Agency of Statistics (INEC). For the United States we use the estimates obtained by Bell and Michael (2002).

We set the population growth rate to $n = 1.55\%$, that corresponds to the average population growth rate between 1990 and 2020. Regarding conditional survival probabilities, we use the population distribution by age for 2017 and compute $\psi_j = pop_{j+1}/pop_j$ for $j = 2, \ldots, 99$, where pop_j is the population of age j. We present these probabilities in Figure 1 and compare them to the ones obtained for the United States (Bell and Michael, 2002). As this figure shows, survival probabilities in Ecuador are significantly lower than in the US, implying that the longevity savings motive is weaker in Ecuador.

[7]This information is available at `http://www.ecuadorencifras.gob.ec/proyecciones-poblacionales/`.

4.2.2. Age-specific productivity

To estimate age-specific productivity we use the definitions in Hansen (1993) to estimate age and labor sector specific productivity profiles. For this we use a repeated cross section of households' labor surveys dataset from 2001 to 2017 (Gachet et al., 2018). We compute real wages using the CPI and then calculate effective hours worked for each household conditional on whether the worker is formal or informal as

$$\tilde{h}_t^{is} = \frac{w_t^{is}}{w_t^s} h_t^i, \tag{19}$$

where w_t^{is} is hourly wage reported by worker i in year t within sector s, w_t^s is the average hourly wage in sector s at time t and h_t^i are normalized yearly hours worked for each worker in a given year.

We include independent workers in our sample. In this case, however, we are not able to observe their hourly market wage, since independent workers report only their total income, which can include both labor and capital income. We thus estimate a Heckman selection model to correct for households self-selection into independent work, and then use the results to estimate the market wage consistent with individuals' observable characteristics.[8]

Then we perform a pooled non-parametric regression to capture how much of the variation in $\tilde{h}_t^{is}$ can be explained by the age of individuals. Specifically, within each sector, we estimate

$$\hat{h}_t^{is} = m^s(j_t^i) + \epsilon_t^{is}, \tag{20}$$

where j_t^i is the age of each individual within each sector in our sample. We interpret $\hat{m}^s(j_t^i)$ as the part of effective labor that can be explained by age, and $\hat{\epsilon}_t^{is}$ as the part that is explained by other factors such as idiosyncratic productivity shocks, ability, etc. To estimate $\hat{m}^s(\cdot)$ we follow

$$\hat{m}^s(j_t^i) = \frac{1}{N^s} \sum_{i=1}^{N^s} \frac{K_h(j - j_t^i)}{\sum_{i=1}^{N^s} K_h(j - j_t^i)} h_t^{is}, \tag{21}$$

where K_h is the Epanechnikov kernel for a given bandwidth h (Epanechnikov, 1969) and N^s is the total number of observations within sector s in our pooled sample. Then, given our age grid $j_g = \{20, 21, \ldots, 64, 65\}$ we compute age-specific productivity as

$$\varepsilon_j^s = \hat{m}^s(j), \tag{22}$$

[8] We provide a detailed description of the data used for this procedure and the results in the Appendix.

where $j \in j_g$. Since the grid over which $\hat{m}^s(\cdot)$ is defined is not j_g necessarily, we use piecewise cubic Hermite interpolating polynomials to obtain ε_j^s (Fritsch and Carlson, 1980). The estimated occupation-specific age-profile of labor productivity is presented in Figure 2.

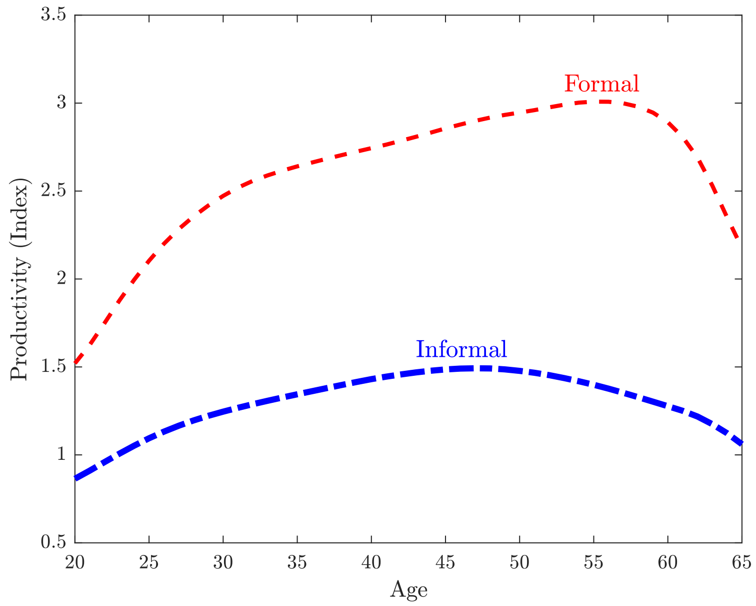


Figure 2: **Age-specific productivity from age 20 to 65 (ε_j).** Productivity for Ecuador is computed using information from the Labor National Survey (ENEMDU).

Our results show the significant productivity gap between formal and informal workers along the life cycle. Formal workers enter the labor market in the model with a productivity level that is 75.9% higher than informal workers, and the gap tends to increase over the life cycle. The productivity dynamics in the two sectors are also very different. While the peak of productivity in the informal sector occurs at around 47 years of age, in the formal sector occurs at age 55, after which there is a severe drop of nearly 20% in the last year before retirement.

These results have important implications in shaping the inequality age profile in our model, since the variance in the cross-section is not only explain by the unexpected productivity shock, but also by the productivity differences that arise due to the type of contract under which the worker is supplying labor. The occupation-specific productivity, moreover, adds to the insurance motives that households have to generate precautionary savings, since moving between productivity levels is entirely governed by the stochastic transitions between type of contracts. We describe such transitions in the following section.

4.2.3. Occupation transitions

One of the main challenges we face is that our model needs to generate a realistic life cycle profile of occupational risk. For this, we opt for estimating occupation transitions directly, which is equivalent to assume that occupation status is exogenous and responds to the embedded structure of the labor market.

We use labor surveys to compute these transitions. For this, we exploit a special feature of these surveys: By design, samples are mobile panels that maintain the same dwellings for two consecutive years every two years. This implies that, for example, surveys for 2015 and 2016 can be considered as a panel at the dwelling level, but although the previous pair 2013-2014 is a panel as well, it is not compatible with the one for 2015-2016.

We concentrate on the available panels between 2011-2015 and match individuals that live in the same dwelling using age and gender focusing only on wage earners. We estimate the transition matrix $\mathbf{P}$ of a Markov chain with states $\mathcal{S} = \{f, i, u\}$. To estimate $\mathbf{P}$ we rely on the frequency of observed transitions between states for all available panels using the surveys' sample weights. Let n_t^s be the number of individuals in year t whose current occupational status is s. Then, the joint probability of being in status s and s' in two consecutive years is simply

$$\mathrm{P}(s, s') = n_t^{s,s'} \left(\sum_{s} \sum_{s'} n_t^s n_{t+1}^{s'} \right)^{-1}, \tag{23}$$

while the marginal probability for each state s in year t is

$$\mathrm{P}(s) = \left(\sum_{s} n_t^{s,s'} \right) \left(\sum_{s} \sum_{s'} n_t^s n_{t+1}^{s'} \right)^{-1}. \tag{24}$$

Combining (23) and (24) using Bayes theorem we obtain

$$p(s', s) = \mathrm{P}(s'|s) = \frac{\mathrm{P}(s, s')}{\mathrm{P}(s)}. \tag{25}$$

To capture the fall (rise) in informality (formality) from age 20 to 25 that we observe in the data, we also estimate initial conditions for the Markov chain corresponding to the share of formal, informal and unemployed workers at age 20. However, these computations are not enough for the model to replicate the life-cycle dynamics that we observe in the data, since our environment features stochastic retirement and survival, and these two ingredients interact with the share of workers alive in each cohort. We thus adjust the estimated transition probabilities from formal to formal, formal to informal, informal to formal and informal to informal using a calibrating procedure until the model exactly replicates the aggregate shares of formal, informal and unemployed workers that we observe in the data. Table 4 shows the results.

The conditional transition probabilities show that movement from formality and informality to other occupational states are rather persistent. In fact, conditional on being informal, the probability of continuing in informality the following period is 77%, and

Current State	Initial Condition	Future State		
		Informal	Formal	Unemployed
Informal	52.13	0.7677	0.1865	0.0458
Formal	34.40	0.0638	0.9107	0.0254
Unemployed	13.47	0.4909	0.2812	0.2279

Table 4: **Estimated Transition Probabilities for Occupation Status and Initial Conditions.** Initial conditions are measured in percentages, while transitions are probabilities. Values are estimated using Labor National Surveys for various years.

the probability of formality continuation is 91%. We also find that it is more likely for an informal worker to become unemployed (5% probability vs 3% for a formal worker) and that unemployed workers are almost twice more likely to find an informal job than a formal one (49% probability vs. 28%).

These results show the rigidity of the formal sector in Ecuador, and how this rigidity (in the form of minimum wages, firing costs, etc.) exclude most of the working-age population from formal contracts. This lack of formality puts additional pressure over the sustainability of the social security fund, something that we address directly in the design of the optimal tax reform.

4.2.4. Other parameters

Our modeled economy features stochastic retirement. To estimate the probability of retirement we use the same idea that we used to compute the age-specific transition matrices between occupational status and compute the transition between economic activity and retirement. For this we use a question in official labor surveys that reads as follows: *Have you received retirement transfers or pensions?* As opposed to transitions between occupational status, we assume that transitions to retirement are permanent.

Since the survey question we are using includes pensions other than retirement (such as transfers to dependents of deceased contributors), there are probabilities different from zero for young households. We are concerned only with retirement transfers, so we set the retirement probability to zero from age 20 to 54, and then to one from 65 onward. This implies that, effectively, we are using the estimated probabilities for ages 55-64 (see Figure 3). Again, we only use wage-earners and probabilities are computed after simulating the model. For completeness, Figure 3 shows the estimated and simulated retirement probabilities. The difference between the two is that the latter is affected by the probability of survival.

The representative firm in the economy follows a Cobb-Douglas specification with respect to capital and labor, with a CES aggregator for formal and informal labor. In section 4.1 we showed how we pinned down the elasticity of substitution between formal and informal labor as part of the calibration procedure. To identify the other parame-

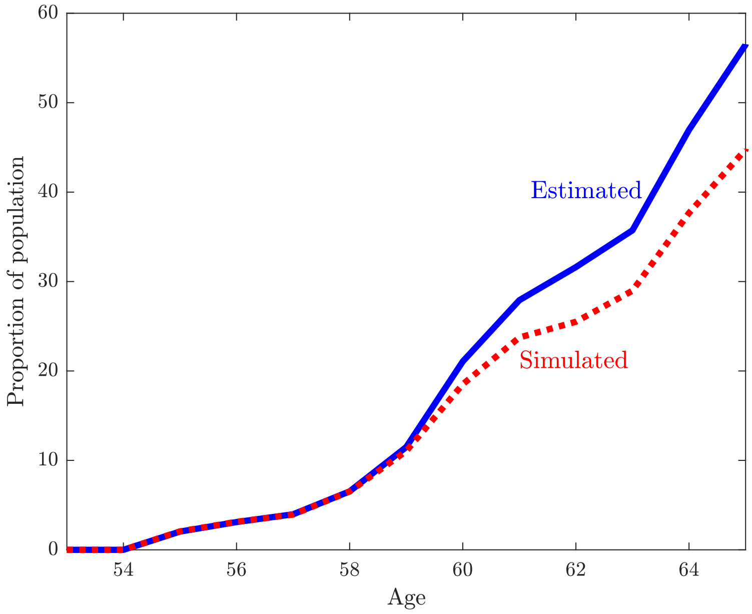


Figure 3: **Age-specific Retirement Probability, From age 53 to 65 (p_j^r).** Simulated retirement probabilities are affected by conditional probability of survival.

ters of the production function, we use administrative firm-level data for formal firms to estimate the following equation

$$\log Q_{it} = \beta_0 + \alpha_0 \log K_{it} + \alpha_1 \log L_{it} + \boldsymbol{\beta}'\mathbf{X}_i + \boldsymbol{\delta}'\mathbf{dt}_t + \boldsymbol{\eta}'\mathbf{di}_i + \mu_{it}, \tag{26}$$

where Q is production measured as total value added, K is the amount of registered assets, L the number of workers and $\mathbf{X}_i$ a matrix of controls at the firm level, $\mathbf{dt}_t$ are time fixed effects and $\mathbf{di}_i$ are firm-level fixed effects. We estimate (33) by a two-step Arellano Bond estimator and set $\alpha = \hat{\alpha}_0$.[9]

Since we are assuming constant returns to scale, each parameter in our production function can be interpreted as the proportion of output that is paid to each input. Thus, the proportion of output paid to formal workers is $\alpha_f(1-\alpha)$, where α_f is the share of labor income paid to formal workers. We estimate α_f from labor surveys. The resulting parameters are presented in Table 5.

Parameter	Description	Value
α	Prop. of output paid to capital	0.4330
α^f	Prop. of income paid to formal	0.6246

Table 5: **Production Function Parameters.** The proportion of output paid to capital is obtained from estimating the model in (33) using a two-step Arellano-Bond estimator. The proportion of income paid to formal workers is computed from National Labor Surveys.

The social security system in our economy manages two transfers targeted to provide partial unemployment and retirement insurance. We assume that both transfers are fully

[9]We provide further details on this estimation in the appendix.

funded by firms' and formal workers' contributions. To setup the proportion of the fund that is used to finance the retirement transfer θ we use information from the Ecuadorian Institute of Social Security (IESS). This information is contained in statistical bulletins were we have data on monthly evolution of beneficiaries and average payments of the different benefits that are managed by IESS.

Based on this data we assume that the only benefits provided by IESS are severance payments (retirement transfers) and unemployment insurance. To compute θ_t, we simply add the total amount used in these transfers and then divide the sum of severance payments by the total of both benefits. We do this using data for 2016 aggregated for the full year and find that $\theta_t = 0.9223$; that is, 92% of the fund is used to cover severance payments.

We also assume that government expenditure is determined as a fixed proportion of total output g. To estimate this parameter, we use annual time series of national accounts for Ecuador from 1965 to 2014. We compute g as the long-term average of the proportion of total Government consumption with respect to GDP and find $g = 0.1358$. Keep in mind that we are only considering current consumption of the General Government in order to compute this ratio.

Finally, the progressive tax scheme applied to labor income requires us to calculate the ratios of the extreme values of the tax scheme y_l and y_h with respect to average income in the economy.[10] Let $\omega_1 \equiv \underline{y}_t/y_t^m$ and $\omega_2 \equiv \overline{y}_t/y_t^m$ where y_t^m is the average of pre-tax labor income of formal workers. For 2017 the progressive tax scheme in Ecuador set (in current USD) $y_l = 11,290$ and $y_h = 115,140$ which, given the average labor income that we observe in the household surveys for that year imply $\omega_1 = 0.8001$ and $\omega_2 = 8.1612$. A summary of these parameters is presented in Table 6.[11] A direct implication of these estimations is that the proportion of households that actually pay labor income taxes in the economy is very low (about 27% of formal workers according to Gachet et al., 2018).

Parameter	Description	Value
θ	Prop. of SS fund used for severance transfers	0.9223
g	% of GDP dedicated to Government consumption	0.1358
ω_1	Lower bound of tax scheme with respect to average income	0.8001
ω_2	Upper bound of tax scheme with respect to average income	8.1612

Table 6: **Other Estimated Parameters.** The value for θ is estimated from statistical bulletins published by the Ecuadorean Institute of Social Security (IESS). The proportion of GDP dedicated to government consumption is estimated as the long-run average using annual time series of National Accounts for Ecuador from 1965 to 2014. Values for ω_1 and ω_2 are estimated using National Labor Surveys.

[10] We do this for simplicity. However, Kindermann and Krueger (2014) relate y_l to the median of the income distribution and y_h to the mean.

[11] The estimated values of all age-specific parameters are available from the authors upon request.

4.3. Numerical solution

We solve the model over a discretized state space using an exponential grid for assets with 21 grid points. The idiosyncratic productivity shock is discretized using a 7-state Markov Chain where grid points and the transition matrix are computed following Tauchen and Hussey (1991). The households' optimization problem is solved using a combination of discretized grid search VFI and interpolated VFI for formal and informal workers, and the Endegenous Grid Method for retirees and unemployed workers. We solve the equilibrium by a simple bisection method. We provide more details on the algorithms in the appendix.

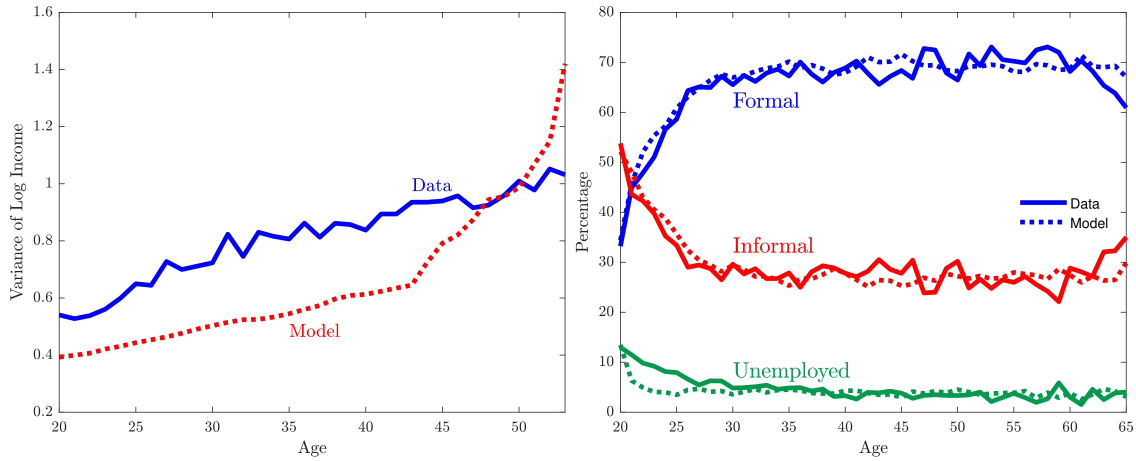
The solution to the calibration minimization problem is obtained by applying the Generalized Pattern Search (GPS) algorithm (Audet and Dennis Jr, 2002) to a weighted quadractic loss function that puts more weight on the capital/output ratio and hours worked targets. All functions involved in the model are evaluated using piece-wise cubic hermite interpolating polynomials (PCHIP) (Fritsch and Carlson, 1980). We use Montecarlo simulations to compute the invariant distribution.

Before turning to the policy experiments, we end this section by assessing the fit of our model to the data and describing the benchmark economy. Table 7 compares the model and the data for the targets we chose for calibration and the percentage of informality, and shows that we have been successful in the calibration procedure. The model, however, produces less inequality compared to what it is observed in the data, and a lower formality wage premium.

Target	Model	Data
Capital-output ratio	3.230	3.245
Hours worked	0.362	0.340
Variance of log income at 20	0.393	0.495
Variance of log income at 50	0.987	0.934
Informality	29.7%	29.6%
Formality wage premium	1.695	1.970

Table 7: **Calibration targets.**

This issue in replicating the observed labor income inequality is evident in Figure 4a. In spite of having enough flexibility, the model generates low variances at age 20, a linear increase up to age 43, and exponential growth from age 44 onward. This is because there are many channels (endogenous and exogenous) that might push inequality in opposite directions in equilibrium, making this calibration particularly complicated. For example, increasing the variance of the idiosyncratic productivity shock encourages households to save and work more hours. This would call for lower values for the discount factor β and higher values for the utility weight of leisure χ, affecting income variability in the opposite direction. In the end, the optimizing algorithm must do its best to balance the two effects, generating a less than perfect fit.



(a) Variance of log income by age

(b) Occupation status by age

Figure 4: **Labor Income Distribution and Percentage of Workers in Each Occupation Status Through the Life Cycle.** Empirical values are estimated using National Labor Surveys.

The good news, however, is in Figure 4b where we show the percentage of formal, informal and unemployed workers along the life cycle generated by the estimated Markov chain that is fed to the model. Our estimations allow the model to replicate a fitted version of what we observe in the data. In particular, the model correctly generates the significant increase in the proportion of formal workers that occurs between ages 20 and early 30s. The model also produces the significant drop in informality during the early years in the labor force and the observed unemployment dynamics at the end of the life cycle; yet, it generates a rapid drop at the beginning of the life cycle –something that does not occur in the data– implying that unemployment persistence may vary with age.

5. The Optimal Reform

We propose a major tax reform that relies on the power of age-dependent taxation to tackle three important problems that most emerging markets face:

1. Demographic change and informality in the labor market make pay-as-you-go social security schemes unsustainable,

2. Inequality is high and state capacity is low, so optimal tax schemes should be efficient and redistributive,

3. Status-quo tax policy in many emerging markets ends up being extremely complicated, tax compliance of personal income taxes is very poor (again, due to high levels

of informality in the labor market, other income-sheltering activities and tax evasion), tax collection relies on broad-base taxes (such as the VAT), so the final result is that the whole tax system becomes regressive or neutral in the best case-scenario.

To achieve this goal, the reform incorporates two key ingredients. First, to solve potential sustainability issues that the pay-as-you-go social security system might exhibit, we move towards a system of private, individual accounts. Second, to simplify the current personal income tax scheme, we replace the progressive labor income tax scheme and the capital income flat tax rate with a single, optimal labor income age-dependent flat tax rate.

Due to the nature of the reform, we need to modify the problem that is solved by the formal worker and the definition of the competitive equilibrium. We do maintain, however, the social security transfer for the unemployed that were previously hired in the formal sector. For this, we take a proportion $(1-\theta_t)$ of the amount collected through the age-dependent tax rate to fund these transfers. The rest of the revenue is transferred to finance government's consumption. With these considerations, the new program for the formal worker is given by,

$$\begin{aligned} & V_t(x_t^f) = \max\left\{U(c_t,\ell_t) + \beta\psi_j \,\mathrm{E}_t\, \tilde{V}_{t+1}^f(x_{t+1})\right\}, \\ \text{s. t.} \quad & a_{t+1} = (1+r_t)a_t + (1-\tau_{jt})y_{jt}^f + Tr_t + G_t - (1+\tau_t^c)c_t, \\ & a_{t+1} \geq \max\{0, b_t^{ss}\}, \quad c_t > 0, \quad 0 \leq \ell_t \leq 1, \end{aligned} \tag{27}$$

where

$$\begin{aligned} \mathrm{E}_t\, \tilde{V}_{t+1}^f(x_{t+1}) \;=\; & (1-p_j^r)\left[\int \sum_{s'=i,f} p(s',f)V_{t+1}(x_{t+1}^{s'})F_\eta(d\eta_{t+1}|\eta_t)\right. \\ & + \left. p(u,f)V_{t+1}(x_{t+1}^u)\right] + p_j^r V_{t+1}^r(x_{t+1}^r), \end{aligned}$$

τ_{jt} is the age-dependent flat tax rate levied on labor income, and b_t^{ss} is the minimum savings requirement introduced by the private accounts pension system.

For the retirees, the optimal reform implies that they no longer have access to a transfer originated in the pensions fund, so all their consumption must be mainly financed by their own savings, so the new problem is

$$\begin{aligned} & V_t(x_t^r) = \max\left\{u(c_t) + \beta\psi_j V_{t+1}(x_{t+1}^r)\right\}, \\ \text{s. t.} \quad & a_{t+1} = (1+r_t)a_t + Tr_t + G_t - (1+\tau_t^c)c_t, \\ & a_{t+1} \geq 0, \quad c_t > 0. \end{aligned} \tag{28}$$

For unemployed and informal workers the changes in their optimization problems are marginal. In particular, the main modification is now that they do not need to pay taxes on their capital income. To avoid repetition, in the following definition we concentrate on stating the necessary modifications for Definition 3.1 to be compatible with the tax reform.

Definition 5.1 (Modified Competitive Equilibrium)**.** *Consider the conditions for the competitive equilibrium in Definition 3.1. Under a private accounts social security system and age-dependent labor income taxation, a competitive equilibrium are sequences $\{\tau_{jt}\}_{t=1}^{\infty}$ and $\{b_t^{ss}\}_{t=1}^{\infty}$ that, besides household's policy functions, firm's production plans, prices, transfers and measures we have that:*

(i) Formal workers and retirees maximize life-time expected utility, so $\{V_t^s, c_t, a_t, \ell_t\}_{t=1}^{\infty}$ for $s \in \{f, r\}$ solve modified problems (27) and (28),

(ii) The transfer for former formal employees during unemployment spells are fully funded by the age-dependent labor income tax revenues, so

$$\lambda_t^u \int \Phi_t(da_t, \cdot, u, f, dj) = (1-\theta_t) \int \tau_{jt} \varepsilon_j^f \eta_t w_t^f \ell_t(x_t^f) \Phi_t(dx_t^f), \tag{29}$$

(iii) The government runs a balanced budget,

$$G_t = \tau_t^c C_t + \theta_t \int \tau_{jt} \varepsilon_j^f \eta_t w_t^f \ell_t(x_t^f) \Phi_t(dx_t^f). \tag{30}$$

All other conditions not explicitly stated in Definition 5.1 are applicable with minor (or no) modifications. Of course, the Definition 3.2 applies directly.

Now we are ready to setup the optimization problem faced by the social planner. We restrict choices to the age-dependent tax scheme that maximizes social welfare in the steady state. For this, we assume a utilitarian welfare function in the planner discounts life time utility of future generations at rate β, so social welfare is equal to average utility in the cross section (Heathcote et al., 2020). Formally,

$$\mathcal{W}(\{\tau_j\}) = \int \frac{c(x)^{1-\sigma_1} - 1}{1-\sigma_1} + \chi \frac{(1-\ell(x))^{1-\sigma_2}}{1-\sigma_2} \Phi(dx), \tag{31}$$

where we have dropped the time subscripts since the welfare function is defined in the stationary equilibrium. The Ramsey problem solved by the planner is to choose a sequence $\{\tau_j\}_{j=1}^{j^r}$ to maximize (31).

We allow for τ_j to be positive (tax rate) or negative (transfer rate). This characteristic implies that we need to impose an additional restriction to the Ramsey problem. In particular, we need for the age-dependent fund to be financed at all times, otherwise, it can potentially produce consumption tax rates that are so high, that households will have to sacrifice most of their income just to be able to pay consumption taxes.

Since households face various sources of uninsured idiosyncratic risk, the planner has strong incentives to generate massive age-dependent transfers for young workers. This is the main mechanism that can produce an unbalanced age-dependent funding scheme, that could trigger extremely high consumption tax rates. To prevent this from happening, we require that the size of the largest transfer never exceeds the largest tax rate, so

$$|\{\min\{\tau_j\}\}| \leq \max\{\tau_j\}.$$

In solving the maximization, we make sure that each element triggers positive welfare variations. We thus proceed in steps. First, we move towards the private accounts social security system and check that the welfare variation is positive. Then, we eliminate labor and capital income and, lastly, we optimize over the age-dependent tax rates from an initial condition in which τ_j increases linearly from $-\exp\{-3\}$ to $+\exp\{-3\}$. To make sure that eliminating the previous PIT tax scheme is optimal, we define the optimization problem over these parameters as well. We solve this optimization using the Generalized Pattern Search (GPS) algorithm (Audet and Dennis Jr, 2002).

6. Results

The optimal tax reform changes the social security system from a pay-as-you-go scheme to a system of private accounts without minimum savings requirement, and replaces both schedules of personal income taxation (progressive labor income and flat capital income tax rate) with an age-dependent labor income flat tax rate. In Figure 5 we present the optimal age-dependent labor income tax rates and compare this schedule with the status quo. To compute the status quo we add the age-specific cross-section average of the effective labor income marginal tax rate, the social security contribution and the capital income flat tax rate.

There are two salient differences between the status quo and the optimal tax schemes. The first one is that the optimal tax reform implies a significant reduction in the tax burden: In the status quo, formal workers pay, on average, a combined marginal tax rate of 36.8% while under the optimal reform, those who pay a positive tax rate pay, on average, 20.4%. The second one is that, even though the status quo tax scheme presents some degree of variation along the life cycle, it is very mild compared to what is obtained under the optimal tax reform. The reason for this is that in the status quo, the variation over the life cycle comes from the differences in age-specific productivity, that immediately translates into variations in the labor income age profile, thus affecting effective marginal tax rates.

There is a third, more subtle, difference between the two schemes, and its related to the percentage of formal workers that contribute through positive labor income tax rates.

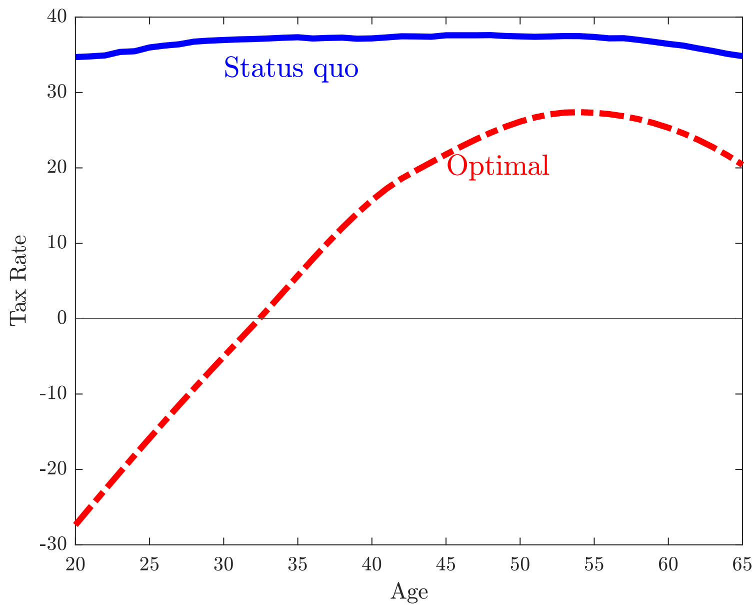


Figure 5: **Optimal Age-dependent Tax Schedule.** All tax rates are presented in percentages. Status quo corresponds to the sum of the cross-section average by age of effective marginal tax rates of labor income, the social security contribution and the flat tax rate on capital income. In the optimal, all these taxes are replaced by a single age-dependent flat tax scheme.

In the status quo, only 57.7% of formal workers pay marginal tax rates that are higher than zero under the progressive tax scheme, and the average marginal tax rate is about 6.3%. On the other hand, under the optimal reform, 70.4% of formal workers contribute with positive age-dependent tax rates, and the overall average tax rate (considering positive and negative rates) is 10.7%. In the end, while the average tax burden faced by all households in the economy is 13.0% in the status quo, it falls to 8.7% under the optimal reform, without considering consumption taxation. The same occurs when we only consider formal workers, where the tax burden falls from 13.0% to 10.9%. Keep in mind that, under the optimal tax reform, the tax scheme includes positive and negative tax rates, and we are considering both in these computations.

Next we study the welfare and aggregates effects triggered by the optimal tax reform. Since the reform encloses major tax cuts and moving to a social security system based on private accounts, despite the long-run welfare gain we might observe severe welfare losses in the short run. Because of this, we analyze the effects of the reform in the long run, and we also study the transitional dynamics.

6.1. Long-term Results

In this section we study how the economy is affected by the reform in the long run. We begin with this analysis because the reform is designed to maximize welfare in steady state. First we concentrate in analysing what occurs to the main aggregates, and then we study the sources of the the welfare gain and the changes triggered over the life cycle profiles.

6.1.1. Main Aggregates and Life Cycle Profiles

In Table 8 we present percentage variations comparing the steady state under the optimal tax reform with the steady state of the status quo. In the long run, the optimal reform produces a welfare gain of 6.1% measured in percentage of compensated equivalent variation. All the gain occurs for working-age population, and its triggered by the significant tax cut implied by the reform. This tax cut, together with the negative rates setup for young, formal workers, increase wealth and disposable income, generating a long-term welfare effect for workers. This way, consumption increases by 6.7% on average, and hours worked fall by 4.9%.

Variable	% Variation
CEV	$+6.1$
Consumption	$+6.7$
Formal	$+10.5$
Informal	$+14.6$
Unemployed	$+15.8$
Retired	-13.5
Hours Worked	-4.9
Formal	-2.9
Informal	-10.5
Labor Supply	-3.6
Formal	-2.5
Informal	-10.0
Capital	$+38.2$
Output	$+13.0$
Consumption Tax	$+1.5$
Gini Coefficients	
Wealth	-8.1
Consumption	-1.8
Disposable Income	$+4.5$

Table 8: **Variations Triggered by Application of Optimal Social Security Contributions Scheme.** Compensated Equivalent Variation (*CEV*) and Aggregate Variations are measured in percentages and correspond to comparisons between steady states. The Gini coefficient variation of disposable income considers working-age population labor income after taxes and transfers.

When we analyze the variations in consumption conditional on occupation status, its clear that the reform implies a significant redistribution of consumption along the life cycle. In this regard, while consumption for formal, informal and unemployed workers increase by 10.5%, 14.6% and 15.8% respectively, consumption among retirees falls about 13.5%. There are two drivers for this redistribution. First, the implementation of private accounts to finance consumption during retirement and, second, the negative tax rates for workers that are just entering the labor market, and that last until workers turn 33.

The wealth effect triggered by the reform, moreover, affects the households' labor supply both under formal and informal contracts. Nonetheless, while formal hours worked fall by 2.9%, informal hours worked plunges 10.5%, provoking a formalization of labor supply. The main channel behind this result is that the optimal age-dependent tax scheme subsidizes formal labor for young cohorts.

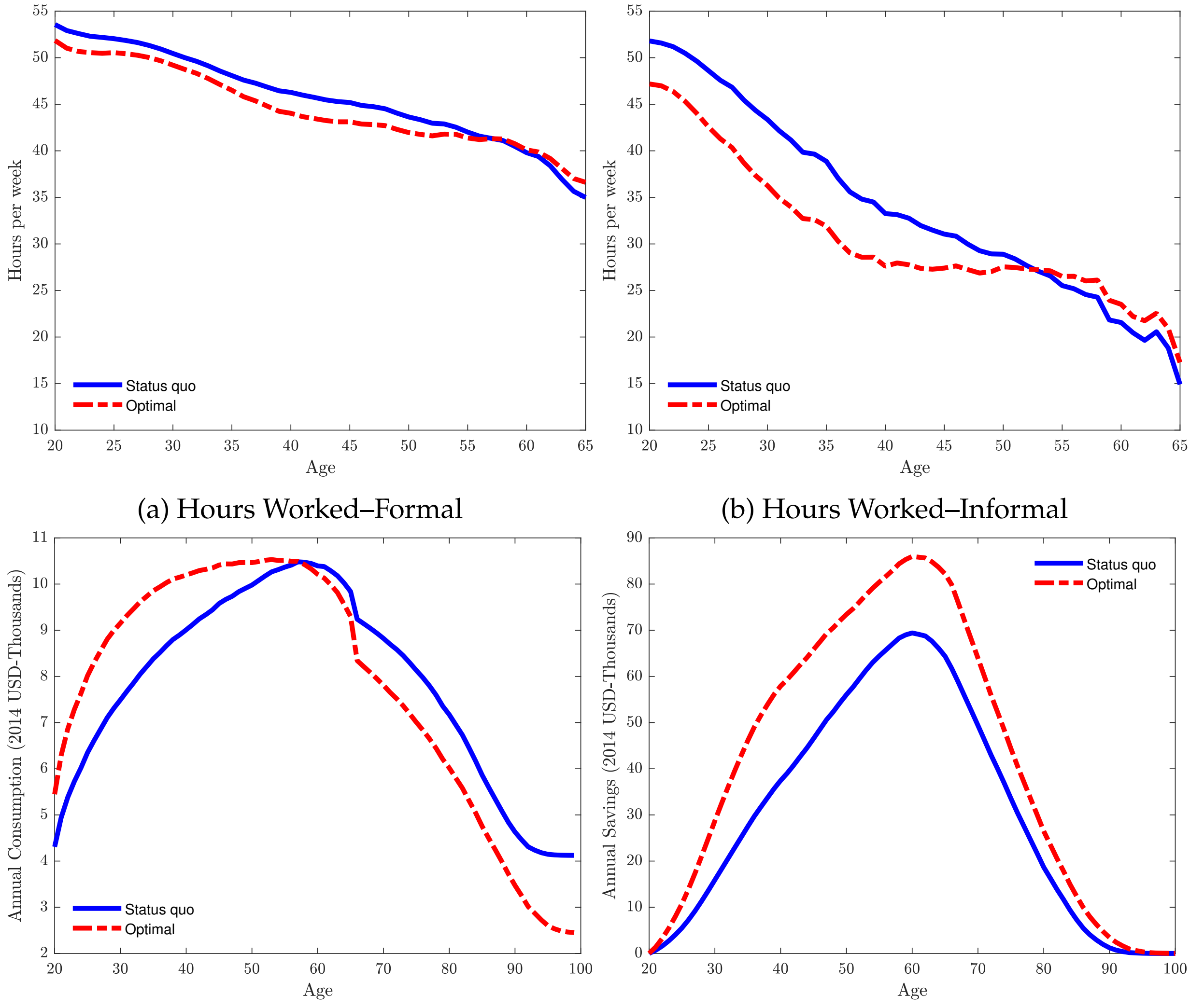
This translates directly into effects in the labor supply. Even though hours worked fall 4.9% with respect to the status quo, labor supply reduction is only 3.6%, implying that hours worked in the post-reform steady state are slightly more productive. These are important results, since they imply that the optimal reform is effective in reducing the level of informality in the labor supply that arises in equilibrium.

The social security part of the reform, together with the abolition of capital income taxation, produce a 38.2% increase in the size of the capital stock in the economy. Since retirement, and thus longevity risk, is no longer insured, households need to accumulate more assets to self-insure against these shocks. This result is in line with previous literature where it is recognized that, when age-dependent taxation is available, capital taxes become redundant (Conesa et al., 2009).

The reason for this is that, if age-dependent taxation is not a possibility, capital taxation becomes optimal because young cohorts tend to over-save in early years to build the stock of assets that will allow them to self-insure against future negative shocks (Aiyagari, 1995; Conesa et al., 2009). By taxing capital, the return on savings falls, thus creating incentives for households to increase consumption. In the presence of age-dependent taxation, on the other hand, the planner can alleviate the tax burden of younger households through negative tax rates, effectively subsidising labor and increasing disposable income. This, in turn, produces a durable income effect that allows households to increase consumption and leisure, thus improving welfare significantly.

The incipient contraction of labor supply and the massive increase in the capital stock produce a 13% increase in output. This result shows how efficient age-dependent taxation is in comparison to the status quo tax scheme, a result that has been repeatedly addressed in previous literature (e.g. Mirrlees, 1971; Weinzierl, 2011; Heathcote et al., 2020).

In terms of tax revenue, the optimal reform performs a major leap in terms of efficiency: Besides funding targeted transfers for young formal workers and unemployment insurance, still contributes about 98.5% of tax revenue that was before generated by the



(a) Hours Worked–Formal

(b) Hours Worked–Informal

(c) Consumption

(d) Savings

Figure 6: **Life Cycle Profiles.** Panels for consumption and savings present the age-specific cross-section averages. Effects for formal, informal and unemployed workers are homogeneous.

personal income tax scheme in the status quo. Because of this, in the optimal tax reform, the consumption tax rate (only) increases by 1.5%.

As we stated at the beginning of this section, reforms in emerging markets are not only concerned with efficiency, but also need to tackle inequality. The last panel in Table 8 presents the long-run variations in the Gini coefficients for wealth, consumption, and after-tax labor income. The reform cuts wealth and consumption inequality by 8.1% and 1.8% respectively. Once again, this result shows the great potential of age-dependent taxation. This reduction in inequality is mainly explained by the inter-generational redistribution triggered by the reform. Disposable income inequality, however, increases by 4.5%, mainly due to the elimination of the progressive labor income tax scheme. This could be corrected if the reform also allows for age-specific progressivity for the labor-income tax scheme, as in Heathcote et al. (2020).

To have a better grasp at the structure of the intergenerational redistribution triggered by the reform in the long run, in Figure 6 we present the age profiles for hours worked (formal in panel 6a and informal in panel 6b), consumption (panel 6c), and savings (panel 6d) comparing the pre- and post-reform steady states invariant distributions. Intergenerational resdistribution is evident in the age profiles for consumption and hours worked. In both cases we observe the wealth effect triggered by the reform at work, increasing consumption and leisure for young workers, with the opposite behaviour for older cohorts. It is also interesting to notice that the supply of formal hours for labor seems to be more inelastic than informal labor supply.

The age-profile of savings, on the other hand, clearly reflects the elimination of the pay-as-you-go social security system and the indirect insurance triggered by the formal labor subsidy implied by the age-dependent tax scheme. The former is the channel that explains the generalized increase in savings that we observe for all ages. The latter is responsible for how fast assets are accumulated during the first 13 years in the labor market, which is exactly the time span in which the age-dependent tax scheme assigns negative tax rates.

6.1.2. Welfare Decomposition

It is clear that the main driver of the welfare gain in the long run is the increase in consumption. However, it is less clear how the intergenerational redistribution of consumption and leisure affects the welfare variation. To have a better understanding of how the redistribution of consumption and leisure affects aggregate welfare, we follow Conesa et al. (2009), and decompose the welfare gain into each of its sources and whether it comes from a variation in levels or from changes in the distributions of consumption and leisure. Given our choice for the utility function, the compensated equivalent variation is given by

$$CEV = \left\{ \frac{(1-\sigma_1)\left[W(c_0,\ell_0) - \chi/(1-\sigma_2)(1-\ell_0)^{1-\sigma_2}\right]+1}{(1-\sigma_1)\left[W(c^*,\ell^*) - \chi/(1-\sigma_2)(1-\ell^*)^{1-\sigma_2}\right]+1} \right\}^{\frac{1}{1-\sigma_2}} - 1,$$

where $W(c,\ell)$ is total welfare given by a specific choice of consumption c and leisure $(1-\ell)$. In general terms, this expression measures the percentage variation in consumption needed to maintain welfare constant in its pre-reform level. From here, consumption CEV_c and leisure CEV_l compensated equivalent variations are defined as

$$\begin{aligned} W(c^*,\ell_0) &= W(c_0(1+CEV_c),\ell_0), \\ W(c^*,\ell^*) &= W(c^*(1+CEV_l),\ell_0), \end{aligned}$$

where $CEV \approx CEV_c + CEV_l$. To see the intuition behind this result, notice that CEV_c measures the variation in consumption needed to compensate consumers for the tax reform *keeping leisure constant*, while CEV_l does the same but starting from the level of welfare obtained by the optimal consumption level c^*, thus measuring only the welfare compensation explained by changes in leisure.

Moreover, within each specific equivalent variation it is possible to pin down the level and distributional effects of the tax reform. In particular, the level effect for the consumption-specific compensated equivalent variation is given by the growth rate of aggregate consumption $CEV_{c\gamma} = C^*/C_0 - 1$, so the distribution effect is the residual $CEV_{cd} = CEV_c - CEV_{c\gamma}$. A similar decomposition can be applied to leisure and we further expand these computations to condition equivalent variations on the workers' occupational status. The results of these calculations are presented in Table 9 for the full reform (column 1) and considering the social security and tax reforms separately (its marginal contribution –columns 2 and 4– and the accumulated welfare gain in each step –columns 3 and 5).

	% Variation				
	Full Reform	**Only SS**		**Replace PIT**	
	Total	**Mg.**	**Acc.**	**Mg.**	**Acc.**
CEV	+6.1	+1.8	+1.8	+4.3	+6.1
Consumption	+7.7	+3.7	+3.7	+4.0	+7.7
Level	+6.7	+4.1	+4.1	+2.6	+6.7
Distribution	+1.1	−0.3	−0.3	+1.4	+1.1
Leisure	−1.6	−1.9	−1.9	+0.3	−1.6
Level	+2.8	+0.7	+0.7	+2.1	+2.8
Distribution	−4.4	−2.6	−2.6	−1.8	−4.4

Table 9: **Welfare Decomposition.** Decomposition is performed over consumption and leisure by component of the full reform. The first component corresponds to an economy where only the pay-as-you-go social security system is replace by private accounts. The second component is when, besides the social security reform, the entire PIT scheme is replaced by the age-dependent flat tax schedule. Columns within each component correspond to the marginal and accumulated welfare variation, respectively.

Overall, we observe that the total welfare gain is the result of a 7.7% increase in welfare explained by its consumption component, and a welfare loss of –1.6% due to the leisure component. The result in consumption is not surprise, and going one step forward shows that the consumption welfare gain is the aggregation of a positive variation of 6.7% due to an increase in aggregate consumption, and a 1.1% welfare gain due to better distribution. These results imply that the implied intergenerational redistribution of consumption along the life cycle is more efficient than the distribution in the status quo.

The results on the leisure component, on the other hand, are not straightforward. The reason for this is that in the aggregate we observe a reduction in hours worked (which translates in more leisure, on average) and, nonetheless, the decomposition shows a 1.6% welfare loss due to leisure. Further decomposing this loss, we find that the increase in aggregate leisure produces, in fact, a 2.8% welfare gain, but this is shadowed by a 4.4% welfare loss explained by the change in the distribution of leisure.

Before turning to the analysis of life-cycle profiles, it is interesting to study how each component of the reform affects welfare. In columns 2 and 3 of Table 9 we present the marginal and accumulated welfare variation, respectively, that can be explained only by replacing the pay-as-you-go social security scheme for a system of private accounts without minimum savings requirements. The overall welfare gain of this part of the reform is of 1.8%, and in this case is clear that the distribution of consumption is an issue, since the partial reform triggers a welfare loss of 0.3% in the long run. The reason for this welfare loss is that the social security reform eliminates the social security transfer for retirees and households need to sustain their consumption during retirement with their own savings. Since households are impatient, however, consumption is pushed forward in the life-cycle, sacrificing consumption at older ages.

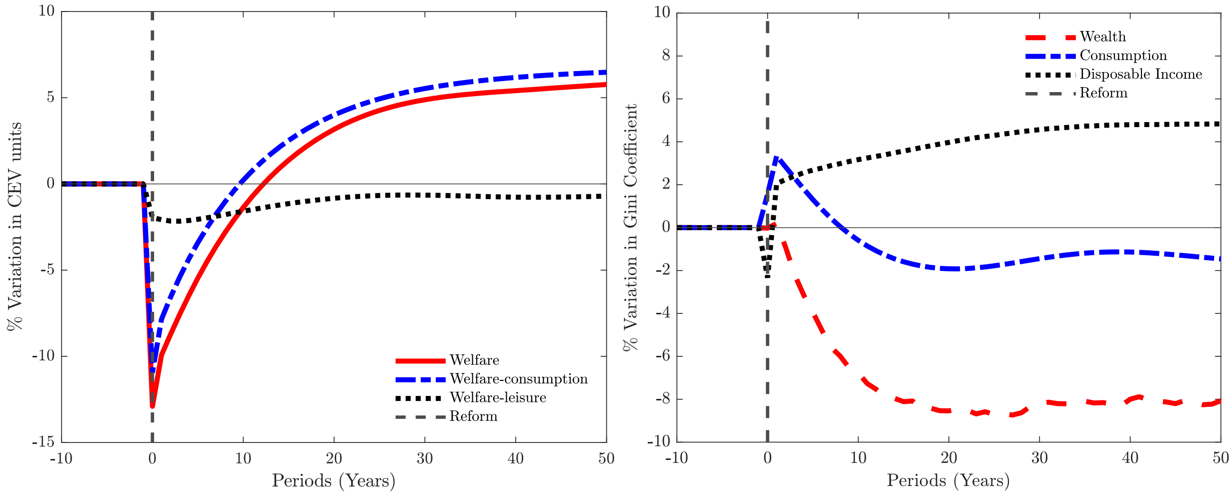
When the social security reform is complemented with the PIT reform and, in particular, with the replacement of the PIT scheme with age-dependent labor income taxation, the reform triggers a marginal welfare gain of 4.3%, and welfare due to consumption distribution compensates the welfare loss generated by the social security reform (gain of 1.4%). The welfare gain comes from liberating income for retirees once capital income taxation is eliminated, and the system of early transfers is in place, allowing households to accumulate savings faster earlier in the life cycle.

6.2. Transition Dynamics

In the previous section we studied the long-run effects of the optimal reform. Due to its nature, however, it is important to analyze the short run effects, since this will allow us to identify affected groups of the population, and inform the designed of contingency policies.

For this, we assume that the policy is applied unexpectedly with a once-and-for-all simultaneous change to the social security system and the PIT schedule. We assume, moreover, that there is perfect-foresight in the economy, so there are no sources of aggregate uncertainty. Keep in mind that we do not optimize during the transition path, but only in the long run. This is why we might observe significant welfare losses in the short run.

In Figure 7 we present aggregate transitional dynamics for welfare and inequality; in both cases, period zero is the period in which the reform occurs. In Panel 7a we present the welfare decomposition that we discussed in the last part of the previous section during the transition path. We are able to do this because we assume that the social planner discount expected life-time utilities of new generations with the same discount rate as the households. Moreover, our model does not feature any ex-ante irreversible invest-



(a) Welfare variation

(b) Inequality

Figure 7: **Welfare variation and inequality during the transition.** Welfare decomposition over the transition measured in compensated equivalent units. Since planner discounts new generations' life time utility with the same discount rate as households, it is possible to perform the same decomposition as in the steady state. Variations are computed with respect to the pre-reform steady state. The inequality dynamics correspond to percentage variation of the Gini coefficient of each variable with respect to the pre-reform steady state along the transition.

ment, so welfare along the transition can be measured the same way as in the steady state (Heathcote et al., 2020).

As expected, the reform produces a 12.8% welfare loss on impact (compensated equivalent variation), and most of this contraction is explained by a severe fall in consumption triggered, mainly, by the reform to social security. We find that it takes about 14 years for the reform to produce positive welfare variations with respect to the pre-reform steady state. The participation of the leisure component in the welfare variation during the transition is stable, and it converges to the long-run equilibrium in about 30 years, much faster compared to the other components.

Short term effects over inequality are also driven by the elimination of the pay-as-you-go social security system, particularly regarding consumption inequality. To see this, in Figure 7b we present percentage variations of the Gini coefficients for assets, consumption and disposable income during the transition. On impact, the elimination of the social security contributory scheme increases consumption inequality in 3.3%.

The mechanism behind this result is the nature of the reform: Older cohorts caught during the period of reform have no chance to adjust savings, so their asset stock internalizes the social security transfer that was expected during their retirement. Since the reform is unexpected, they suddenly need to cut back on current consumption, while young formal workers enjoy negative tax rates implemented by the PIT age-dependent tax reform. Consumption inequality thus increases, showing higher Ginis for 9 years after the reform. As time goes by, however, older cohorts die and new cohort enter the economy with the understanding that retirement consumption depends entirely on the asset stock that they manage to accumulate during their life cycle. This reduces consumption dispersion, thus reducing inequality over time.

Regarding disposable income (labor income of working-age population after taxes and transfers), inequality falls 2.3% on impact, but quickly bounces back one year after, and then converges monotonically to its value in the long run. As we explained before, the increase in disposable income inequality is expected, as we are replacing a progressive tax scheme with an age-dependent flat tax schedule.

The dynamics of wealth inequality present monotonic convergence to its optimal steady-state value. In this case, all cohorts immediately adjust savings after the reform, so cross-section dispersion tends to fall as time advances.

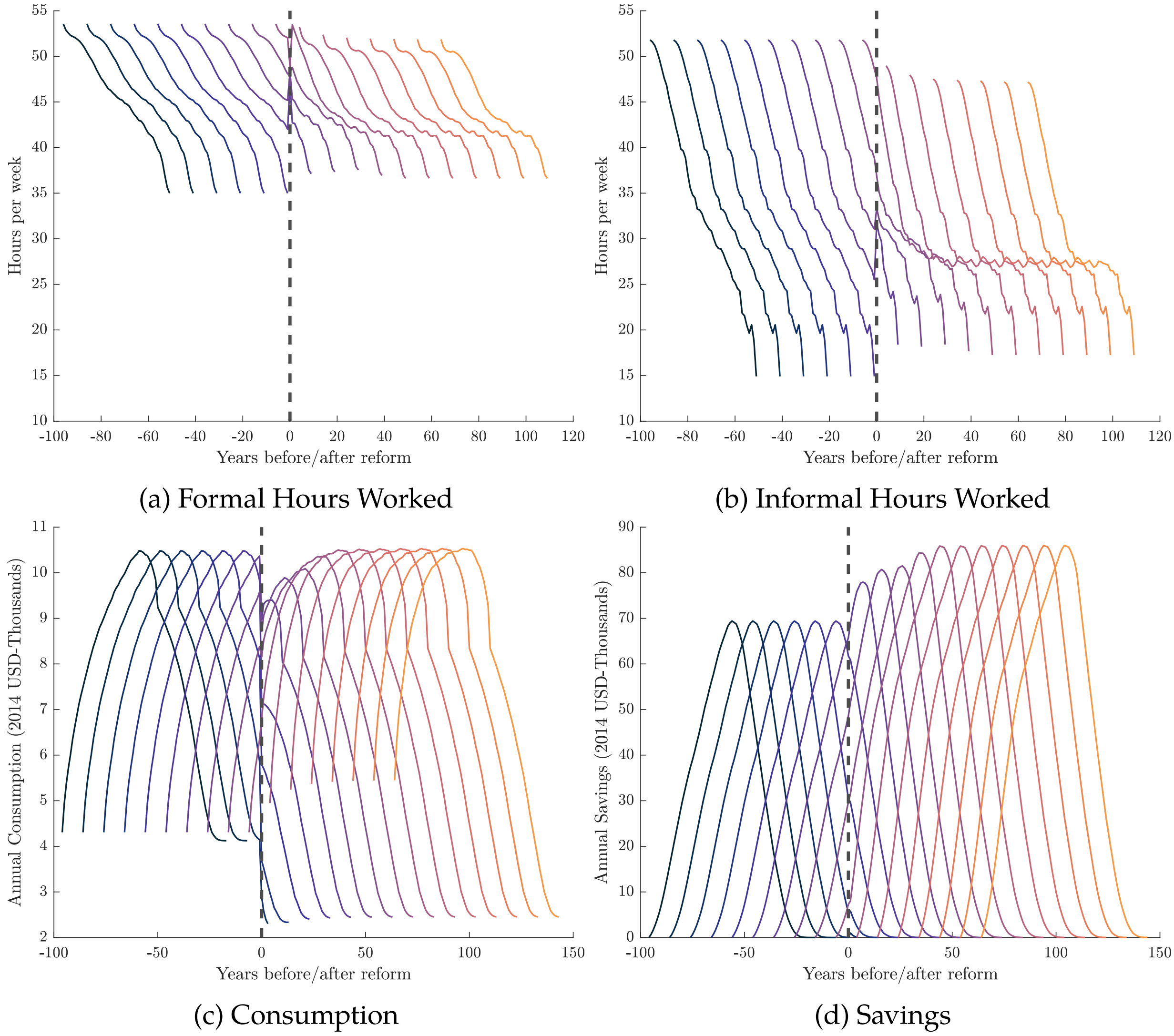
We finish this section by studying the dynamics of hours worked, consumption and savings for different cohorts during the transition. This exercise provides valuable information, since it allows us to understand heterogeneous transitions that different cohorts exhibit after the reform. We do this in Figure 8 for cohorts born 95 years before the reform through 65 years after, every 10 years; lighter colors representing younger cohorts.

The effect of the reform is very heterogeneous across different cohorts. In Panel 8a we present the life-cycle profiles of formal workers during the transition. In the period of the reform, we observe that there is a sizable increase in formal hours worked, and the effect is stronger for younger cohorts. Once the optimal policy is in place, younger cohorts adjust their labor supply downward because of the income effect triggered by the negative tax rates implied by the age-dependent tax scheme, while older cohorts face a period of increasing hours worked, that later in time falls to the new steady-state value (which is lower than the transition, but higher compared to the previous steady state).

The transition dynamics on the informal sector are similar to the formal sector in the sense that labor supply increases on impact for all cohorts alive during the period of the reform, but the effect is stronger for older cohorts. As time advances, the dynamics of labor supply for younger and older cohorts is similar to what occurs in the formal sector, again reflecting the income effect implied by the reform (see Panel 8b).

This increase in hours worked that we observe on impact explains part of the welfare loss that we observe during the transition, since in all cases we observe a reduction in leisure, which is particularly concentrated in younger formal workers. The reason for this behavior is that younger workers are able to quickly internalize the absence of the social security transfer when they reach retirement, so they adjust their decisions to accelerate capital accumulation and self-insure against the longevity risk.

The effect over consumption, on the other hand, is much more heterogeneous depending on the age of the cohort at the moment of the reform (Panel 8c). Very young cohorts benefit from the tax cut (and the age-dependent transfer that is part of the reform), so they increase consumption faster compared to what would occur in the absence of the reform. Older cohorts, on the older hand, need to severely reduce consumption,



(a) Formal Hours Worked

(b) Informal Hours Worked

(c) Consumption

(d) Savings

Figure 8: **Life Cycle Profiles during the transition.** In each case, lighter colors represent younger cohorts. Period 0 corresponds to the year of the reform, which is also marked with the vertical dotted line.

thus increasing the speed of assets accumulation to try to self-insure as much as possible for longevity risk. Years after the reform, it is evident how the new tax structure benefits younger cohorts, because they are able to increase consumption during all the transition; and something similar occurs with middle-aged workers. Retirees, on the other hand, practically suffer all the consumption contraction that is evident in the long run on impact.

This discussion shows that, for the reform to be feasible in the long run, a compensation mechanism has to be designed to smooth the transition for those groups of the population that are more negatively affected by the reform. A typical compensation mechanism for this type of reforms (we refer particularly to the social security reform) consists on charging a temporary tax on savings stocked on private accounts to finance retirement transfers for those who where close to or already retired. Another option, that does not affect private accounts directly, is to finance these retirement transfers with part of the tax revenue collected with the age-dependent tax scheme. In this case, one sacrifices government revenue that translates into a temporary increase of consumption taxes to finance government consumption during the transition.

Which option is better, for how long have to be in place, and how much welfare loss can be avoided during the transition, are important issues that could be addressed with a dynamic optimization problem to be solved by the planner. Due to computation power restrictions, however, this problem is out of the scope of this paper.

7. Sensitivity Analysis

The main ingredient in the benchmark model is stochastic occupation transitions between the formal and the (less-productive) informal sectors, and unemployment. Broadly speaking, this is just an additional source of idiosyncratic risk, and one could ask how this shock is any different from a simple idiosyncratic productivity shock enhanced with some additional ingredients. In fact is no different but, as we show in this section, its an important part of our model and key to explain some of our most salient results.

To check the relevance of occupational risk, we build an alternative model in which we only consider the formal sector. For this, we build an economy with full employment and discipline the model using data corresponding only to formal workers, when necessary. In building this alternative setting, we do our best to keep it as comparable as possible to the original economy with occupational risk. The main additional ingredient in the full-employment economy, however, is a fixed-ability shock determined at birth, which is necessary to replicate the variance of log-income age profile (we follow Conesa et al., 2009, in this).[12]

[12]We include a full description of the model and its calibration in Appendix C.

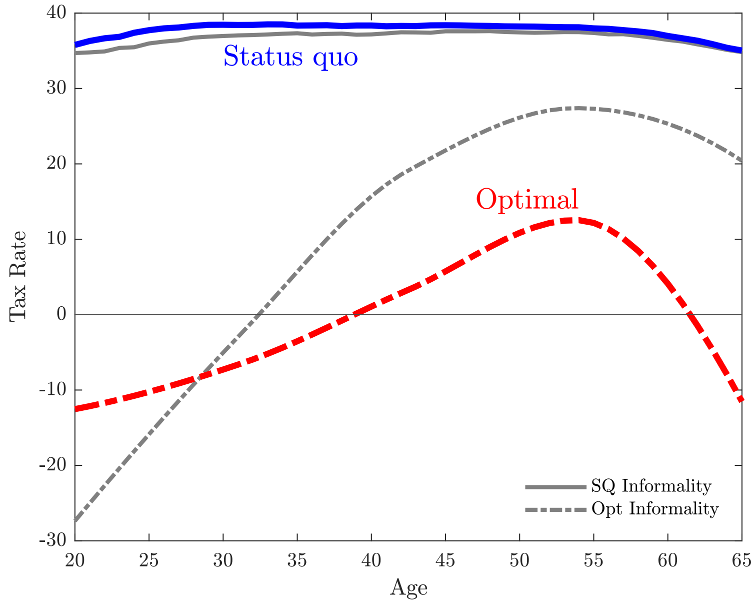


Figure 9: **Optimal Age-dependent Tax Schedule.** Comparison of optimal tax reform under economy with occupation transitions and full employment economy. All tax rates are presented in percentages. Status quo corresponds to the sum of the cross-section average by age of effective marginal tax rates of labor income, the social security contribution and the flat tax rate on capital income. In the optimal, all these taxes are replaced by a single age-dependent flat tax scheme.

In Figure 9 we compare the optimal age-dependent tax scheme to the status quo in the full-employment economy, and also present their counterparts in the economy with occupational risk. Again, the optimal reform includes moving from a pay-as-you-go social security to a system of private accounts without minimum savings requirement, and a full reform to the PIT tax scheme where progressive labor income and flat capital income tax schedules are replaced by a unique age-dependent tax schedule.

When full-employment is assumed, the optimal age-dependent tax scheme implies a deeper tax cut compared to the economy with informality: While the average contribution rate in the economy with informality is about 10.7%, in the full-employment economy positive tax rates are completely compensated with negative taxes (subsidies) both at the beginning and at the end of the life cycle. In fact, all the tax revenue that is collected with the age-dependent tax schedule is used entirely to finance the system of transfers implied by the optimal reform. The mechanism that explain this result is simple: In the absence of occupational risk, the planner devotes the age-dependent tax schedule to safeguard agents against idiosyncratic productivity risk (negative tax rates for young workers) and longevity risk (negative tax rates for workers that are about to retire).

In the aggregate, however, the optimal reform have very similar implications in the two economies, except for some important differences (we present aggregate variations in Table 10). We first describe the similarities. The welfare gain triggered by the reform in the economy with full employment is 7.2% measured in compensated equivalent variation (compared to 6.1% in the economy with occupational risk). Due to the nature of the

Variable	% Variation	
	Informality	Full Employment
CEV	$+6.1$	$+7.2$
Consumption	$+6.7$	$+9.0$
Working age	$+11.6$	$+14.4$
Retired	-13.5	-12.9
Hours Worked	-4.9	-4.0
Labor Supply	-3.6	-2.7
Capital	$+38.2$	$+47.2$
Output	$+13.0$	$+16.4$
Consumption Tax	$+1.5$	$+90.4$
Gini Coefficients		
Wealth	-8.1	-4.3
Consumption	-1.8	$+5.7$
Disposable Income	$+4.5$	$+5.8$

Table 10: **Variations Triggered by Application of Optimal Social Security Contributions Scheme.** Compensated Equivalent Variation (*CEV*) and Aggregate Variations are measured in percentages and correspond to comparisons between steady states. The Gini coefficient variation of disposable income considers working-age population labor income after taxes and transfers. Informality refers to the economies where workers face stochastic occupation transitions.

reform, moreover, the decomposition of this welfare variation is very similar to the results in the economy with occupational risk, in the sense that all the welfare gain is explained by an increase in the aggregate consumption.

In this regard, the variations in consumption in the full-employment economy are a bit stronger compared to the economy with occupational risk, but qualitatively the results are the same: The increase in consumption is concentrated in working-age population, and retirees are severely affect by a significant contraction in consumption. From here, the changes in hours worked, labor supply, capital and output, all move in the same directions, with higher changes in the economy with full-employment in the cases of capital and output.

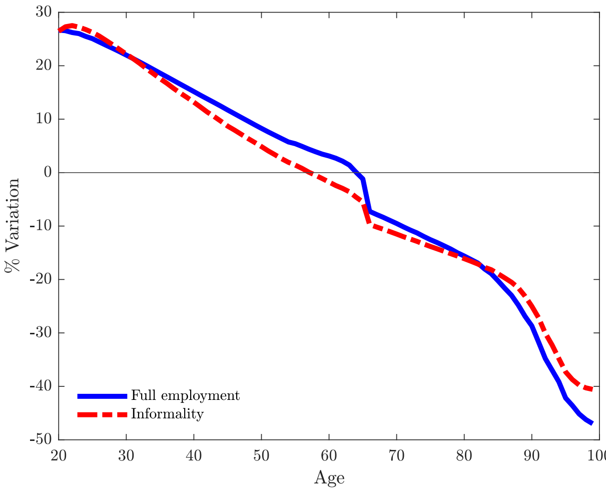
The differences, however, arise when we look at how government expenditure is financed, and the overall effect of the reform over inequality. Since all the tax revenue collected through labor income taxation is used to finance age-dependent transfers, the government needs to finance all its expenditure with consumption taxes (in the absence, of course, of the possibility to issue debt). Because of this, the consumption tax rate almost doubles under the optimal reform in the full-employment economy. Recall, in this case, that with occupation risk, the consumption tax rate only increased in 1.5% compared to the status quo.

Regarding inequality, in the presence of occupational risk the optimal reform is very effective in overcoming wealth and consumption inequality, mainly because the optimal reform triggers an income effect that affects all the working-age households, thus allowing them to increase consumption regardless of their labor status. On the contrary, when full employment is assumed, even though the reform reduces wealth inequality (although only half of the reduction in the economy with occupational risk), it increases consumption and disposable income inequality. The reason for the increase in consumption inequality is that, when full employment is assumed, the Gini coefficient reflects the augmented gap between the consumption of working-age population and retirees. Regarding the inequality of disposable income, the reasons are the same as in the economy with occupational risk: It is explained by the removal of the progressive labor income tax scheme.

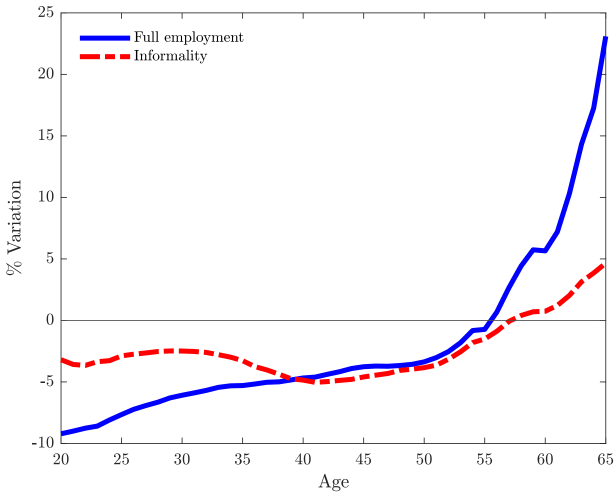
To have a better understanding of why we have these differences in inequality variations under the two alternative settings, in Figure 10 we show the differences in age profiles between pre- and post-reform settings under the full-employment assumption (blue lines) and the economy with occupational risk (red, dotted lines) for average consumption and assets (Panels 10a and 10c), and for hours worked in the formal sector (Panel 10b).

The variations in consumption are very similar in the two economies, and this makes sense because of the consumption smoothing mechanism implied by households' optimization over the life-cycle. The main differences arise when we look at how hours worked are distributed over the life cycle, and how assets are accumulated. Regarding the former, in the economy with occupational risk formal hours worked are more inelastic compared to the economy with full employment. The reason is that when the worker enjoys a formality spell, she has all the incentives to take advantage of the higher levels of productivity implied by formality, since this will allow her to improve her self-insurance against future negative shocks. When there is full employment, on the other hand, these large shifts in future productivity are not present, so hours worked gain more elasticity, particularly towards the end of the productive years.

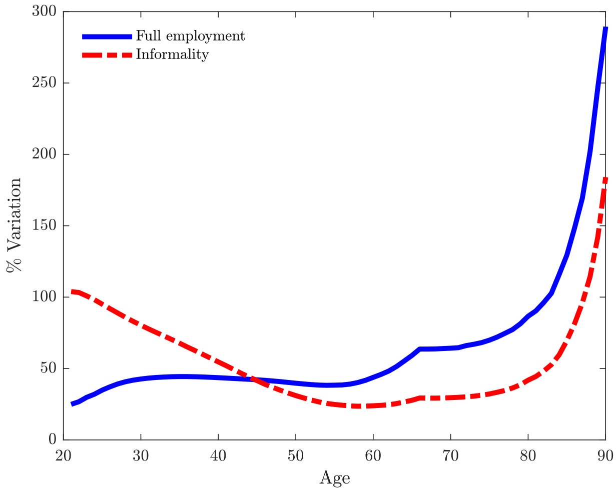
Regarding assets accumulation, in the full-employment economy the presence of negative taxes at birth allow households to accumulate assets faster. The accumulation rate, however, is much lower compared to what occurs in the economy with occupational risk. The reason for this is that in the latter, negative transfers do no last as much, and the risk of severe drops in future productivity is higher. During the last productive years, moreover, workers face positive tax rates in the economy with occupational risk, while in the full-employment economy workers enjoy negative transfers just before retirement, thus allowing them to significantly increase the assets stock to insure themselves against longevity risk. This occurs at the beginning of the life cycle when occupational risk is present.



(a) Consumption



(b) Hours Worked



(c) Savings

Figure 10: **Percent Variations between Pre- and Post-reform Steady States for Economies with Informality and Full Employment.** For consumption and assets we compare cross-section age-specific averages for the entire population. For hours worked we compare the variations only for formal workers.

8. Conclusion

In this paper we showed that a major reform to social security and the personal income tax scheme its not only optimal, but in many ways, desirable, specially in the context of informal labor markets. For this, we used an overlapping generations economy where households faced uninsured idiosyncratic risk and partially insured occupational risk, age- and sector-specific productivity and, stochastic retirement and life spans.

The optimal reform consists on moving from a pay-as-you-go social security scheme to a system of private accounts, and replacing progressive labor income taxation and capital income taxation with a single, age-dependent labor income tax schedule. We find that the optimal tax reform triggers a 6.1% welfare gain in the long run, but at the cost of a 12.8% cut in the short run, which is mainly explained by the social security reform. In the long run, moreover, the reform produces a reduction in both wealth and consumption inequality, but an increase in the dispersion of disposable income due to the elimination of progressive income taxation. We also showed that the fall in inequality is stronger in the presence of informality compared to a setting were we assume full employment.

The main shortcoming in our approach is that we do not ask the planner to find the optimal policy throughout the entire transition path (remember that the optimization we propose occurs in the steady state implied by the policy). We do, however, make a major effort to disentangle the sources of welfare losses in the short run, and also the groups of the population that are more negatively affected by the reform. The purpose of this exercise is that, knowing about the potential losses and recognizing the most affected groups, the reform could be accompanied by temporary compensation policies. We could also improve the results regarding inequality by allowing age-dependent progressivity, as proposed in Heathcote et al. (2020). All these additional features, however, are marginal with respect to the main contribution that we offer in the paper.

The paper does, nonetheless, leave the door open for interesting venues for future research. One of the main question that remains unanswered by our effort is how the optimal tax reform would change if we also consider the extensive margins involved in labor supply endogenous decisions. Here, we model informality and unemployment as exclusion mechanisms, in the sense that these are shocks that cannot be controlled by the household. I there is some degree of endogeneity in labor supply extensive margins, how would different formalization policies (such as further enforcement of social security contributions) affect the design of social security and tax reforms? How can a complete reform, as the one we proposed here, would affect the extensive margin in the households' decision of becoming informal? How would this affect unemployment?

References

Aiyagari, S. R. (1995): "Optimal Capital Income Taxation With Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of political Economy*, 103, 1158–1175.

Akerlof, G. A. (1978): "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *The American Economic Review*, 68, 8–19.

Alloush, M., C. Chartouni, R. Gatti, and J. Silva (2013): "Informality and exclusion: evidence from matched employer-employee data for Lebanon and Syria," *IZA Journal of Labor Policy*, 2, 18.

Altamirano-Montoya, A., S. Berstein, M. Bosch, M. García-Huitrón, and M. Oliveri (2018): *Presente y futuro de las pensiones en América Latina y el Caribe*, Inter American Development Bank.

Alvarez-Parra, F. and M. Toledo (2016): "On the sources of productivity losses in an economy with informality," .

Amaral, P. S. and E. Quintin (2006): "A Competitive Model of the Informal Sector," *Journal of Monetary Economics*, 53, 1541–1553.

Antón, A. (2014): "The Effect of Payroll Taxes on Employment and Wages Under High Labor Informality," *IZA Journal of Labor & Development*, 3, 20.

Audet, C. and J. E. Dennis Jr (2002): "Analysis of Generalized Pattern Searches," *SIAM Journal on Optimization*, 13, 889–903.

Barillas, F. and J. Fernández-Villaverde (2007): "A Generalization of the Endogenous Grid Method," *Journal of Economic Dynamics and Control*, 31, 2698–2712.

Bell, F. and M. Michael (2002): "Life Tables for The United States Social Security Area 1900–2100," Tech. Rep. Actuarial Study 120, Social Security Administration.

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

——— (2015): "The Labor Market Effects of Introducing Unemployment Benefits in an Economy with High Informality," *European Economic Review*, 75, 1–17.

Bosch, M. and W. Maloney (2007): *Comparative analysis of labor market dynamics using Markov processes: An application to informality*, The World Bank.

Canelas, C. (2018): "Informality and poverty in Ecuador," *Small Business Economics*, 1–19.

CARROLL, C. D. (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics letters*, 91, 312–320.

CONESA, J. C. AND C. GARRIGA (2000): "Reforma del Sistema de Seguridad Social y Adquisición de Formación," *Investigaciones Económicas*, 24, 271–295.

CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): "Taxing Capital? Not a Bad Idea After All!" *American Economic Review*, 99, 25–48.

CONESA, J. C. AND D. KRUEGER (2006): "On the Optimal Progressivity of the Income Tax Code," *Journal of Monetary Economics*, 53, 1425–1450.

DE PAULA, A. AND J. A. SCHEINKMAN (2010): "Value-Added Taxes, Chain Effects, and Informality," *American Economic Journal: Macroeconomics*, 2, 195–221.

DOLIGALSKI, P. AND L. E. ROJAS (2020): "Optimal Redistribution with a Shadow Economy," Tech. rep., European University Institute.

EPANECHNIKOV, V. A. (1969): "Non-parametric Estimation of a Multivariate Probability Density," *Theory of Probability & Its Applications*, 14, 153–158.

EROSA, A. AND M. GERVAIS (2002): "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, 105, 338–369.

EROSA, A. AND T. KORESHKOVA (2007): "Progressive Taxation in a Dynastic Model of Human Capital," *Journal of Monetary Economics*, 54, 667–685.

FARHI, E. AND I. WERNING (2013): "Insurance and Taxation Over the Life Cycle," *Review of Economic Studies*, 80, 596–635.

FERNÁNDEZ, C. AND L. VILLAR (2017): "The Impact of Lowering the Payroll Tax on Informality in Colombia," *Economia*, 18, 125.

DATASET, IRS (2020): "Registry of Business Groups and Big Corporations (Catastros)," Data retrieved from *Servicio de Rentas Internas*, Catastros, https://www.sri.gob.ec/web/guest/catastros.

DATASET, SUPERCÍAS (2020): "Balance Sheets for Registered Firms in Ecuador, 2006-2019," Data retrieved from *Superintendencia de Compañías, Valores y Seguros*, Balance Sheets by ISIC code, https://appscvsmovil.supercias.gob.ec/portalInformacion/sector_societario.zul.

FORTIN, B., N. MARCEAU, AND L. SAVARD (1997): "Taxation, Wage Controls and the Informal Sector," *Journal of Public Economics*, 66, 293–312.

FRITSCH, F. N. AND R. E. CARLSON (1980): "Monotone Piecewise Cubic Interpolation," *SIAM Journal on Numerical Analysis*, 17, 238–246.

Gachet, I., D. Grijalva, P. Ponce, and C. Uribe-Teran (2018): "Ecuador's Income Tax Dataset (2001-2017): Methodology and Income Distribution Analysis," USFQ Working Paper.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003): "Optimal Indirect and Capital Taxation," *The Review of Economic Studies*, 70, 569–587.

Granda, C. and F. Hamann (2015): "Informality, Saving and Wealth Inequality in Colombia," *IDB Working Paper Series*.

Grijalva, D. F., I. Gachet, P. Lucio-Paredes, and C. Uribe-Teran (2018): "The Political Economy of Trade Policy in Ecuador: Dollarization, Oil, Per - sonalism and Ideas," *Inter-American Development Bank, Washington, DC, Unpublished*.

Gruber, J. (1997): "The Incidence of Payroll Taxation: Evidence from Chile," *Journal of Labor Economics*, 15, S72–S101.

Haanwinckel, D. and R. Soares (2020): "Workforce Composition, Productivity, and Labor Regulations in a Compensating Differentials Theory of Informality," *University of Chicago, Becker Friedman Institute for Economics Working Paper*.

Hansen, G. D. (1993): "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked," *Journal of Applied Econometrics*, 8, 71–80.

Heathcote, J., K. Storesletten, and G. L. Violante (2020): "Optimal progressivity with age-dependent taxation," *Journal of Public Economics*, 104074.

Heer, B. and A. Maussner (2009): *Dynamic general equilibrium modeling: computational methods and applications*, Springer Science & Business Media.

Hintermaier, T. and W. Koeniger (2010): "The Method of Endogenous Gridpoints with Occasionally Binding Constraints among Endogenous Variables," *Journal of Economic Dynamics and Control*, 34, 2074–2088.

Karabarbounis, M. (2016): "A Road Map for Efficiently Taxing Heterogeneous Agents," *American Economic Journal: Macroeconomics*, 8, 182–214.

Kindermann, F. and D. Krueger (2014): "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," Tech. rep., National Bureau of Economic Research.

Krueger, D. and A. Ludwig (2013): "Optimal Progressive Labor Income Taxation and Education Subsidies When Education Decisions and Intergenerational Transfers are Endogenous," *American Economic Review*, 103, 496–501.

——— (2016a): “On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrium,” *Journal of Monetary Economics*, 77, 72–98.

——— (2016b): “On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium,” *Journal of Monetary Economics*, 77, 72–98.

Krueger, D., K. Mitman, and F. Perri (2015): “Macroeconomics and heterogeneity, including inequality,” *Handbook of Macroeconomics*.

Kucera, D. and L. Roncolato (2008): “Informal employment: Two contested policy issues,” *International Labour Review*, 147, 321–348.

Lagarias, J. C., J. A. Reeds, M. H. Wright, and P. E. Wright (1998): “Convergence properties of the Nelder–Mead simplex method in low dimensions,” *SIAM Journal on optimization*, 9, 112–147.

Maloney, W. F. (2004): “Informality Revisited,” *World development*, 32, 1159–1178.

Margolis, D. N., L. Navarro, and D. Robalino (2014): “Unemployment Insurance, Job Search, and Informal Employment,” *Social Insurance, Informality, and Labour Markets: How to Protect Workers While Creating Good Jobs*, 112.

McGrattan, E. R. and E. C. Prescott (2017): “On Financing Retirement With an Aging Population,” *Quantitative Economics*, 8, 75–115.

McKiernan, K. (2020): “Social Security Reform in the Presence of Informality,” *Review of Economic Dynamics*.

Medvedev, D. and A. M. Oviedo (2016): “Informality and Profitability: Evidence from a New Firm Survey in Ecuador,” *The Journal of Development Studies*, 52, 412–427.

Meghir, C., R. Narita, and J.-M. Robin (2015): “Wages and Informality in Developing Countries,” *American Economic Review*, 105, 1509–46.

Mirrlees, J. A. (1971): “An Exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies*, 38, 175–208.

Mondragon-Velez, C., X. Pena, D. Wills, and A. Kugler (2010): “Labor Market Rigidities and Informality in Colombia [with Comment],” *EconomÃa*, 11, 65–101.

Morales, L. F. and C. Medina (2017): “Assessing the Effect of Payroll Taxes on Formal Employment: The Case of the 2012 Tax Reform in Colombia,” *Economía*, 18, 75–124.

Ndiaye, A. (2017): “Flexible Retirement and Optimal Taxation,” *Mimeo, Northwestern University*.

Perry, G. E., O. Arias, P. Fajnzylber, W. F. Maloney, A. Mason, and J. Saavedra-Chanduvi (2007): *Informality: Exit and exclusion*, The World Bank.

Tauchen, G. and R. Hussey (1991): "Quadrature-based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica: Journal of the Econometric Society*, 371–396.

Ulyssea, G. (2010): "Regulation of Entry, Labor Market Institutions and the Informal Sector," *Journal of Development Economics*, 91, 87–99.

——— (2018): "Firms, Informality, and Development: Theory and Evidence from Brazil," *American Economic Review*, 108, 2015–47.

Uribe-Teran, C. (2020): "Endogenous Tax Compliance, the Elasticity of Taxable Income and Optimal Labour Income Taxation," Tech. rep., Universidad San Francisco de Quito.

Weinzierl, M. (2011): "The Surprising Power of Age-Dependent Taxes," *The Review of Economic Studies*, 78, 1490–1518.

Williams, C. C. and Y. Youssef (2015): "Theorising Entrepreneurship in the Informal Sector in Urban Brazil: A Product of Exit or Exclusion?" *The Journal of Entrepreneurship*, 24, 148–168.

Zenou, Y. (2008): "Job Search and Mobility in Developing Countries. Theory and Policy Implications," *Journal of development economics*, 86, 336–355.

A. More on the Estimation of Age-specific Productivity

[TO BE COMPLETED]

B. Production Function Estimation

The firm in our baseline model hires formal and informal labor, and capital. We assume that it uses a Cobb-Douglas production function that exhibits constant returns to scale on capital and labor. Total labor, moreover, is obtained with a CES aggregator over formal and informal inputs. In this section of the appendix we present the empirical strategy, data description and main results related to the estimation of the parameters of the production function.

B.1. Empirical Strategy

The production function we use in the baseline model can be written as

$$Y_t = AK_t^{\alpha}\left[\alpha_f L_f^{\gamma} + (1-\alpha_f)L_i^{\gamma}\right]^{\frac{1-\alpha}{\gamma}}, \tag{32}$$

where, as usual, α is the share of national income paid to capital and $1-\alpha$ is the labor share. For the calibration of our occupational risk model we need to pin down the value of α, since α_f is determined with labor surveys and γ is set by assumption.

As it will become clear in the following section, in our data we can only observe the number of workers that are hired in formal firms under formal contracts. We assume, thus, that each firm i at time t uses a Cobb-Douglas technology to produce final consumption goods. Formally,

$$Y_{it} = AK_{it}^{\alpha}L_{it}^{1-\alpha}\exp\{\beta'\mathbf{X}_i + \delta'\mathbf{dt}_t + \eta'\mathbf{di}_i + \mu_{it}\},$$

where $\mathbf{X}_i$ is a matrix of controls at the firm level, $\mathbf{dt}_t$ are time fixed effects and $\mathbf{di}_i$ are firm-level fixed effects. Taking logs on both sides of the previous equation we obtain

$$\log Y_{it} = \beta_0 + \alpha_0 \log K_{it} + \alpha_1 \log L_{it} + \beta'\mathbf{X}_i + \delta'\mathbf{dt}_t + \eta'\mathbf{di}_i + \mu_{it}, \tag{33}$$

where $\beta_0 = A$, $\alpha_0 = \alpha$ and $\alpha_1 = 1-\alpha$. Estimation of (33) by OLS or panel fixed effects might be biased because of endogeneity. Instead, we opt for the two-step Arellano-Bond estimator to allow the model for a partial adjustment mechanism that might reflect frictions proper of emerging economies. Once this variable is accounted for, the strict exogeneity assumption in the fixed effects estimation is violated and coefficients are biased. The two-step Arellano-Bond GMM procedure corrects for this bias using higher order lags of the dependent variable as instruments.

B.2. Data

The data that we use to estimate (33) its a combination of administrative data that is publicly available in Ecuador. First, we use the balance sheet information that all legally registered firms in Ecuador are required to report to the Firms' Superintendency (*Superintendencia de Compañías, Valores y Seguros, Supercías by the spanish acronym*, which is the same form (Form 101) that firms file as part of their Income Tax Return each year (Dataset, Supercías, 2020). Moreover, we gain access to data on workers per firm using the dataset by Grijalva et al. (2018).

From this dataset we have enough accounting information to compute value added, the capital stock and the total wage bill at the firm level. Moreover, we know the industry at which each company belongs to (ISIC 6 digits code) and information on the geographical location of the firms' headquarters. The dataset published that is publicly available in the Supercías web page also contains the firms' fiscal ID, which allows us to match this data with information that is also made public by Ecuador's Internal Revenue Service (IRS).

In this regard, Ecuador's IRS carries a detailed registry of two type of firms that are relevant from a tax administration perspective. First, the IRS has done an extraordinary job in identifying firms that work within Business Groups (or *Economic Groups* in Ecuador's tax administration jergon). A Business Group is a set of local firms and persons that are related through the composition of the shareholders of each firm and, as a group, are owners (or shareholders) of firms offshore (Dataset, IRS, 2020).

In the same vein, the IRS also identifies big corporations (*Grandes Contribuyentes*) which are firms that, due to the volume of their operations, are subject to additional specific legislation and are required to file additional returns compared to regular firms. Again, the IRS publishes the registry of big corporations, so the data is publicly available (Dataset, IRS, 2020). Same as the Business Groups registry, the Big Corporations registry contains the firms' fiscal ID, which allows us to match the information with the Supercías data.

In Table 11 we present descriptive statistics regarding the main variables used for the estimation classified according to the industry to which each firm belongs to. We only consider goods in our estimations, so the industries that are included in our sample are Agriculture, Manufacture, Construction and Wholesale & Retail. We present pooled statistics for years 2015, 2016 and 2017: Averages in main row and standard deviations in parenthesis. Value added and capital are measured in thousands of 2007 US Dollars, workers are presented in number of persons, and the proportions of firms that belong to Business Groups or that are classified as Big Corporations are in percentages.

Most of the firms in our sample concentrated in the Wholesale & Retail industry, although Manufacture is the biggest in terms of sales and capital stock. Agriculture, on the other hand, concentrates more workers (on average) and a higher proportion of firms that

Variable	Agriculture	Manufacture	Construction	Wholesale & Retail
Value Added	1,383.8	2,200.1	521.8	879.7
	(9,186.8)	(10,222.3)	(3,521.1)	(5,643.8)
Capital	2,603.1	4,485.3	1,454.7	1,952.8
	(15,552.2)	(18,909.1)	(12,278.4)	(14,491.5)
Workers	75.3	63.0	21.4	28.1
	(330.6)	(184.7)	(127.9)	(175.8)
Business Group	11.0	8.1	2.3	4.1
	(31.3)	(27.2)	(15.1)	(19.9)
Big Corporation	28.9	24.3	29.7	22.8
	(45.3)	(42.9)	(45.7)	(41.9)
Observations	6,419	9,803	5,421	29,151

Table 11: **Descriptive Statistics.** All values are means within each industry. Standard deviation in parenthesis. Value added and capital are measured in thousands of 2007 US dollars, workers are presented in number of persons. Economic group and big corporation are percentages. We present pooled statistics for years 2015, 2016 and 2017.

are connected to Business Groups. Agriculture and Construction exhibit higher proportions of firms that are classified as Big Corporations.

B.3. Results

Before going into the discussion of the main results, it is worth explaining why we consider that tax-driven firm classifications as controls are important for the production function estimation. The reason for this is that we are working with administrative data that firms file for tax purposes. If we do not control for these firms' characteristics, we would be omitting important features of the firms that have a definite impact on its behavior as taxpayer, thus affecting the values they feed to their tax returns.

With this in mind, in Table 12 we present the estimation results of four different specifications: OLS, Fixed Effects, One-step Arellano-Bond and Two-step Arellano-Bond, in that order. In the OLS model we also include industry and province fixed effects. In all cases we include time fixed effects.

Our results show that the share of income paid to capital in Ecuador is significantly higher to the typical calibration for the US. In our preferred specification we see that, on average, 43.3% of firms' value added is used to pay for capital services, vs. 36% that is usually estimated for the US. This is why we set $\alpha = 0.433$ in the model we use for Ecuador.

	(1) OLS	(2) Fixed Effects	(3) One-step Arellano-Bond	(4) Two-step Arellano-Bond
Capital (log)	0.408***	0.443***	0.425***	0.433***
	(0.013)	(0.033)	(0.034)	(0.033)
Workers (log)	0.562***	0.510***	0.509***	0.504***
	(0.017)	(0.024)	(0.021)	(0.021)
Business Group	–0.208***	0.071	0.063	0.064
	(0.055)	(0.045)	(0.047)	(0.047)
Big Coorporation	0.508***	0.488***	0.496***	0.496***
	(0.063)	(0.090)	(0.026)	(0.026)
Interaction	0.176***	–0.065	–0.069	–0.079
	(0.063)	(0.055)	(0.051)	(0.051)
Time FE	YES	YES	YES	YES
Industry FE	YES	NO	NO	NO
Province FE	YES	NO	NO	NO
Constant	YES	YES	YES	YES
Observations	50,794	50,794	29,406	29,406
Firms	21,359	21,359	16,766	16,766
Wald test	2.00	2.16	4.30	4.02

Table 12: **Estimation of Cobb-Douglas Parameters.** Columns (1) and (2) includes clustered standard errors at the industry level (ISIC 3-digits) in parenthesis, columns (3) and (4) includes WC-robust standard errors in parenthesis. Controls include dummy variables that identify firms that belong to economic groups and that are classified as big corporations by Ecuador's IRS (both time-varying), and its interaction. For the Wald test we use the null $H_0 : \alpha_0 + \alpha_1 = 1$ and report the F-value for models (1) and (2), and the χ^2 for models (3) and (4). For models (1) and (2) we cannot reject the null hipothesis, while for models (3) and (4) we reject the null at the 5% significance level.

Something interesting that we find in this estimation is that the labor share falls as the especification becomes more robust: It goes from 56.2% in the OLS specification to 50.4% in the Two-step Arellano-Bond estimator. Because of this, although we cannot reject the null hypothesis of constant returns to scale in models (1) and (2), we do reject the null in the Arellano-Bond estimations. We still assume constant returns in the model, however, mainly because rejection only occurs at the 5% significance level. Finally, we find that the big corporation classification is the only tax-related classification that matters showing that, on average, this group of firms produce about 49.6% more value added than regular firms.

C. The Economy with Full Employment

The main difference in the economy with full employment is that, instead of looking at the whole structure, we only concentrate on the formal labor market. This assumption has consequences for most of the model that we presented in the paper, and is much closer to the typical economy in Conesa et al. (2009) or Krueger and Ludwig (2016a). The main difference with respect to this literature is that we still assume stochastic retirement.

C.1. Households

The economy is populated by bachelor households that enter the labor market at model age 1, and are randomly distributed among possible idiosyncratic productivity shocks. Differently to the benchmark economy, and in the absence of occupation-specific productivity, we assume that households can be born into one of two possible ability shocks. All retirees have access to the pensions fund, and each period workers are endowed with one unit of time that has to be distributed between leisure and work. Except for occupation transitions, demographics are the same as the benchmark economy. We use the full-employment economy only to compare long-term equilibrium, so we only describe the stationary economy.

C.1.1. Income Process

Households' income comes from wages determined in equilibrium w, an idiosyncratic productivity shock η, age-specific productivity ε_j for all $j \in \{1, \ldots, J\}$, fixed ability shock determined at birth $\mu \in \{-\sigma^{\mu}, \sigma^{\mu}\}$ and the amount of hours supplied by workers ℓ. The log of labor supply can thus be written as

$$\log y_j = \log \varepsilon_j + \log \eta + \log \mu + \log w + \log \ell.$$

As in the benchmark economy, we assume that the idiosyncratic productivity shock follow an $AR(1)$ process in logs, so

$$\log \eta' = \rho_{\eta} \log \eta + \epsilon',$$

where we use prime notation to denote period $t+1$ in steady state. As before, ρ_{η} is the coefficient of autocorrelation and ϵ is normally distributed with zero mean and variance σ_{η}^2. In this case, $F\eta(\eta'|\eta)$ is the stationary cumulative distribution function.

C.1.2. Recursive Formulation

The state space in this economy is much simpler compared to the benchmark. For notation tractability, however, we present again the state space as

$$x = \{a, \eta, \mu, s, j\},$$

where $s \in f, r$ implying that the household is either hired in the formal sector or retired. With this definition, the problem solved by working-age households in the stationary equilibrium can be written as,

$$\begin{aligned} & V(x) = \max\{U(c,\ell) + \beta\psi \,\mathrm{E}\,\tilde{V}^f(x')\}, \\ \text{s. t.} \quad & a' = (1+R)a + (1-\tau^{ss})y_j - T((1-\tau^{ss})y_j) + Tr + G - (1+\tau^c)c, \\ & a' \geq 0, \quad c > 0, \quad 0 \leq \ell \leq 1, \end{aligned} \tag{34}$$

where

$$\mathrm{E}\,\tilde{V}^f(x') = (1-p_j^r)\int V(x')F\eta(d\eta^p r|\eta) + p_j^r V((x^r)'),$$

is the expectation taken over the possibility of early retirement and $x^r = \{a, j\}$ is the state space for retirees. During retirement, thus, the households' problem can be written as

$$\begin{aligned} & V(x^r) = \max\{u(c) + \beta\psi_j V((x^r)')\}, \\ \text{s. t.} \quad & a' = (1+R)a + b^r + Tr + G - (1+\tau^c)c, \\ & a' \geq 0, \quad c > 0. \end{aligned} \tag{35}$$

C.1.3. Firm

The firm is much simpler in this economy. Since there is no diversity in labor supply, the production function takes a Cobb-Douglas production function, so

$$Q(K, L) = AK^\alpha L^{1-\alpha},$$

where α is the share of output paid to capital. Profit maximization implies the following expressions for the wage and the interest rate:

$$r = \alpha\frac{q}{k} - \delta, \tag{36}$$

$$w = (1-\alpha)\frac{q}{k}, \tag{37}$$

where $q = Q/L$ is output per effective labor and $k = K/L$ is capital per effective labor.

C.1.4. Social Security

In the status quo, the social security system runs a pay-as-you-go contributory scheme. Workers contribute by means of a flat payroll tax and we assume that the social security fund is always balanced. The transfer for retirees b^r is thus determined in equilibrium.

C.2. Government

The government is identical to the benchmark economy in the status quo. It uses tax revenue that comes from labor income, capital income and consumption taxation to finance an exogenous stream of government consumption $G = gQ$, where $0 \leq g \leq 1$.

C.3. Equilibrium

We now turn to the formal definition of the competitive equilibrium in the full-employment economy. We build upon the definitions stated in Section 3.5 keeping in mind that we only focus on the stationary equilibrium. We do so in Definition C.1.

Definition C.1 (Stationary Competitive Equilibrium). *Given capital income and consumption tax rates τ^c, τ^k, the progressive labor income tax scheme $T(\underline{\tau}, \overline{\tau}, \underline{y}, \overline{y})$ and the payroll tax rate τ^{ss}, an stationary competitive equilibrium in the full-employment economy are functions for households $\{V, c, a, \ell\}$, a production plan for the firm $\{K, L\}$, a pension transfer b^r, prices $\{r, w\}$, the transfer derived from accidental bequest Tr and the invariant distribution $\Phi \in \mathcal{M}$ such that:*

(i) Households maximize their life-time expected utility, so $\{V, c, a, \ell\}$ solve problems (34) and (35),

(ii) The production plan $\{K, L\}$ maximizes profits of the firm, so prices $\{r, w\}$ satisfy equations (36) and (37),

(iii) The social security system's budget constraint is satisfied, so

$$b^r \int \Phi(dx^r) = \int \tau^{ss} \varepsilon_j \eta \mu w \ell(x^f) \Phi(dx^f), \tag{38}$$

where $x^f = \{a, \eta, \mu, j\}$,

(iv) The accidental bequest lump-sum transfer is given by,

$$Tr' = \int (1 - \psi_j) a(x) \Phi(dx), \tag{39}$$

(v) Markets clear:

$$C = \int c(x)\Phi(dx), \tag{40}$$

$$K' = \int a'(x)\Phi(dx), \tag{41}$$

$$L = \int \varepsilon_j \eta \mu \ell(x^f)\Phi(dx^f), \tag{42}$$

$$K' + C + G = Q(K, L) + (1-\delta)K, \tag{43}$$

(vi) The government runs a balanced budget,

$$G = \tau^c C + \tau^k K + \int T((1-\tau^{ss})\varepsilon_j \eta \mu w \ell(x^f)). \tag{44}$$

This definition of the stationary competitive equilibrium can be adjusted, following the same logic as we did in the paper, to incorporate the structural changes embedded in the optimal tax reform.

C.4. Estimation and Calibration

We take advantage of the calibration and estimation procedure that we already completed for the economy depicted in the paper, and adjust some parameters to reflect the fact that, in the full-employment economy, we only look at the formal sector. To avoid repetition, we only present here the parts in which the calibration of the full-employment economy diverts from what we already did in the paper.

We recalibrate the model to match similar targets and in the paper. Now, however, we use the standard deviation of the fixed ability shock σ_μ to match the variance of log labor income at age 20. The result of the calibration procedure are presented in Table 13.

Parameter	Target	Value
σ_1	Literature	2.0000
σ_2	Literature	3.0000
χ	Average hours worked	1.0778
β	Steady-state capital-output ratio	0.9700
σ_μ	Variance of log income at age 20	0.2048
σ_η	Variance of log income at age 50	0.2074
ρ_η	Linear increase in variance of log income	0.9900

Table 13: **Calibrated Parameters for Households in the Full-employment Economy.** Parameters are computed minimizing a weighted quadratic loss function except for σ_1, σ_2 and ρ_η.

Regarding the calibration of the representative firm, the only difference is that there is no informal labor input, so we do not need to calibrate the elasticity of substitution. The rest of parameters remain unchanged. Something similar occurs with the parameters for the social security system and the government.

Another important difference occurs in how we estimate age-specific productivity. In this case, we only concentrate on formal workers to perform this estimation, and we set $_1 = 1$, so the rest of the productivity profile represents average productivity growth along the life cycle. The final output is a re-centered version of the age-specific productivity profile depicted in Figure 2.

D. Algorithm for the Numerical Solution

We solve two versions of the model: The stationary general equilibrium and the transitional dynamics. The solution of the stationary equilibrium is simple and uses initial guesses for the main aggregates involved in the model to apply a bisection method. In this section of the appendix, we concentrate on describing the algorithms involved in the solution of the households' problem, and then the algorithm used for the simulation of the transitional dynamics.

D.1. The Households' Problem

At any point in time, households face idiosyncratic productivity shocks and stochastic occupation transitions while facing a non-linear tax scheme. For this reason, we use a combination of the Endogenous Grid Method (Carroll, 2006; Barillas and Fernández-Villaverde, 2007; Hintermaier and Koeniger, 2010), and Value Function Iteration to solve the model. The complication arises in that the non-linearity of the tax scheme can potentially produce local maxima. To prevent the algorithm from getting stuck in local optima, we use the results of grid-search optimization as initial guesses for continuous optimization routines. For the latter, we use the Nelder-Mead simplex method as described in Lagarias et al. (1998). In what follows we explain in detail all the algorithms involved in the solution of the households' problem.

Algorithm 1: Set model's parameters and grids

Result: Parametrized model and grids
setup parameters;
exponential grid for assets $\mathbf{a}_g$;
grid for idiosyncratic productivity shock $\mathbf{eta}_g$ (using Tauchen and Hussey, 1991);
transition matrix for idiosyncratic shock $\mathbf{P}_\eta$;
matrix of occupation transitions $\mathbf{P}_s$;

Algorithm 2: Solve problem for retirees

Result: Policy functions for retirees: $V(a,j), a'(a,j), c(a,j)$
$\to$ Start EGM;
while $a_i \in \mathbf{a}_g$ **do**
  use finite economy to compute consumption during last period alive;

$$\begin{aligned} c(a_i, J) &= \left[(1+(1-\tau^k)r)a_i + b^r + G + Tr\right]/(1+\tau^c); \\ V(a_i, J) &= U(c(a_i, J)); \end{aligned}$$

end
while $1 \leq i_j < J$ **do**
  define current age going backwards $j = J - i_j$;
  while $a_i \in \mathbf{a}_g$ **do**
    compute consumption in endogenous grid using Euler equation;

$$c_i^e(a_i^e(a_i,j),j) = U^{-1}(c(a_i,j+1));$$

    solve for endogenous grid using households' budget constraint;

$$a_i^e(a_i,j) = (a_i + (1+\tau^c)c^e(a_i^e(a_i,j),j) - b^r - Tr - G)/(1+(1-\tau^k)r);$$

    use PCHIP interpolation to find $a_i'(a_i,j)$;
    recover consumption on assets grid using budget constraint $c_i(a_i,j)$;
    check that borrowing constraint is satisfied;
    check that consumption non-negativity constranint is satisfied;
    update value function $V(a_i,j) = U(c(a_i,j)) + \beta\psi_j V(a'(a_i,j+1),j+1)$;
  end
end

The problem for working age population is slightly more difficult, since we cannot solve formal, informal and unemployed workers' problems for all agents separately (the problems are interconnected because of occupation transitions in the future). The strategy we take consists on solving the three problems for a given age, and then iterate backwards all the way to age 1. The problems for formal and informal workers are solved using grid search VFI enhanced with a local minimization routine, while the problem for the unemployed is solve using the Endogenous Grid Method. We describe this process in detail in the following algorithms.

Algorithm 3: Solve problem for formal workers

Result: Policy functions for formal workers:
$V(a,\eta,f,j), a'(a,\eta,f,j), c(a,\eta,f,j), \ell(a,\eta,f,j)$

$\rightarrow$ for a given age $j\ldots$;

setup fine grids for assets $\mathbf{a}_g^f$ and hours worked $\mathbf{n}_g^f$;

while $a_i \in \mathbf{a}_g$ **do**

  while $\eta \in \mathbf{eta}_g$ **do**

    perform grid search over fine grids $\mathbf{a}_g^f$ and $\mathbf{n}_g^f$ using VFI;

    use $(a')_i^*$ and n_i^* as initial conditions for local optimizer using interpolated VFI (PCHIP);

    obtain optimal assets and hours worked ;

    compute consumption using households' budget constraint ;

    update policy functions V, a', c, ℓ ;

  end

end

Algorithm 4: Solve problem for informal workers

Result: Policy functions for informal workers:
$V(a,\eta,i,j), a'(a,\eta,i,j), c(a,\eta,i,j), \ell(a,\eta,i,j)$

$\rightarrow$ for a given age $j\ldots$;

similar to algorithm 3 but setting labor-income taxes to zero ;

Algorithm 5: Solve problem for unemployed

Result: Policy functions for unemployed: $V(a,u,\tilde{s},j), a'(a,u,\tilde{s},j), c(a,u,\tilde{s},j)$

$\rightarrow$ for a given age $j\ldots$;

if $\tilde{s} == f$ **then**

  $b^u = \lambda^u$;

else

  $b^u = 0$;

end

solve unemployed households' problem using EGM (similar to algorithm 2) ;

Algorithm 6: Solve problem for working-age population

Result: Policy functions for all workers and all ages

while $j = 1, \ldots, j^r$ **do**

  algorithm 3;

  algorithm 4;

  algorithm 5;

end

Once we obtain the policy functions, we compute the invariant distribution using Montecarlo simulations. In this simulation we assume that each cohort consists of 1,500 households at birth, so the result is a panel of households that we follow during 81 years, resulting in 76,027 observations. With this simulated dataset we compute aggregates, check converge in the bisection method where we are solving for the general equilibrium, and keep iterating until convergence is achieved.

We solve this economy using a 21-point exponential grid for assets, 7-point grid for the idiosyncratic productivity shock and 150 grid points for the fine grids for assets and working hours that we use in the grid search step for the solution of formal and informal workers problem.

We use MATLAB to find the solution, so we vectorize whenever possible. If vectorization is not an option (this occurs when we solve the formal and informal workers' problems, and when we perform the Montecarlo simulation), we use the MATLAB Coder to produce stand-alone MEX functions precompiled in C++. Each iteration of the bisection method to solve the General Equilibrium takes about 12 seconds, and solving for the equilibrium needs 25 iterations, taking about 5-6 minutes. With this, the converge difference in the interest rate is below $|1e^{-6}|$ and the excess demand is just above $|2.054e^{-5}|$ for the economy in the status quo (this numbers are very close for the alternative scenarios).

To solve for the transitional dynamics we follow Heer and Maussner (2009, pp. 406-411). In particular, we use the algorithm based on an initial guess for a finite-time path for main aggregates. We set the path for 240 periods and solve for the policy functions going backwards, keeping in mind that the last period in the simulation corresponds to the steady state under the optimal policy reform. Then, we move forward in time using Montecarlo simulations to update the distributions and compute aggregates that are compatible with the policy functions. We keep iterating until the paths for aggregates converge.