

Tkhir, Anna-Mariia

**Conference Paper**

## Education and Tax Policies in the Presence of Informality

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Tkhir, Anna-Mariia (2021) : Education and Tax Policies in the Presence of Informality, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/242394>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Education and Tax Policies in the Presence of Informality\*

Anna-Mariia Tkhir<sup>†</sup>  
Goethe University Frankfurt

February 6, 2021

**Preliminary Draft!**

Comments are welcome!

## **Abstract**

This paper analyzes the aggregate long-run effects of education and tax policies in the presence of a large informal sector. I develop an overlapping-generation life-cycle model with heterogeneous agents and incomplete markets, where agents make educational and occupational choices. Education is costly but ensures a wage premium in the future. Education costs are partially subsidized by the government. Individuals choose to operate either in the formal or in the informal sector given that the education premium is higher in the former. However, formal workers must pay progressive income taxes. I calibrate my model to Brazil and assess the effects of education subsidies on the overall economy. The increase in the education subsidy improves educational attainment, expands formalization, and increases welfare. The subsidy rate which covers 100% of education costs maximizes welfare leading to the overall welfare gain of around 94% measured in consumption equivalent variation. Due to a significant positive effect on the tax base, the education subsidy is self-financing in the long run. Tax schedule reforms have mild effects on the aggregate measures of education. However, they alter the educational composition of labor. Reducing the level of income tax or flattening the tax schedule leads to the reallocation of highly educated and productive individuals into the formal sector.

***JEL Classifications:*** I24, E26, J24.

***Keywords:*** Human Capital, Informal Economy, Education Policy, Tax Policy.

---

\*I am grateful to Alessandro Di Nola, Nicola Fuchs-Schündeln, Leo Kaas, Alexandr Ludwig, Almuth Scholl, Nawid Siassi and Oliko Vardishvili for insightful comments.

<sup>†</sup>[tkhir@wiwi.uni-frankfurt.de](mailto:tkhir@wiwi.uni-frankfurt.de). Postal address: Goethe University Frankfurt, Theodor-W.-Adorno-Platz 360629 Frankfurt, Germany.

# 1 Introduction

Developing countries feature low levels of educational attainment and high levels of informality, defined as unofficial production undeclared with a purpose of tax and social security avoidance (Schneider et al., 2010). Empirical studies show that informal workers are on average less educated and receive a lower education premium compared to the formal workers (Araujo et al., 2013; Lopez Garcia, 2015; Fairris and Jonasson, 2016; Meghir et al., 2015). Additionally, according to Filho (2012) the rise in educational attainment explains up to 60% of the informality reduction in Brazil.<sup>1</sup>

Against the backdrop that education is an important determinant of informality, it may be reasonable to encourage human capital accumulation in countries that feature large shadow economies. On the one hand, education subsidies make schooling more affordable and allow individuals to accumulate more human capital. For educated individuals, the formal sector is more attractive due to higher returns to education. On the other hand, education subsidies have to be financed with additional taxation. Raising taxes to finance the education subsidy may encourage individuals to become informal.

This paper sheds light on the interrelation between education and tax policies in the presence of an informal sector. In particular, I answer several questions: How do education subsidies affect the size of informality? What are the revenue and the welfare-maximizing levels of the subsidy? What are the consequences of the tax schedule change when allowing for educational and occupational decisions?

To answer these questions, I develop a two-sector overlapping generation model with endogenous educational and occupational choices. In the first stage of their life, individuals make their educational choice given their inherited assets and innate learning ability. The total cost of education consists of the monetary and time cost. The monetary cost is partially subsidized by the government. While making an educational choice, individuals are aware of the future education premium differing between the formal and informal sectors. Following empirical evidence, I assume that the education premium is higher in the formal sector (Arbex et al., 2010; Araujo et al., 2013). Apart from the difference in education premia, by definition sectors differ in terms of tax compliance and social security coverage. Formal employees pay progressive income taxes. Informal employees, however, do not pay income taxes but bear a fixed utility cost associated with the lack of social security. The government collects tax revenues to finance its expenditures including education subsidies, pension payouts, and government consumption.

I calibrate my model to Brazil, a country with a high level of informality and a low level of educational attainment. I use the Brazilian National Household Sampling Survey (PNAD) which contains detailed information about the individual characteristics, educational attainment, labor income, and hours of work for both workers in the formal and

---

<sup>1</sup>More empirical evidence is provided in Section 3.

informal sectors. The model matches well the characteristics of the formal and informal workers, the distribution of education within the formal and informal sectors as well as the indicators of the overall economy.

In equilibrium, formal employees have a higher average education level compared to informal employees. In line with the data, the prevailing share of formal workers has a completed high school degree. Although there is an incentive to invest a lot in education to enjoy a high education premium in the formal sector, progressive taxation prevents individuals from investing too much. Moreover, analyzing the share of informality by years of education reveals a U-shape pattern in the model and the data. This pattern indicates that the share of informality reduces with schooling at first but then rises again for individuals who have acquired tertiary education. Two key assumptions of the model, the difference in education premia and progressive taxation, suffice to explain this pattern.

In the first step, I vary the share of publicly financed education costs. The increase in the education subsidy makes education more affordable, increases the average level of education, and expands formalization. If the government carries all the monetary costs of education (subsidy = 100%) compared to the current level (subsidy = 33%), the share of highly-educated individuals increases from around 17% to over 68% and the share of informal workers reduces from 17% to 9%. It is important to note that the increase in education does not monotonically lead to formalization. When the subsidy increases from 90% to 100%, the share of highly educated rises from 56% to 68% but the share of informal workers slightly increases from 8.63 to 8.70%.

Higher human capital accumulation as well as a higher share of formal workers expand the tax base. Due to this effect, the optimal education subsidy is self-financed in the long run and does not require additional tax hikes. On the contrary, the government can reduce the income tax rates in the long run. The government budget surplus spikes for the subsidy level of 90%. The combination of higher income, formalization, and reduction in hours worked leads to a considerable welfare gain of 94% measured in consumption equivalent variation (from here onwards, *cev*) for the subsidy level of 100%.

In the second step, I analyze the effects of a change in the income tax schedule. I focus on two policies: (1) the change in the income tax level and (2) the income flat tax reform. Reducing the level of income taxation is effective in decreasing informality. A 15*pp* reduction in a tax rate reduces the share of informal workers by 6*pp*. At the same time, human capital accumulation responds mildly to the tax level change. In total, reducing the tax level from around 25% to 10% leads to a welfare gain of 16% *cev*. However, this gain comes at a cost of a substantial government budget deficit. This means that tax reduction is not a sustainable policy in the long run.

Finally, I perform a flat tax reform such that all formal workers face the same level of income taxation. I find that this policy has small effects on the aggregate levels of informality and educational attainment but it changes the composition of the labor force.

In particular, this reform flattens the U-shape of the informality level with respect to years of education. Under such a scenario, individuals enjoy a welfare gain of 4.3% which is lower compared to the policies analyzed above. However, the income tax revenue to GDP ratio stays fixed such that there is no additional deficit accumulation.

The rest of the paper is organized as follows. In the next section, I review the related literature. In Section 3, I present stylized facts about the educational attainment and informality in Brazil. Next, I discuss the model and the calibration in Section 4 and 5, respectively. Finally, I present the results and draw conclusions.

## 2 Literature Overview

This study relates to an extensive literature analyzing the impact of education subsidies on human capital accumulation and labor participation. [Krueger and Ludwig \(2013\)](#) analyze the optimal combination of progressive income taxes and education subsidies in a model with endogenous human capital, borrowing constraints, and incomplete financial markets. They find that progressive taxes distort labor supply as well as weaken the incentives to acquire education. They also claim that the latter distortion can potentially be mitigated by an education subsidy. Similarly, [Bovenberg and Jacobs \(2005\)](#) and [De Fraja \(2002\)](#) study optimal education subsidies and find that education policy offsets some of the tax-induced distortions on learning. [Guvenen et al. \(2014\)](#) abstract from the question of optimality and focus on the distortionary impact of progressive taxation on human capital accumulation which affects the income distribution and inequality.

The studies mentioned above look at education subsidies and tax distortions in a theoretical framework suitable for developed countries. However, in developing countries, the existence of a large informal sector may have important implications.

First, in the presence of informality, individuals may adjust both the intensive and the extensive margins of their labor supply such that they can reduce hours of work or operate informally as a response to government policies. Second, human capital accumulation may have asymmetric returns in the formal and in the informal sector leading to the reallocation of labor with important fiscal consequences. The empirical evidence indicates that low-educated individuals are more likely to be informal ([Arbex et al. \(2010\)](#), [Lopez Garcia \(2015\)](#), [Meghir et al. \(2015\)](#)). A hypothesis in the informality literature is that some individuals are not skilled enough to operate in the formal sector and are rationed out to informality (see [Maloney \(2004\)](#) for a discussion). However, more recent studies state that unskilled low-educated individuals rationally choose to become informal due to a comparative advantage in this sector ([Lopez Garcia \(2015\)](#), [Araujo et al. \(2013\)](#), [Fairris and Jonasson \(2016\)](#)). In particular, [Lopez Garcia \(2015\)](#) shows that human capital accumulation and preferences for job amenities explain up to 72% of transition between the informal and the formal sector.

The literature focusing on education policies and informality is surprisingly scarce. [Haanwinckel and Soares \(2017\)](#) emphasize that improvements in the labor force schooling and skill level may be more effective in reducing informality in the long term compared to other commonly used instruments. The authors document that informality is strongly responsive to the composition of the labor force. A labor force with higher levels of schooling generates incentives for firms to grow and formalize which decreases informality in the long run.

[El Badaoui and Rebiere \(2013\)](#) develop a search and matching model of a dual labor market to theoretically analyze the impact of increasing access to education on employment flows. In their model, the formal sector is reserved for educated and trained-on-the-job workers, and the informal sector is accessible to all workers. They find that a rise in access to education reduces the size of the informal sector if financed with an external subsidy. However, if an education subsidy is financed by a tax on formal sector firms, it reduces the labor market efficiency. Similarly, [Bobba et al. \(2018\)](#) build a search and matching model where firms and workers are forming formal and informal matches. Workers choose the level of schooling before entering the labor market. The authors study the role of the dual social security system in Mexico and find that changing the social security system can increase output, schooling, and long-term productivity at a small fiscal cost.

The study of [D'Erasmus et al. \(2014\)](#) focuses on the demand side for skilled and unskilled labor by firms that may operate either formally or informally. They analyze the role of institutions in shaping the demand for human capital and the level of informality. The authors show that countries with a low degree of debt enforcement and high costs of formalization are characterized by relatively lower stocks of skilled workers, larger informal sectors, and low efficiency.

Importantly, [Berniell \(2020\)](#) in a similar framework with human capital investments, occupational choice and an informal sector theoretically explains the cross-country differences in the level of entrepreneurship, informality, and human capital investment. She finds that a higher level of informality discourages human capital investments for workers but incentivizes these investments for entrepreneurs.

In contrast to the studies mentioned above, this paper analyzes education financing in the presence of an informal sector in a quantitative general equilibrium framework. I focus on individual choices and labor supply decisions. Due to a rich heterogeneity of individuals and endogenous educational and occupational choices the model allows to analyze the impact of education and tax policies on labor composition, fiscal and welfare implications. I also study the impact of progressive taxation in the set-up with endogenous human capital accumulation and endogenous informal employment contributing to the literature described above.

### 3 Empirical Evidence, Brazil

In this section, I provide empirical facts on the education system and education outcomes, as well as labor informality in Brazil.

The education system in Brazil can be roughly divided into three levels: primary, secondary and tertiary education. Primary education is mandatory for children from age 6 to 14. Secondary education (high school) is available only if the elementary school is finished and normally lasts for another three years. Tertiary education is conditional on finishing a high school degree and has different length, between three and up to eight years, depending on the area of specialization and an academic level, bachelor's or master's.

According to the [OECD \(2018\)](#) report, educational attainment in Brazil is low compared to other OECD countries. Only 69% of the 15-19 years old and 29% of the 20-24 years old are enrolled in education institutions compared to the OECD average of 85% and 42%, respectively. Strikingly, over half of Brazil's adult population (25-64 years old) have not completed secondary education, which is more than double the OECD average.

There are considerable differences in how education is financed depending on the level of education. Over 80% of primary and secondary students attend public institutions which are free of charge. On the contrary, over three-quarters of bachelor's students in Brazil attend private universities which charge a wide range of tuition fees ([OECD \(2019\)](#)). Although spending in public institutions increased by 19% between 2010 and 2016, spending per student is still below the OECD average at USD 14200. Direct spending on public institutions represent about 1% of the country's GDP. Including transfers to households, public spending on education account for 1.4% of GDP. Brazilian Campaign for the Right to Education (2018) reports that in order to reach the goals stated in the National Education Plan, Brazil has to invest three to five times more in education per student compared to the current investment.

Apart from a relatively low level of educational attainment, another important feature of the Brazilian labor market is a high level of informality. [Schneider et al. \(2010\)](#) report that the size of the shadow economy in Brazil is around 32.5% of GDP and the share of informal labor ranges between 15% and 40% depending on the definition of informality ([Meghir et al. \(2015\)](#)). In this study, I use a conservative definition of informality, such that I define a worker as informal if she reports to be an employee but has not signed a working contract and does not pay social security contributions. I abstract from the self-employed which are often considered to operate informally for two reasons. First, from the Brazilian micro data PNAD it is impossible to identify whether a self-employed works formally or informally. Second, the costs and benefits of education can be very different for the employees and the self-employed for whom the entrepreneurial talent is potentially more important than the educational attainment.

Table 1 reports the share of individuals who have completed primary (low-educated),

secondary (middle-educated) and tertiary education (highly educated) within formal and informal sectors. Almost a half of those who are formally employed have a high-school degree, around 29% are low educated and around 17% are highly educated. In the informal sector, on the contrary, the majority of workers are low-educated 49%, and the share of highly educated individuals is lower compared to the formal sector. Additionally, running a Probit regression and controlling for age, gender, region and race reveals that the level of education has negative significant effect on the probability of operating informally.<sup>2</sup> This result is consistent with the empirical studies of [Araujo et al. \(2013\)](#) and [Fairris and Jonasson \(2016\)](#) among others.

Table 1: Education Distribution within Sectors

Education	Formal	Informal	
Primary, %	28.71	48.67	
Secondary, %	54.65	40.36	
Tertiary, %	16.65	10.97	
Total, %	100	100	
Sector size, %	82.4	17.6	100

*Notes:* The table compares educational attainment between the formal and the informal sectors. Data source is PNAD from 2010 to 2015. The definition of informality is conservative and excludes self-employed, focusing solely on the informal employees.

It is important to understand how the educational attainment effects future wages of employees. In order to shed light on the education premia in the formal and informal sector, I run a wage regression for both sectors controlling for age, race and region. I focus on male workers only such that I do not have to control for gender. Table 15 illustrates the results. The dependent variable is a gross hourly wage in a particular sector. First, the coefficients for education are positive and statistically significant for formal as well as informal workers. Second, the education premium is larger in the formal sector. One additional year of studying increases the wage in the formal sector by 7.5% in contrast to 6.9% in the informal sector. [Araujo et al. \(2013\)](#) evaluates the Brazilian informal labor market and documents a similar set of facts about the education premium in the formal and the informal sector as in this section. Additionally, [Arbex et al. \(2010\)](#) conducts a detailed empirical investigation of the education premium in the informal sector.

Following these empirical facts, I build a two-sector life-cycle model where individuals invest in their human capital through education and can choose a sector of occupation, formal or informal. Education requires time and monetary input. A part of monetary costs is financed by the government. Both formal and informal workers receive an education

<sup>2</sup>See Appendix 8.1, Table 14 for the details.



premium with a higher return to human capital for the former. Additionally, formal workers face progressive labor taxes whereas informal workers escape taxation.

## 4 The Model

I build a dynamic general equilibrium overlapping generations model in the spirit of [Auerbach and Kotlikoff \(1987\)](#) with income and lifespan uncertainty as in [Huggett \(1996\)](#). Individuals go through three stages of their life: (1) the young, education period; (2) the working period and (3) retirement. In the first stage of their life, individuals make the educational choice given their asset position and innate learning ability. Education requires time and monetary investment. A certain share of education costs is subsidized by the government. In the second stage of life, individuals make an occupational choice. They may work in either a formal or an informal sector and may switch between them during their working lives. The formal sector features a higher education premium and individuals enjoy higher gross wages compared to informal employees. However, formal workers must pay progressive income taxes. Informal workers, on the contrary, do not pay taxes but bear a fixed utility cost associated with informal employment. This utility cost reflects the lack of social security in the informal sector. Individuals retire when they reach the retirement age and receive pension benefits. Additionally, there are formal and informal firms that hire formal and informal labor, respectively. The government collects tax revenues to finance its expenditures which include education subsidies, pension payouts and government consumption.

### 4.1 Households

#### Demographics

The economy is populated by overlapping generations of individuals. Individuals enter the economy at the age of  $j_0$  and face lifespan uncertainty. The conditional probability of survival from age  $j$  to age  $j + 1$  is denoted as  $\psi_j$ . The maximum possible age is  $j = J$ , with  $\psi_J = 0$ . The size of a new cohort grows at a constant rate  $n$ . Bequests  $b$  are assumed to be collected and distributed to the young population to mimic the inter-generational transfers. Individuals enter the economy with initial assets  $a_0$  drawn from a log-normal distribution  $\text{Lognormal}(\mu_a, \sigma_a)$ .<sup>3</sup> The implication of this assumption is that the young individuals start their life-cycle with different initial financial conditions.

Households differ with respect to their age  $j$ , an asset position  $a$ , an idiosyncratic productivity  $\varepsilon$ , an innate learning ability  $e$  and a level of education  $h$  which is defined after the

---

<sup>3</sup>The mean of the distribution  $\mu_a$  is defined by the accidental bequests, whereas the dispersion  $\sigma_a$  is set to match the asset distribution in Brazil. Further discussion is provided in Section 5.

first stage of life and does not change afterwards.<sup>4</sup> Asset markets are incomplete, that is, households cannot insure against shocks to working productivity  $\varepsilon$ . Moreover, individuals are borrowing constrained.

## Preferences and Endowments

Households maximize the expected sum of discounted utility given by:

$$E_{j_0} \sum_{j=j_0}^J \beta^{j-j_0} \psi_j u(c_j, \tilde{l}_j), \quad (1)$$

where  $\beta \in (0, 1)$  is the time discount factor,  $c_j$  and  $\tilde{l}_j$  is consumption and leisure at the age of  $j$  respectively. The total time endowment is normalized to one.

## Educational Choice

Individuals enter the economy with an innate learning ability  $e \in (0, 1)$  drawn from a truncated normal distribution  $N(\mu_e, \sigma_e)$  and with the initial assets  $a_0$ . The young split their time endowment between leisure  $\tilde{l}_j$ , working time  $l$  and education time  $q$ , hence  $\tilde{l}_j + l + q = 1$ . The effective time spend on education is heterogeneous and depends on the innate learning ability  $e$ . The higher the learning ability, the lower is the required time investment, reflecting the fact that for more able individuals it is easier to acquire a certain level of human capital  $h$ , hence  $q = \xi(e)h$  and  $\xi'(e) < 0$ . Apart from the time cost, education requires a monetary investment  $\kappa$  per unit of human capital  $h$ .  $\kappa$  is interpreted as an annual tuition fee and  $h$  as the amount of academic years which defines the level of education. The share  $s$  of education expenditures is covered by the government to reflect the publicly financed share of education costs. Hence, the total monetary cost of education borne by a student is  $(1 - s)\kappa h$ .

The young individuals can work  $l$  hours and receive labor income  $\varepsilon w l$  which depends on the idiosyncratic labor productivity  $\varepsilon$ , wage rate  $w$  and hours of work  $l$ . For simplicity, I assume that the wage rate in the first period is equal to the wage rate in the informal sector  $w = w^I$  and stays fixed across counterfactual experiments. Additionally, the young can borrow up to a borrowing limit  $\underline{A}$ .

Individuals choose education level, consumption, labor supply and savings in order to maximize their expected lifetime utility subject to their budget constraint.

$$V^Y(j, a, \varepsilon, e, h) = \max_{c, h, l, a'} (u(c, l, h) + \beta \psi_j E[V(j+1, a', \varepsilon', h) | \varepsilon]) \quad (2)$$

---

<sup>4</sup>For simplicity, I abstract from the human capital accumulation during the working life.

subject to

$$(1 + \tau_c)c + a' \leq (1 + r(1 - \tau_k))a_0 - (1 - s)\kappa h + \varepsilon w l \quad (3)$$

$$a' \leq \underline{A} \quad (4)$$

where  $\tau_c$  and  $\tau_k$  are the consumption and capital tax, respectively.

## Occupational Choice

Individuals enter the second stage of their life-cycle with a human capital  $h$  accumulated in the first stage. In every period of the working period, an individual decides whether to work formally or informally. The maximization problem of the working population is given:

$$V(j, a, \varepsilon, h) = \max_{sec \in \{F, I\}} \{V^F(j, a, \varepsilon, h), V^I(j, a, \varepsilon, h)\}, \quad (5)$$

where  $V^F(\cdot)$  and  $V^I(\cdot)$  are the value functions associated with becoming a formal and an informal worker, respectively.

## Formal Workers

In the formal sector, individuals receive income  $y^F$  which depends on the individual sector-specific labor productivity  $\eta^F$ , the formal market wage  $w^F$  and hours of work  $l$ . The labor productivity  $\eta^F$  consists of two components. First, an idiosyncratic labor productivity  $\varepsilon$  which is drawn from a finite-state Markov process with transition probability given by  $F(\varepsilon'|\varepsilon)$ . Second, a human capital component  $\exp(\theta^F h)$  which depends on the human capital accumulated in the first period  $h$  and a sector specific return to human capital  $\theta^F$ .

Formal workers pay consumption tax  $\tau_c$  and capital tax  $\tau_k$ , as well as progressive labor income tax  $T_l(\cdot)$  which is a function of gross income  $y^F$ .<sup>5</sup> Individuals choose consumption, labor supply and savings in order to maximize their expected lifetime utility subject to their budget constraint.

$$V^F(j, a, \varepsilon, h) = \max_{c, l, a'} (u(c, l) + \beta \psi_j E[V(j + 1, a', \varepsilon', h)|\varepsilon]) \quad (6)$$

subject to

$$y^F = \eta^F w^F l$$

$$\eta^F = \varepsilon \cdot \exp(\theta^F h)$$

$$(1 + \tau_c)c + a' = (1 + r(1 - \tau_k))a + y^F - T_l(y^F) \quad (7)$$

---

<sup>5</sup>I use a HSV functional form and calibrate the shape of the tax schedule to match empirical moments from the data.

## Informal Workers

In the informal sector, workers get income  $y^I$  which depends on the market informal wage  $w^I$ , their individual productivity  $\eta^I$  and hours of work  $l$ . As before the labor productivity  $\eta^I$  consists of two components, namely: an idiosyncratic labor productivity shock  $\varepsilon$  and a human capital component  $\exp(\theta^I h)$ . The return to the human capital in the informal sector is different from the one in the formal sector and is assumed to be lower in accordance with the empirical evidence discussed in Section 3, hence  $\theta^I \leq \theta^F$ . Importantly, informal workers do not pay the labor tax, in contrast to the formal employees. However, they bear a fixed utility cost of informality  $u_I$  (Fernandez-Bastidas (2018), Busato and Chiarini (2004)). This cost captures the lack of social security and potentially worse working conditions in the informal sector which I do not model explicitly.

$$V^I(j, a, \varepsilon, h) = \max_{c, l, a'} (u(c, l) - u_I + \beta \psi_j E[V(j+1, a', \varepsilon', h) | \varepsilon]) \quad (8)$$

subject to

$$\begin{aligned} y^I &= \eta_I w^I l \\ \eta^I &= \varepsilon \cdot \exp(\theta^I h) \\ (1 + \tau_c)c + a' &= (1 + r(1 - \tau_k))a + y^I \end{aligned} \quad (9)$$

## Retirees

After reaching the retirement age  $J_R$  individuals stop working and receive a pension benefit  $p$ . I abstract from the possible heterogeneity of pension benefits. Retirement excludes the possibility to work. Retirees choose their consumption and saving in order to maximize:

$$V^R(j, a) = \max_{c, a'} (u(c, 0) + \beta \psi_j V^R(j+1, a')) \quad (10)$$

subject to

$$(1 + \tau_c)c + a' = (1 + r(1 - \tau_k))a + p. \quad (11)$$

## 4.2 Firms

### Formal Firms

Formal firms hire formal labor  $L_F$  and rent capital  $K$  in order to produce. A representative formal firm operates under a Cobb-Douglas production function with a technological parameter  $A^F$ . Formal firms pay wage  $w^F$  to their employees. Capital depreciates at the rate  $\delta$ .

Hence, formal firms maximize their profits by solving the problem of the form:

$$\pi^F = \max_{K, L_F} (A^F K^\alpha L_F^{1-\alpha} - w^F L_F - (r + \delta)K), \quad (12)$$

where  $\alpha \in (0, 1)$  is the share of capital in the production function.

### Informal Firms

Informal firms hire only the informal labor  $L_I$  in order to produce. [La Porta and Shleifer \(2014\)](#) document that informal firms are much more labor intensive than the formal ones. I follow [Ihrig and Moe \(2004\)](#), [Busato and Chiarini \(2004\)](#) and assume that informal firms do not use capital and produce with linear production function in labor and a productivity parameter  $A^I$ .

The optimization problem of a representative informal firm reads as follows:

$$\pi^I = \max_{L_I} (A^I L_I - w^I L_I). \quad (13)$$

## 4.3 Government

The government collects revenues from the consumption and capital tax paid by all individuals. Additionally, it receives the labor income tax payments from formal workers. The revenues are spent on education subsidies, pension payouts and government consumption  $G$ .

## 4.4 Timing of Events

The sequence of events in this economy unfolds as follows:

1. Individuals enter the economy at the age of  $j = j_0$  with assets  $a_0$  and an innate learning ability  $e$ . They choose how much to consume and save, as well as how much time to work and learn.
2. In each period of the working stage, individuals choose their occupation either in the formal or in the informal sector,  $o \in \{F, I\}$ . In both sectors, the workers choose their consumption, savings and hours of work.
3. Finally, at the age of  $j = J_R$  individuals retire and receive their pension benefits  $p$ . At the age of  $j = J$  individuals die with certainty.

## 4.5 Stationary Competitive Equilibrium

For a given set of exogenous demographic parameters  $\{n, \psi_j\}$  and a given stream of government expenditures  $\{G\}$ , a competitive equilibrium is characterized by sequences of household policies and value functions  $\{c, l, h, a', o, V^Y, V^F, V^I, V^R\}$  for each state  $x \in \{j, a, \varepsilon, e, h\}$ , production plans  $\{Y, K, L_F, L_I\}$ , sequences of taxes  $\{\tau_c, \tau_k, T_l(\cdot)\}$ , education subsidy and pension benefit  $\{s, p\}$ , sequences of prices  $\{w^F, w^I, r\}$ , as well as a distribution of individuals  $\mu$  that satisfy the following conditions:

1. Taking prices as well as tax and other policy parameters as given, individuals choose their consumption, labor input, savings, investment in education and occupation in order to maximize their expected discounted utility specified by equations (2), (5), (6), (8) and (10).
2. Factor prices are determined competitively:

$$r = \alpha A^F K^{\alpha-1} L_F^{1-\alpha} - \delta,$$

$$w^F = (1 - \alpha) A^F K^\alpha L_F^{-\alpha},$$

$$w^I = A^I.$$

3. The initial total amount of assets is equal to the amount of assets left by the deceased.

$$A_0 = \int_x a'(x)(1 - \psi_j) d\mu(x).$$

4. The labor, capital and the goods markets clear:

$$L_F = \int_x \mathbf{1}^F \eta^F l(x) d\mu(x),$$

$$L_I = \int_x \mathbf{1}^I \eta^I l(x) d\mu(x),$$

$$K = \int_x a(x) d\mu(x),$$

$$Y = \int_x c(x) d\mu + G + \tilde{I},$$

where  $\mathbf{1}^F$  and  $\mathbf{1}^I$  are the indicator functions, taking the value of one if the individual is employed formally or informally, respectively.  $\tilde{I}$  is the total investment and  $Y$  denotes the total production in the economy comprising production in both sectors.

5. The government budget constraint is balanced:

$$G + \int_x \mathbf{1}^R p \cdot d\mu(x) + \int_x \mathbf{1}^Y s \cdot \kappa \cdot h(x) d\mu(x) = \int_x (\tau_k r \cdot a(x) + \tau_c c(x)) d\mu(x) + \int_x \mathbf{1}^F (T_l(y^F(x))) d\mu(x)$$

where  $\mathbf{1}^R$  and  $\mathbf{1}^Y$  are the indicator functions, taking the value of one if the individual is a retiree or a student, respectively.

6. The distribution of individuals across states is stationary:

$$\mu(x) = R_\mu [\mu(x)].$$

## 5 Calibration

My main source of data for this study is the Brazilian National Household Sampling Survey (PNAD) for a period from 2010 to 2015. PNAD is a survey conducted annually by the Brazilian National Statistics Bureau (IBGE) and investigates a wide range of population characteristics such as education, labor and income. PNAD is a cross-sectional data which I pool for calculating main calibration targets regarding educational attainment, occupation and income. For estimating a stochastic ability process, I rely on the labor force survey of Brazil, Pesquisa Mensal de Emprego (PME). The advantage of the latter is the panel dimension, however, it has a smaller coverage in terms of survey questions. Both data sets are conducted by the IBGE and are complementary to each other.

For the purpose of this paper, I focus on male individuals of age 25-65 who are household heads.<sup>6</sup> Following the literature, I define an individual to be working informally if she is an employee but did not sign a working contract (working card) with the exception for public employees who do not have to sign it ([Ulyssea 2018](#)). I do not consider individuals who report to be unpaid workers, unemployed or self-employed. It is evident that a substantial share of the self-employed in Brazil operate informally. However, the effect of education on the entrepreneurial ability may be different from the effect of education on the productivity of employees (see, e.g., [Berniell \(2020\)](#)). Hence, I abstract from the self-employed in this study.<sup>7</sup> Another possible concern is that individuals may operate formally and informally simultaneously making it difficult to categorize an individual in either of a group. In my sample, less than three percent of individuals report to have more than one job making this concern negligible. The summary statistics of my sample are presented in the Appendix.

The rest of this section is organized as follows: First, I report the functional forms and parameters which are fixed outside the model. Then I discuss internally calibrated

---

<sup>6</sup>I use this sample to derive targets for the working population in the model.

<sup>7</sup>See [Di Nola et al. \(2021\)](#) analyzing tax evasion by self-employed.

parameters which are set to match selected data targets. Finally, I present the model fit along several dimensions of the data and discuss the model performance.

## 5.1 Functional Forms and Externally Fixed Parameters

The demographic structure of the economy is set externally. The population growth  $n$ , is fixed to 1.8% which is the number used by [Jung and Tran \(2012\)](#) and refers to the annual population growth in Brazil.<sup>8</sup> The conditional surviving probabilities,  $\{\psi_j\}_{j=0}^J$ , are taken from the life-tables provided by the World Health Organization.<sup>9</sup> Individuals enter the economy at the age of  $j_0 = 15$  when they start making their education choice. Since primary education in Brazil is compulsory I assume that attending a school till the age of 15 is not a choice. I set the retirement age to 60 years which is close to the average retirement age in Brazil<sup>10</sup> and the maximum age in the economy to 90 years.

Table 2: Externally Calibrated Parameters

Parameter	Description	Value	Source
<i>Demographics</i>			
$n$	population growth	1.8%	Jung and Tran (2012)
$\{\psi_j\}_{j=0}^J$	surviving probabilities	Figure 8	WHO
$J$	maximum age	90 years	WHO
$J_R$	retirement age	60 years	World Bank (2017)
<i>Preference and ability process</i>			
$\sigma$	rel. risk aversion = 2	3.5	standard value
$\rho_\eta$	persistence	0.776	micro data (PME)
$\sigma_\eta$	standard deviation	0.354	micro data (PME)
<i>Technology and production</i>			
$\alpha$	capital share	0.36	standard value
$A^F$	TFP, formal sector	1	normalization
<i>Government</i>			
$\tau_{hsv}$	income tax, progresivity	0.05	Gobetti and Orair (2017)

The utility function  $u(\cdot)$  is defined as:

$$u(c, 1 - l) = \frac{(c^\nu(1 - \xi(e)h - l)^{1-\nu})^{1-\sigma}}{1 - \sigma}.$$

where  $\nu$  denotes the intratemporal elasticity of substitution between consumption and leisure and  $\sigma > 0$  is the risk aversion parameter. I assign a value of 3.5 to  $\sigma$  which results

<sup>8</sup>For computational purposes one period in the model is equivalent to five years. In this section I report annual values. For the computations I have to translate annual values to their 5-year period equivalents. For example, the population growth for a model period is  $(1 + n)^5 - 1 = 0.093$  or 9.3%.

<sup>9</sup>The right panel of Figure 8 shows the conditional surviving rates along age.

<sup>10</sup>On average individuals retire when they are 58 ([World Bank \(2017\)](#)).



in the coefficient of relative risk aversion being around 2 which is a standard value in the macroeconomic literature. For the time cost of education,  $\xi(e)$ , I follow [Krueger and Ludwig \(2013\)](#) and assume a linear decreasing function:  $\xi(e) = 1 - e$  where  $e \in (0, 1)$ . After the education period, a term  $\xi(e)h$  disappears since  $h = 0$ .

I estimate a stochastic process for labor productivity  $\varepsilon$  following two steps, as it is standard in the literature (see, e.g., [Guvenen \(2009\)](#) and [Heathcote et al. \(2017\)](#)). First, I regress hourly wages on the observable household characteristics such as age, education and experience in order to obtain a measure of labor income residuals  $\varepsilon_t$ . Second, I model the residuals as a first-order autoregressive process:

$$\log \varepsilon_{t+1} = \rho_\varepsilon \log \varepsilon_t + \zeta_{\varepsilon,t+1}, \quad (14)$$

where  $\zeta_{\varepsilon,t+1} \sim N(0, \sigma_\varepsilon^2)$ . I estimate this process and obtain the persistence parameter  $\rho_\varepsilon = 0.78$  and the dispersion parameter  $\sigma_\varepsilon = 0.35$ . I approximate the stochastic process in (14) by a discrete Markov chain following the procedure described in [Tauchen and Hussey \(1991\)](#).

The parameter  $\alpha$  represents the capital share in the formal sector production and is set to 0.36 which is standard in the macroeconomic literature (see, e.g., [Conesa and Krueger \(1999\)](#)). The total factor productivity of the formal sector  $A^F$  is normalized to unity.

For the calibration of the progressive income taxation, I follow [Heathcote et al. \(2017\)](#) and consider a functional form:

$$T_l(y) = y - \lambda_{hsv} y^{1-\tau_{hsv}},$$

where  $T_l(y)$  is the tax rate at the income level  $y$ ,  $\tau_{hsv} > 0$  is a measure of progressivity of the tax schedule and  $\lambda_{hsv}$  is a parameter that governs the average tax rate. I use the data from [Gobetti and Orair \(2017\)](#) who report average tax rates by income brackets in Brazil to pin down the shape of the tax function  $\tau_{hsv}$ . I recover a value of 0.05 which is consistent with the estimates by [Bick et al. \(2020\)](#).  $\lambda_{hsv}$  is calibrated internally to match the labor income tax revenue to GDP ratio in Brazil. The recovered tax function as well as the comparison to the data is presented in the [Figure 7](#).

All externally set parameters are reported on the annual basis in [Table 2](#).

## 5.2 Internally Calibrated Parameters

After setting the external parameters, there are 16 remaining parameters that are calibrated internally, namely: the discount factor  $\beta$ , the capital depreciation rate  $\delta$ , the weight on consumption in the utility function  $\nu$ , two parameters in the learning ability distribution  $\mu_e$  and  $\sigma_e$ , the education subsidy rate  $s$ , the annual tuition fee  $\kappa$ , the utility cost of informality  $u_I$ , the rate of return to human capital in the formal and the informal sectors  $\theta^F$  and  $\theta^I$ , the TFP in the informal sector  $A^I$ , the pension  $p$ , consumption and

capital taxes,  $\tau_c$  and  $\tau_k$ , the level of labor income taxation  $\lambda_{hsv}$  and the dispersion of the initial asset distribution  $\sigma_a$ . I calibrate these parameters to match 19 selected moments from the data listed in Table 3.

Table 3: Internally Calibrated Parameters

Parameter	Description	Value	Target
$\beta$	discount factor	0.97	interest rate
$\delta$	depreciation rate	0.06	capital-output ratio
$\nu$	weight on consumption	0.42	average hours worked
$\mu_e$	learning ability, mean	0.4	share of low, middle and
$\sigma_e$	learning ability, st. dev.	0.8	highly educated
$s$	subsidy rate	0.33	education expenditure to GDP
$\kappa$	tuition fee	0.35	tuition fee to GDP
$u_I$	cost parameter	0.25	share of informal workers
$\theta_F$	education premium	2.0	income profile by sector and education
$\theta_I$	education premium	1.2	income profile by sector and education
$A_I$	TFP, informal sector	0.29	income profile by sector and education
$p$	pension	0.195	pension exp. to GDP
$\lambda_{hsv}$	labor income tax, level	0.75	labor income tax revenue to GDP
$\tau_c$	consumption tax	0.23	goods and services tax revenue to GDP
$\tau_k$	capital tax	0.08	property tax revenue to GDP <sup>a</sup>
$\sigma_a$	initial asset distribution, st. dev.	1.75	Gini wealth

<sup>a</sup>The major components of property income are interest, dividends, and rent.

It is well understood that all the model parameters affect all the targets, but we can nonetheless outline which data moment is most informative about a specific parameter. The interest rate and the capital-output ratio identify the discount factor  $\beta$  and the depreciation rate  $\delta$ . The weight on leisure  $1 - \nu$  is pinned down by matching the average labor supply in Brazil, measured as the percentage of time endowment devoted to market work.

The parameters for the distribution of the innate learning ability  $\mu_e$  and  $\sigma_e$  are set to match the share of low, middle and highly educated individuals in the data. In the data, I define low-educated as those who have attained primary mandatory education, middle-educated are those who have a high-school degree and highly-educated are those who have acquired tertiary education. In the model, I define low-educated as those who choose to invest zero hours in education. My model starts at the age of 15 such that the primary education has been attained already. I consider this as a reasonable assumption since the primary education is compulsory in Brazil and is not a choice. In the model, the educational choice is the level of education defined by the academic years. In Brazil, high-school and tertiary education take on average 8 years, 3 of which are required to complete a high-school degree and another 5 years to finish a university degree. I define

the middle-educated in the model as those who have completed  $(3/8) * 100\% = 37.5\%$  academic years which can be interpreted as a high-school degree. Highly educated in the model are those who have attained more than 37.5% of possible academic years.<sup>11</sup>

There is no direct value for the share of public education financing in Brazil, hence, the subsidy rate  $s$  is calibrated to match the government education expenditure to GDP ratio. The annual tuition fee  $\kappa$  is identified by matching the average tuition fee to GDP per capita in Brazil.

The utility cost of informality  $u_I$  identifies the share of informal workers in the economy. The education premia parameters  $\theta_F$  and  $\theta_I$  as well as the TFP in the informal sector  $A_I$  (three parameters) are used to match the average gross hourly wages in two sectors for three education levels (6 moments).

I determine the pension in the economy  $p$  so that the pension expenditure to GDP ratio is consistent in the data and in the model. The value for the labor income tax level  $\lambda_{hsv}$  is identified by matching the labor income tax revenue to GDP ratio. Finally, I set the parameters for the consumption and capital tax to 23% and 8%, respectively, to match the tax revenues to GDP ratio for consumption and capital tax.<sup>12</sup>

To capture the heterogeneity of assets at the time when the educational decision takes place, I assume that  $a_0$  is drawn from a log-normal distribution  $\text{Lognormal}(\mu_a, \sigma_a)$ . The mean of the distribution  $\mu_a$  is defined by the accidental bequests left by the deceased, whereas the dispersion parameter  $\sigma_a$  is identified by matching the asset Gini coefficient and the share of assets held by the bottom 90% of population. The recovered value is  $\sigma_a = 1.75$ . Table 3 summarizes the recovered values for the internal parameters. The presented values are annualized.

### 5.3 Model Fit

Before analyzing the mechanism of the model and discussing the results, I present how the model performs in terms of matching the data and discuss its limitations.

The model fit is presented in Table 4. The annual interest rate in the model is close to the annual real interest rate in the data. The latter is taken from the data provided by the Federal Reserve Bank of St. Louis.<sup>13</sup> The capital-output ratio as well as the labor supply in the model and in the data are also consistent. The data target for the capital-output ratio is taken from the empirical study of Filho (2002). The labor supply is defined as the percentage of time endowment devoted to market work. The data target was calculated

---

<sup>11</sup>Note that the level of education  $h$  is a discrete variable taking limited amount of values, hence the time invested in education  $q$  should not be interpreted as hours but rather as years invested in education.

<sup>12</sup>The empirical counterpart for the consumption tax is the tax on goods and services whereas for the capital tax - the property tax.

<sup>13</sup>Note, that I take the nominal interest rate of Treasury bills in Brazil and adjust for the inflation over the period of 2010-2015. This time range is chosen to stay consistent with the period I use in PNAD.

using PNAD for a pooled sample of workers both formal and informal during the period of 2010-2015.

Table 4: Model Fit, Targeted Moments

<b>Moments</b>	<b>Data</b>	<b>Model</b>	<b>Source</b>
Interest rate	5%	5.2%	FRED
Capital-output ratio	3	3.05	Filho (2002)
Labor supply	36.32%	36.46%	PNAD (2010-2015)
Low educated	32.22%	30.23%	PNAD (2010-2015)
Mid. educated	52.13%	52.2%	PNAD (2010-2015)
Highly educated	15.65%	17.56%	PNAD (2010-2015)
Educ. exp. to GDP	1.4%	1.24%	OECD
Tuition fee to GDP per cap.	91.43%	95.54%	World Bank
Share of informal workers	17.6%	17.02%	PNAD (2010-2015)
Pension exp. to GDP	11%	11.51%	OECD
Labor tax rev. to GDP	11%	11.34%	IMF <sup>a</sup>
Consumption tax rev. to GDP	14.2%	14.26%	IMF
Capital tax rev. to GDP	1.3%	1.41%	IMF
Bottom 90%	50.00%	50.3%	Statista <sup>b</sup>
Gini wealth	72.86	74.36	Fortune <sup>c</sup>

<sup>a</sup>data-imf.org. Government Finance Statistics. Brazil 2010-2015.

<sup>b</sup><https://www.statista.com/statistics/754724/wealth-distribution-income-share-brazil/>

<sup>c</sup><http://fortune.com/2015/09/30/america-wealth-inequality/>

Importantly, the model captures the shares of low, middle and highly educated individuals reasonably well. The definition of education categories in the data and in the model are explained above. Figure 9b shows the recovered distribution of the innate learning ability  $e$ .

OECD (2019) reports that the direct spending on public education institutions in Brazil accounts for around 1% of GDP in Brazil. Including transfers to households raises this number to 1.4% of GDP. The model generates a value of 1.24%. The tuition fee for tertiary education in Brazil is around 9500EUR<sup>14</sup> which is 10686USD with the exchange rate of 1USD = 0.89EUR. The average GDP per capita in Brazil over the period of 2010-2015 is 11688USD (World Bank data base). Hence, the tuition fee to GDP per capita ratio in Brazil is  $10686/11688 = 0.914$  or 91.43%. The model counterpart is 95.54%. Admittedly, the tuition fee may be heterogeneous for different individuals. Since I do not have information about the distribution of the tuition fees in Brazil I target the average value.

<sup>14</sup><https://www.master-and-more.eu/en/how-much-does-education-cost-around-the-world/>

Table 5: Model Fit, Wage Profile across Sectors and Education Levels

Moments	Data	Model
Formal		
low-educated	0.64	0.55
mid-educated	0.84	1.03
high-educated	2.44	2.10
Informal		
low-educated	0.49	0.14
mid-educated	0.74	0.45
high-educated	2.18	1.93

Gross hourly wages by a respective group normalized by the mean gross hourly wage in population. Data source: PNAD (2010-2015).

The share of the informal employees in the economy is matched well. I calculate the share of informal workers from PNAD using a conservative definition of an informal worker described above. Around 17.6% of private sector employees did not sign a working contract and operate informally.

The model generates the pension-expenditure to GDP ratio of 11.51% which corresponds to the value in the data of 11% reported by [OECD \(2016\)](#). The consumption and capital tax revenues to GDP ratios in the data and in the model are consistent. The empirical counterpart for the consumption tax is the tax on goods and services whereas for the capital tax, the property tax. The data targets are reported by the IMF in Government Finance Statistics. Labor income tax revenue consists of taxes on payroll and workforce as well as social contributions. The labor income tax revenue to GDP ratio is equal to 11% in the data and 11.34% in the model.

According to the empirical evidence, the poorest 90% of the population in Brazil hold around half of the total wealth and the remaining 10% of the population hold another half. Moreover, the wealth Gini coefficient is around 72.86%. The model replicates these numbers with the dispersion parameter  $\sigma_a = 1.75$ . [Figure 9a](#) shows the recovered distribution of initial assets.

[Table 5](#) shows the fit of gross hourly wages, normalized by the mean gross hourly wage in the economy, across sectors of production for different education groups. There are 6 targeted moments and only three parameters in the model I use to match them. Specifically, I calibrate returns to human capital in the formal and informal sectors  $\theta^F$  and  $\theta^I$  as well as the TFP in the informal sector  $A^i$ . The latter governs the relative level of wages in the informal sector. The model outcome is the result of multiple endogenous factors such as the market wages in both sectors, human capital accumulation and self-selection into the sector of production. Given the complexity of interactions, the model performance is satisfactory but it should be taken as a model's limitation that the wages

of low-educated informal workers are lower in the model compared to the data.

Table 6: Distribution of Education within Sectors

Education	Formal		Informal	
	Data	Model	Data	Model
Low, %	28.71	23.72	48.67	61.98
Middle, %	54.65	57.6	40.36	25.91
High, %	16.65	18.68	10.97	12.11
Total, %	100	100	100	100

Table 6 illustrates the within sector composition of educational attainment in the data and in the model which were not explicitly targeted. The model's shares of low, middle and highly educated individuals in the formal sector are close to the data counterparts. However, the model exaggerates the share of low educated individuals in the informal sector. Additionally, Table 7 presents the level of informality by the level of education. Both in the data and in the model, the share of informal workers is the highest for the lowest level of education. Interestingly, the level of informality does not decrease with the accumulation of the years of education but rather follows a U-shape. The share of the informal workers drops for the group of the middle-educated but then rises again for the highly educated individuals. I describe how the model can explain this pattern in Section 6.

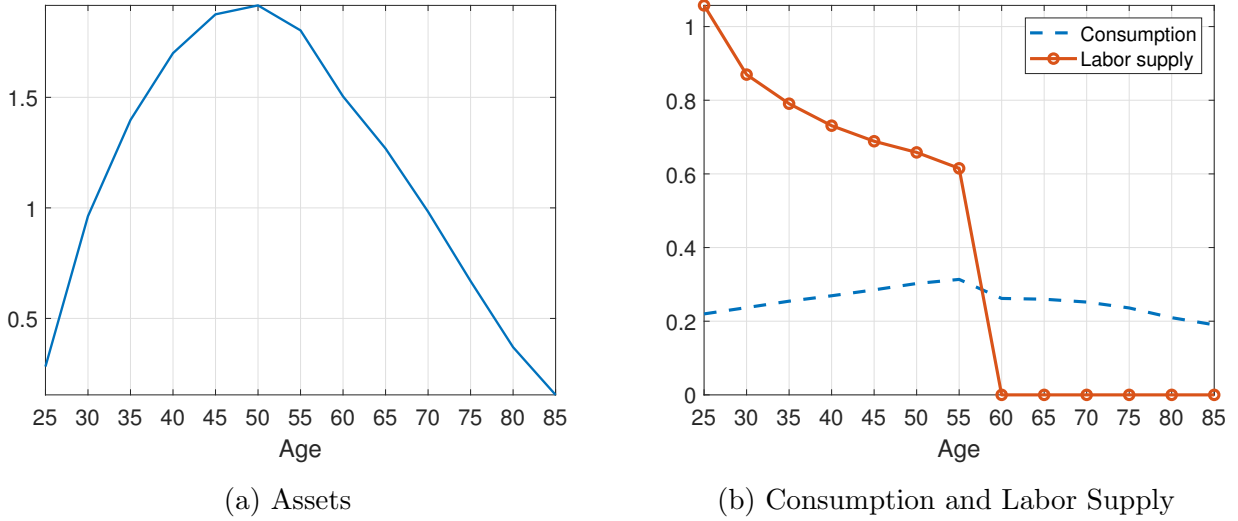
Table 7: Level of Informality by Education Level

	Data	Model
Low, %	26.56	34.9
Middle, %	13.62	8.45
High, %	15.22	11.73

## 6 Results

In this section, I first discuss the aggregate performance of the model. In particular, I explain the life-cycle patterns of consumption, labor supply, and aggregate savings for the whole population. Next, I present the optimal educational and occupational choices in the model and highlight their interrelation. Finally, I assess the outcome of the policy experiments which include an increase of the education subsidy and a change in the tax schedule. For the latter experiment, I consider both a change in the level of taxation as well as a change in the tax schedule progressiveness.

Figure 1: Means over the Life-cycle



*Notes:* The left panel illustrates the mean asset accumulation over the life-cycle. The right panel shows the mean life-cycle profiles of consumption and labor supply measured in efficiency units for the whole population. Both figures show the results in model units.

## 6.1 Aggregate Performance of the Model

Figure 1 illustrates average life cycle profiles of consumption, labor supply in efficiency units, and asset accumulation for the whole population. The outcome is standard for an OLG model.

Individuals accumulate assets during their working period to insure themselves against negative income shocks as well as to smooth their consumption in the retirement period. Figure 1a shows that individuals start to deplete their accumulated assets after the retirement age. Since there is no bequest motive in the model the savings in the last period of life drop to zero.

The labor supply follows a declining pattern over the life cycle due to the wealth effect. With higher accumulated assets workers tend to value leisure more and work fewer hours. Figure 1b shows labor supply measured in efficiency units which comprises both hours worked as well as the labor productivity. Figure 13 shows hours of work only. After reaching the retirement age individuals stop working, hence the labor supply reduces to zero. With the increasing demand for leisure over the life-cycle, the household favors an increasing consumption path. During retirement, individuals receive pension benefits that are lower than their labor income, and the average consumption level decreases.

Assets, consumption, and labor supply separately for the formal and the informal employees as well as hours of work and the level of informality over age are reported in the Appendix 8.3.

## 6.2 Educational and Occupational Choice

In this subsection, I explain an educational and an occupational choice of individuals which are two key elements of the model. The left panel of Figure 2 shows the level of education attained by the asset position for different levels of the innate learning ability. The level of education is normalized between zero and one and is a discrete variable. We should think of the level of education as a number of academic years such that a value of 0.2 corresponds to one year<sup>15</sup>. To attain a certain level of education an individual has to invest both time and money. The benefit of education is the education premium which ensures higher wages in the future. A high learning ability  $e$  means that a student learns faster than her peers and the effective time cost of education is comparatively lower. Hence, students with a high learning ability attain a higher education level. Similarly, the higher the initial amount of assets the higher is the chosen education level, since for wealthy individuals education is more affordable. It is important to note that after a certain level of assets, low-learning-ability students choose a lower level of education which is a result of the wealth effect. For wealthy individuals, leisure becomes more important than the education premium in the future. The distribution of the initial assets and the innate learning ability are presented in Figure 9a and Figure 9b, respectively. It is important to understand the relationship between educational and occupational choices. Figure 2b depicts occupational choice for different levels of education and labor productivity. The value of one refers to the formal occupation whereas the value of zero to the informal occupation.<sup>16</sup> The education level affects the occupational choice differently depending on the labor productivity  $\varepsilon$ . Low educated and low productive workers choose to be informal. For them, the difference between the education premia in the formal and in the informal sectors is relatively small. Hence, the untaxed informal sector is more attractive. Highly educated workers who face a high productivity shock also choose to be informal. For them the difference in the education premia is high but the increase in taxation due to its progressive nature prevents them from operating formally. Individuals with average characteristics like those who have (1) average labor productivity and average educational attainment, (2) a low level of education but high labor productivity, or (3) a high level of education but low labor productivity choose to operate formally. For them, a higher education premium in the formal sector out-weights the labor income taxation.

Table 6 and Table 7 presented above show the equilibrium outcome of the educational and occupational choices in the model and data. The interaction between the education premia in both sectors and the progressive taxation results in a U-shape pattern of the informality level along educational attainment presented in Table 7.

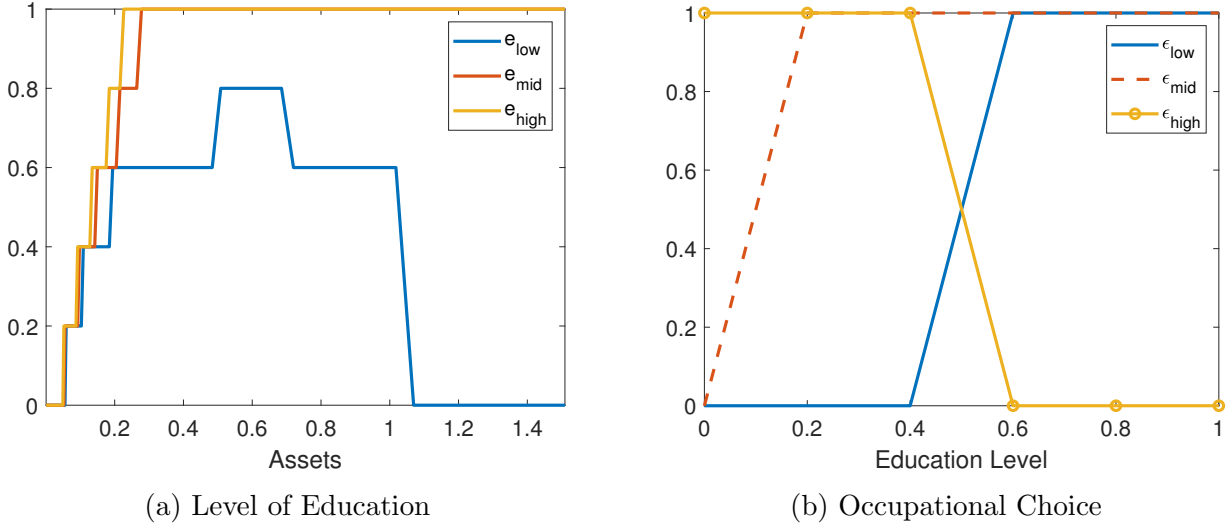
---

<sup>15</sup>Note that one period in the model is five years.

<sup>16</sup>All other variables of the state space are fixed to their average values.



Figure 2: Educational and Occupational Choice



*Notes:* The left panel shows the level of education for different innate learning abilities and assets. The right panel depicts an occupational choice for different education level and labor productivity  $\epsilon$ . The value of one means the formal occupation whereas the value of zero means the informal occupation.

### 6.3 Policy Experiments

In this subsection, I analyze the effects of different policy reforms. In the first step, I study the impact of an increase in the education subsidy on educational attainment, occupational choice, and aggregate outcomes. Next, I analyze the effects of tax policies on educational and occupational choices with a focus on fiscal and welfare implications.

#### Education Subsidy

In this experiment, I vary the subsidy rate  $s$  from 0 to 1 which means that the government finances from 0 up to 100% of education costs. The results are presented in Table 8.

Subsidizing education makes it more affordable and individuals become more educated. If the government carries all the monetary costs of education (subsidy = 1), the share of highly-educated individuals increases from around 17% to over 68% and the share of low-educated individuals drops to zero. Figure 14b compares the education policy functions for the benchmark economy and an economy where the subsidy level is 100%. Individuals with all types of innate learning ability increase their human capital investment.

In general, increasing educational attainment reduces the share of informal employees. Increasing the subsidy from the benchmark value of 0.33 to 0.9 reduces informal employment from 17.02% to 8.63%. Increasing the subsidy further to 100% drives the share of informality slightly up to 8.7%. This result indicates that the increase in the education level does not necessarily lead to expanding formalization. It does so, only if the expected

gain from a higher education premium in the formal sector out-weights the tax cost.

The combination of higher human capital accumulation and the reallocation of labor from the informal to the formal sector leads to multiple aggregate effects. First, individuals become richer due to the increased human capital which affects individual productivity and wages. Hence, workers can accumulate more assets and reduce hours of work. The increase in asset supply reduces the interest rate and increases the formal market wage in the economy. On the contrary, the increase in the formal labor supply reduces the formal market wage such that the resulting wage rate remains relatively constant. The increase in both physical and human capital leads to a considerable increase in GDP.

Welfare measured in terms of consumption equivalent variation (cev)<sup>17</sup> increases as the subsidy is raised. This result is driven by three forces: (1) the utility of consumption, (2) the disutility of labor, and (3) the disutility of operating informally. First, with the increase in the education subsidy, individuals become richer and can afford more consumption. Second, workers reduce their hours of work from around 36% of their time endowment in the benchmark economy to nearly 28% in the economy with the 100% subsidy level. Third, the reduction in informality decreases the number of individuals suffering the utility cost associated with worse working conditions and a lack of social security in the informal sector. All in all, the welfare gain from subsidizing education is large. To keep individuals indifferent between the two economies where the subsidy rate is 0.33 and 1, one has to increase the consumption of the benchmark economy by 94% or to increase the income by 32%.

Figure 3 shows the fiscal outcomes of the education policy. The increase in the subsidy rate increases the education expenditures exponentially. The blue dashed line shows that education expenditures measured in model units increase from zero to 3 as the subsidy increases from zero to 100%. The education subsidy to GDP ratio increases from 1.24% to 6.29% (see an extended Table in the Appendix). At the same time capital, consumption, and income tax revenues increase. The latter increase due to an increase in the tax base from formalization and an increase in income due to a higher education level. The considerable expansion of consumption and income tax revenues covers the raise in education expenditures. This means that the education subsidy is self-financed in the long run and additional tax revenues can be redistributed back to the population to further increase their welfare.

Figure 3b shows the difference between the total tax revenues and the total expenditures. The positive value means that there is a fiscal surplus whereas the negative value means that there is a fiscal deficit. Since the increase in the education subsidy leads to

---

<sup>17</sup>I report this statistic for the working population only since the young population suffers a substantial disutility due to a very low consumption level and high time investment in education. The value function for the young is times lower compared to the working population and a slight change in their utility drives the total result.

fiscal surpluses, in Table 9 I show the results for the education policy in which the labor taxation is reduced to clear the government budget. The results with fiscal neutrality are qualitatively similar to those explained above but quantitatively lead to even higher welfare gains.

Table 8: Education Subsidy Change

Subsidy $s$	0.0	0.1	0.2	0.3	0.33	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Share of inf., %	21.97	21.16	19.09	17.29	17.02	16.39	13.68	12.75	9.80	9.04	8.63	8.70
Low educated, %	52.08	49.53	37.97	30.67	30.23	28.49	15.22	12.71	4.42	1.40	0.47	0.39
Mid. educated, %	36.59	37.84	47.09	52.46	52.20	50.36	60.52	58.48	59.69	54.62	43.86	30.86
Highly educated, %	11.34	12.63	14.94	16.87	17.56	21.15	24.27	28.81	35.88	43.97	55.67	68.75
Interest rate, %	5.28	5.24	5.24	5.21	5.20	5.17	5.18	5.12	5.12	5.03	4.97	4.87
Wage ratio	1.72	1.73	1.73	1.73	1.74	1.73	1.73	1.74	1.74	1.75	1.76	1.77
GDP	0.33	0.33	0.35	0.36	0.37	0.37	0.39	0.41	0.43	0.45	0.47	0.50
Hours work, %	37.88	37.64	37.10	36.54	36.36	35.87	35.14	34.26	33.51	32.18	30.48	28.36
Capital	0.20	0.20	0.21	0.22	0.22	0.23	0.24	0.25	0.26	0.27	0.29	0.30
Welfare, hev	-0.06	-0.05	-0.03	-0.01	0.00	0.01	0.06	0.09	0.14	0.19	0.25	0.32
Welfare, cev	-0.14	-0.12	-0.07	-0.01	0.00	0.03	0.15	0.22	0.37	0.52	0.69	0.94

*Notes:* This table summarizes the education policy effects. A gray shaded column refers to the benchmark economy with the subsidy level of 33%. The range for the subsidy goes from zero to 100%. All tax rates remain the same as in the benchmark economy. Prices are allowed to adjust.

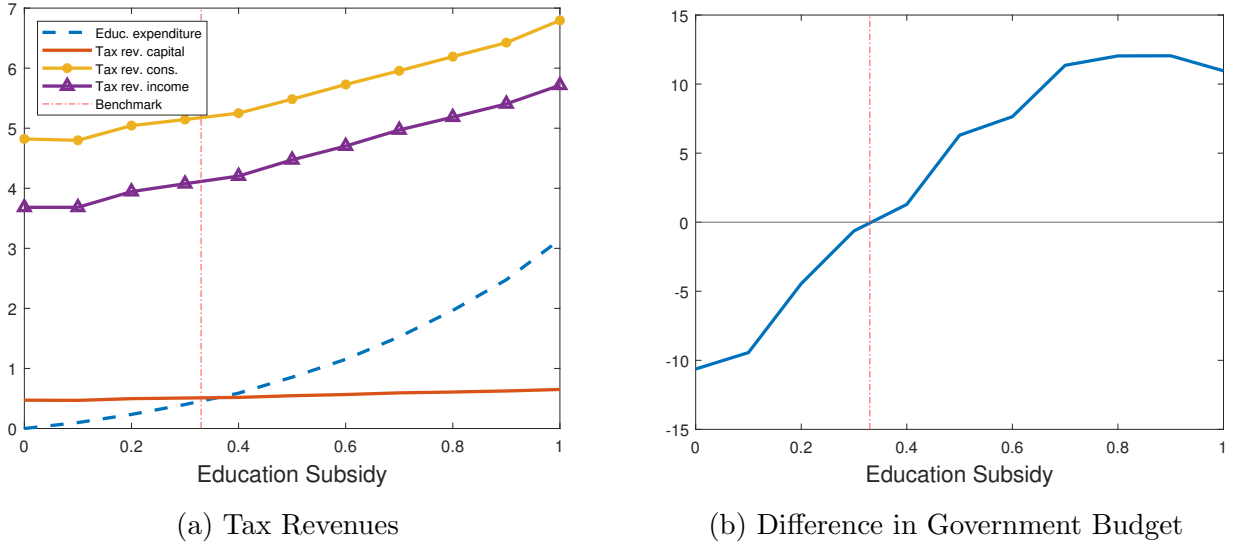
Table 9: Education Subsidy Change, Fiscal Neutrality

Subsidy $s$	0.0	0.1	0.2	0.3	0.33	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Share of inf., %	29.80	29.13	23.52	17.29	17.02	16.39	12.78	11.91	8.71	7.90	7.47	7.38
Low educated, %	52.29	49.86	38.04	30.67	30.23	28.49	15.23	12.70	4.41	1.38	0.45	0.37
Mid. educated, %	36.18	37.14	46.86	52.46	52.20	50.36	60.95	58.81	60.07	55.41	44.59	31.66
Highly educated, %	11.53	12.99	15.11	16.87	17.56	21.15	23.83	28.49	35.52	43.21	54.96	67.97
Interest rate, %	5.19	5.16	5.20	5.21	5.20	5.17	5.15	5.09	5.06	4.99	4.93	4.84
Wage ratio	1.73	1.74	1.73	1.73	1.74	1.73	1.74	1.74	1.75	1.75	1.76	1.77
GDP	0.31	0.32	0.34	0.36	0.37	0.37	0.40	0.41	0.44	0.46	0.48	0.50
Hours work, %	36.93	36.64	36.40	36.54	36.36	35.87	35.23	34.30	33.53	32.16	30.42	28.17
Capital	0.18	0.19	0.20	0.22	0.22	0.23	0.24	0.25	0.27	0.28	0.29	0.31
Welfare, hev	-0.09	-0.08	-0.04	-0.01	0.00	0.01	0.07	0.10	0.16	0.21	0.27	0.34
Welfare, cev	-0.20	-0.18	-0.10	-0.01	0.00	0.03	0.17	0.25	0.42	0.57	0.75	1.02

*Notes:* This table shows the education policy effects with fiscal neutrality. An income tax rate  $\lambda_{hsv}$  adjusts to clear the budget. A gray shaded column refers to the benchmark economy with the subsidy level of 33%. Prices are allowed to adjust.

For completeness, Table 16 presents the results of the education policy in partial equilibrium such that the prices of labor and capital do not adjust. Both qualitatively and quantitatively the results are similar to those in general equilibrium (see Table 8).

Figure 3: Fiscal Outcomes of Education Policy



*Notes:* These figures display the fiscal effects of the education policy which increases the subsidy rates. A red dashed vertical line refers to the benchmark economy. The left panel shows education expenditures and tax revenues from different sources in model units multiplied by 100. The right panel illustrates the change in the fiscal balance such that positive values refer to fiscal surpluses whereas negative values are fiscal deficits.

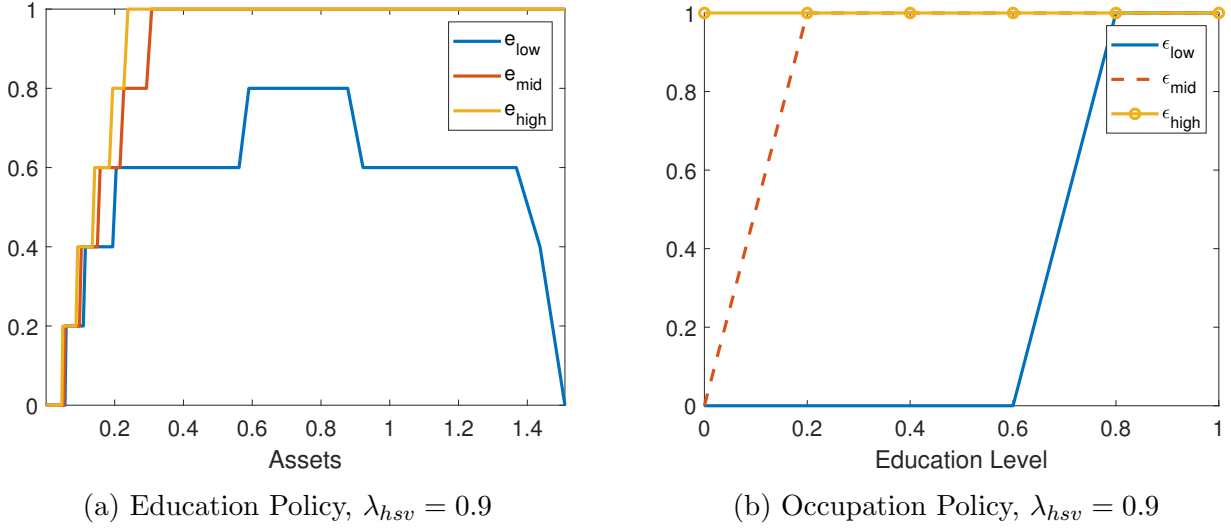
### Tax Schedule, Income Tax Level

In this subsection, I analyze the effects of tax policies on the occupational choice and human capital accumulation. I focus on two policies: (1) the change in the income tax level and (2) the income flat tax reform.

Table 10 shows the results of the income tax level reduction. Note that the parameter  $\lambda_{hsv}$  in the tax function defines the level of taxation. The higher the  $\lambda_{hsv}$  is, the lower is the average tax rate. The value of  $\lambda_{hsv}$  in the HSV specification can be roughly interpreted as the tax level of  $1 - \lambda_{hsv}$ . Hence, I vary the tax level from around 10% to 40% with the benchmark value of 25% ( $\lambda_{hsv} = 0.75$ ).

Decreasing the level of taxation favors formalization but has a modest effect on human capital accumulation. The change from the benchmark tax level of 25%, to the tax level of 10%, reduces informality by around 6pp. The share of low educated individuals stays relatively constant whereas the share of highly educated slightly reduces. In order to understand this result I plot educational and occupational choices for  $\lambda_{hsv} = 0.9$  in Figure 4. Comparing the left panel of Figure 4a with the benchmark educational policy in Figure 2a shows that the tax reduction has little effect on the educational choice of the middle and highly talented students. However, those with a low level of innate learning ability and relatively high assets invest more time in education. Additionally, Figure 4b depicts

Figure 4: Educational and Occupational Choices, Labor Tax Reduction



*Notes:* The left panel shows the level of education for different innate learning abilities and assets. The right panel depicts an occupational choice along educational attainment for different labor productivity  $\varepsilon$ . The value of one means the formal occupation whereas the value of zero means the informal occupation.

the occupational choice for  $\lambda_{hsv} = 0.9$ . Comparing with the benchmark economy, highly productive individuals (yellow line with circles) are more likely to operate formally for high levels of education (x-axis), whereas low productive workers (blue solid line) with average education are more likely to be informal. All in all, income tax level reduction has little effect on the educational choice but enhances formalization for highly educated individuals.

As a response to lower income taxation, individuals work on average more hours and accumulate more assets. As a result, output increases in the economy. Reducing income taxation leads to considerable welfare gains of up to 6% measured in Hicksian equivalent variation (hev) and 16% measured in consumption equivalent variation (cev). As before this result is an outcome of three channels. First, the increase in net income of the formal workers leads to a welfare gain. Second, the increase in hours work reduces welfare due to forgone leisure. Third, the reallocation from the informal to the formal sector leads to a welfare gain, due to better working conditions and social security in the formal sector.

Although the welfare effects of subsidizing education and reducing income taxation are similar the fiscal outcomes are quite different. Figure 5a depicts tax revenue changes in model units for all types of taxes. Reducing the level of taxation from the benchmark value of 25% to 10% decreases income tax revenues by a factor of four. It is important to note that this reduction is mitigated by formalization and an increase in hours worked. Hence, the change in the tax base cannot compensate for a drop in the tax rate leading

to considerable fiscal deficits shown in Figure 5b. This fiscal deficit has to be financed using another type of tax or by accumulating government debt which is not sustainable in contrast to the education policy discussed above. Table 17 reports the results in partial equilibrium. In general, the outcome is similar to the one presented in Table 10.

Table 10: Income Tax Level Change

$\lambda_{hsv}$	0.60	0.65	0.70	0.75	0.80	0.85	0.90
Share of inf., %	30.81	26.79	23.72	17.02	13.77	11.93	11.00
Low educated, %	30.73	30.44	30.28	30.23	30.19	30.22	30.29
Mid. educated, %	49.99	50.42	51.59	52.20	52.77	53.68	54.03
Highly educated, %	19.29	19.14	18.13	17.56	17.04	16.11	15.68
Interest rate, %	4.84	5.10	5.17	5.20	5.17	5.03	4.85
Wage ratio	1.77	1.74	1.73	1.73	1.73	1.75	1.77
GDP	0.33	0.34	0.35	0.37	0.37	0.38	0.39
Hours work, %	35.96	35.50	35.35	36.36	36.92	37.39	37.62
Capital	0.18	0.20	0.21	0.22	0.23	0.25	0.26
Welfare, hev	-0.07	-0.04	-0.02	0.00	0.02	0.04	0.06
Welfare, cev	-0.15	-0.10	-0.06	0.00	0.05	0.11	0.16

*Notes:* This table provides the results of the tax level change policy. Parameter  $\lambda_{hsv}$  defines the level of income taxation in the economy. The higher the  $\lambda_{hsv}$ , the lower is the income tax level. A gray shaded column refers to the benchmark economy with  $\lambda_{hsv} = 0.75$ . All other tax rates remain unchanged. Labor and capital prices adjust.

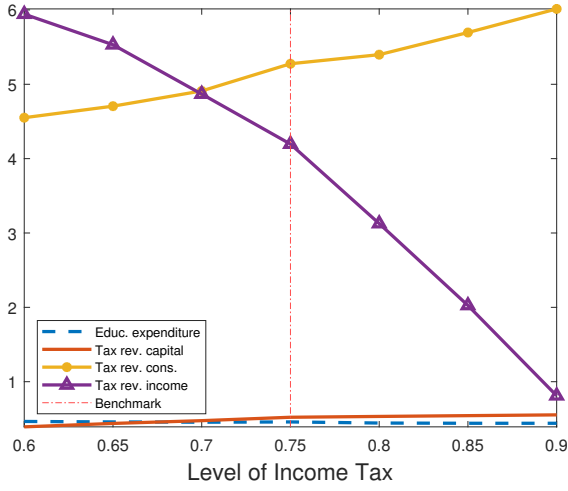
## Tax Schedule, Income Tax Progressivity

In this subsection, I perform a flat tax reform and analyze the consequences in terms of the aggregate outcomes. Instead of varying the progressivity parameter  $\tau_{hsv}$  of the HSV tax function, I analyze the extreme scenario where everyone in the formal sector faces the same labor income tax rate such that  $\tau_{hsv} = 0$ .<sup>18</sup> In doing so, I pin down the new tax rate  $\lambda_{flat}$  which matches the income tax revenue to GDP ratio in Brazil. The recovered level is 18.5% which is lower compared to an approximate tax level in the benchmark economy  $1 - \lambda_{hsv} = 1 - 0.75 = 0.25$  or 25%.

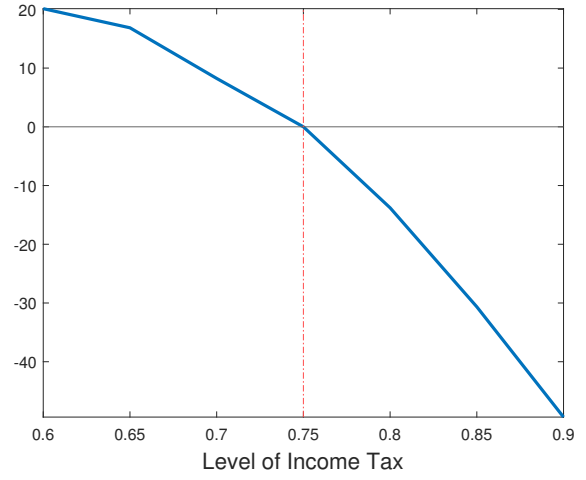
Figure 6 presents an educational and occupational choice in the economy with a flat tax rate. Compared to the benchmark economy, low-talented students with high assets choose to invest more in their human capital. In other words, the flat tax reform mitigates the wealth effect for low learning ability individuals. However, due to the skewness of the asset

<sup>18</sup>Varying the HSV progressivity parameter may result in controversial conclusions since the HSV tax function is negative for low levels of income. In other words, the change in  $\tau_{hsv}$  changes the tax rate for the rich and the tax subsidies to the poor at the same time.

Figure 5: Fiscal Outcomes of Tax Policy



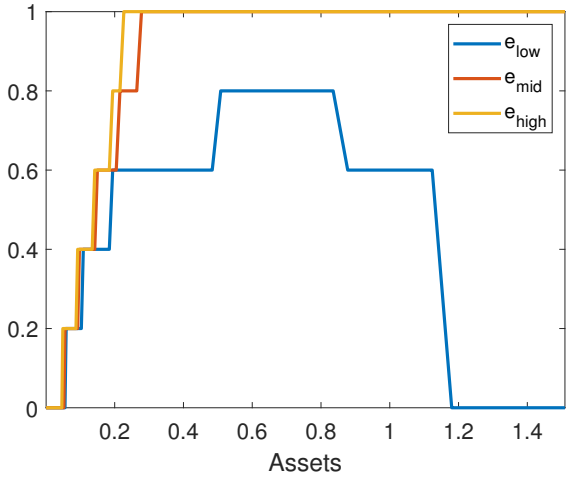
(a) Tax Revenues



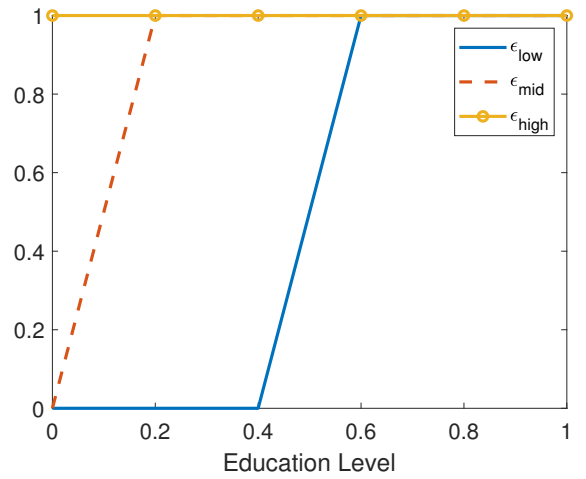
(b) Difference in Government Budget

*Notes:* These figures display the fiscal effects of the tax policy which reduces the income tax level. A red dashed vertical line refers to the benchmark economy. The left panel shows education expenditures and tax revenues from different sources in model units multiplied by 100. The right panel illustrates the change in the fiscal balance in percents such that positive values refer to fiscal surpluses whereas negative values are fiscal deficits.

Figure 6: Educational and Occupational Choices, Flat Tax Reform



(a) Education Policy, Flat Tax



(b) Occupation Policy, Flat Tax

*Notes:* The left panel shows educational choice for different innate learning abilities and assets. The right panel depicts an occupational choice along educational attainment for different labor productivity levels  $\epsilon$ . The value of one denotes the formal occupation whereas the value of zero indicates the informal occupation.

Table 11: Flat Tax Reform

	Benchmark	Flat Tax	Flat Tax, PE
Share of inf., %	17.02	16.34	16.5
Low educated, %	30.23	30.1	30.05
Mid. educated, %	52.20	52.43	52.46
Highly educated, %	17.56	17.47	17.48
Interest rate, %	5.20	5.08	5.20
Wage ratio	1.74	1.74	1.74
GDP	0.37	0.38	0.38
Hours work, %	36.36	36.21	36.18
Capital	0.22	0.24	0.25
Welfare, hev	0.00	0.017	0.018
Welfare, cev	0.00	0.041	0.043
Tax rev. income to GDP	11.34	11.35	11.28
Flat tax rate	-	0.185	0.185

*Notes:* This table reports the effects of the flat tax reform. In the contra-factual economy the government substitutes the status-quo tax-transfer schedule with a flat income tax which reaches the same income tax revenue to GDP ratio as in the benchmark economy. The recovered value is 18.5%. All other policy parameters are kept unchanged.

Table 12: Level of Informality by Education Level

	Benchmark $s = 0.33 \lambda = 0.75 \tau = 0.05$	Education Reform $s = 1$	Reduced Tax $\lambda = 0.9$	Flat Tax $\tau = 0$
Low, %	34.9	1.28	25.32	36.17
Middle, %	8.45	5.97	4.85	8.05
High, %	11.73	9.97	4.49	7.06

*Notes:* This table reports the share of informal workers by three groups of education. As before the low-educated are those who have attained primary education, the middle-educated - a high school degree, the highly-educated - tertiary education. The second column refers to the benchmark economy with subsidy rate of 33% and tax function parameters  $\lambda = 0.75$  and  $\tau = 0.05$ . The third column illustrates the share of informal employees by education level in the economy where 100% of education costs are subsidized. The next column shows the informality pattern for an economy with reduced labor tax rate  $\lambda = 0.9$ . The last column refers to the flat tax scenario with  $\tau = 0$ .



distribution, the change in the education policy function does not lead to a substantial increase in the education outcomes. Figure 6b shows that highly educated with high ability operate formally. Recall that in the benchmark economy with progressive taxation, this type of worker was choosing the informal sector. With the flat income tax, highly educated and highly productive workers face a lower tax rate and have incentives to operate formally.

Table 11 reports the aggregate statistics for the flat reform scenario. Flattening the tax schedule leads to mild effects in terms of educational outcomes but incentivizes formalization. Note that in the benchmark economy poor formal workers receive a tax subsidy. In the contractual economy, tax transfers are absent such that poor individuals have no incentive to operate formally. Hence, the recovered number for the share of informality 16.34% masks the reallocation of poor individuals into the informal sector and rich individuals into the formal sector. Additionally, Table 12 shows the level of informality by educational attainment. Recall that in the benchmark economy the level of informality follows a U-pattern along the education dimension. Given that the flat tax reform does not change the aggregate levels of informality and educational attainment much, it is important to note that a flat tax reform eliminates the U-shape pattern.

Hours of work respond mildly to the policy reform. Although flattening the tax schedule for the rich should motivate them to work more hours, the increase in the net income and asset accumulation makes the wealth effect more pronounced. Asset accumulation also promotes capital and output. Individuals experience a welfare gain of around 1.7% and 4.1% measured in  $hev$  and  $cev$ , respectively. The welfare gain is a cumulative result of the (1) increase in consumption which raises welfare, and (2) reduction of informality which reduces the informality cost and has a positive effect on well-being. Since hours of work do not change much the role of this channel for the welfare effects is less relevant. The effects in the partial equilibrium are similar to those in the general equilibrium. Keeping the prices fixed, however, leads to slightly higher welfare gains.

Comparing the flat tax reform with the previous reform of the tax level reduction highlights the fact that flattening the tax schedule has milder effects on the overall economy. Welfare gains of the flat tax reform are smaller but do not lead to fiscal deficits meaning that this reform is sustainable and does not require further interventions into the tax system.

## 7 Conclusions

In this paper, I examine the long-run effects of education and tax policies in the presence of the informal sector in a general equilibrium life-cycle model. I quantify the responses of the Brazilian economy to the change in the education subsidy rates as well as the change in the labor income tax schedule. For the latter, I analyze the change in the income tax level and income tax progressiveness.

I develop a heterogeneous agent overlapping generations life-cycle model in which individuals endogenously choose the level of education, sector of occupation, and hours of work. In the first period of their life, individuals choose the level of education taking into account the costs and benefits of schooling. Each education level requires time and monetary investments. The time cost depends on the innate learning ability and contributes to the disutility associated with the foregone leisure. The monetary cost represents the tuition fee and is partially subsidized by the government. The benefit of acquiring a higher education level is the increase in individual human capital which results in higher productivity and higher wages in the future. The return to human capital is different across sectors of production. In the formal sector, the return to education level is higher than in the informal sector. Additionally, informal employees bear a fixed utility cost of informality associated with the lack of social security and worse working conditions. On the contrary, formal workers face progressive income tax rates.

I calibrate my model to Brazil which features a high level of informality and a low level of educational attainment. This means that there is room for more human capital accumulation and labor formalization. My model matches well the aggregate macroeconomic indicators such as education expenditures and labor income tax revenues to GDP ratios, the share of the informal employment, the distribution of educational attainment and wages across sectors of occupation and levels of education. The model is also capable of replicating a U-shape pattern of the level of informality over education level.

In equilibrium, the share of informality is high for the group of low educated individuals and it reduces for the group of the middle educated.<sup>19</sup> However, the share of informal employees rises again for the highly educated. The model can explain this U-shape pattern of informality along years of education by the interaction between the education premia in both sectors and the progressive taxation. The difference between the return to human capital in both sectors is low for the low-educated. Hence, they opt for the informal sector where there is no taxation. Although the education premium for the highly educated is considerably higher in the formal sector, high labor income tax rates make this sector unattractive for highly productive individuals.

An increase in the education subsidy makes education more affordable, increases the average level of educational attainment, expands formalization, and increases tax revenues. The latter effect is the composition of a higher tax base due to formalization and increase in income resulting from higher human capital accumulation. Hence, the education subsidy is self-financed in the long run and does not require additional tax hikes. Increasing the subsidy level from the current 33% to 90% increases the share of highly educated from 17% to 56% and reduces informality by around 8 pp. Although the additional increase in the subsidy further improves the educational attainment, the share of informal workers

---

<sup>19</sup>Low educated are defined as those who have primary education, middle educated are those who attained a high-school degree, and highly educated are those who have tertiary education.

rises. It illustrates the fact that the increase in education outcomes does not necessarily result in a larger formal sector. Subsidizing 100% of education costs leads to a welfare gain measured in consumption equivalent variation of 94%, meaning that the government would have to double the consumption in the benchmark economy to make individuals indifferent between the two scenarios.

Although the education subsidy is self-financed in the long run, an interesting question is what are the effects of the change in the income tax schedule in an economy with human capital accumulation and occupational choice. In the first step, I vary the level of taxation and observe that decreasing the tax rate favors formalization but has a modest effect on human capital accumulation. Similar to the education policy, reducing the labor income taxation leads to a welfare gain but at a cost of accumulating substantial fiscal deficits. Hence, this policy is not sustainable in the long run.

In the second step, I perform a flat tax reform meaning that all individuals in the formal sector face the same level of labor taxation. Flattening the tax schedule, reduces informality and changes the educational composition of the formal and informal sectors. Highly educated individuals are more likely to opt for a formal sector. As a result, the flat tax reform eliminates the U-shape pattern of the informality along the education dimension. This reform leads to a welfare gain of around 4% measured in cev keeping the labor income tax revenue to GDP ratio fixed.

To conclude, in a developing country with a high level of informality, education policies have positive aggregate effects in the long run. Due to a positive effect on the tax base, there is room for a self-financed education subsidy. This analysis abstracts from the transitional dynamics which may have important implications for the welfare conclusions.

## References

- Araujo, Ricardo Azevedo, Paulo Roberto Amorim Loureiro, and Nathalia Almeida de Souza**, “An Evaluation of the Brazilian Informal Labor Market from 1995 to 2008,” *Journal of Economic Studies*, 2013, 40, 71–87. [1](#), [2](#), [3](#), [3](#)
- Arbex, Marcelo, Antonio F. Galvao, and Fábio Augusto Reis Gomes**, “Heterogeneity in the Returns to Education and Informal Activities,” Insper Working Papers 216, October 2010. [1](#), [2](#), [3](#)
- Auerbach, Alan J. and Laurence J. Kotlikoff**, “Dynamic Fiscal Policy,” *Cambridge University Press, Cambridge*, 1987. [4](#)
- Badaoui, Eliane El and Therese Rebiere**, “Education, Informality, and Efficiency: A Matching Model for a Developing Economy,” *Revue d’Economie Politique*, 05 2013, 123, 423–441. [2](#)
- Bailey, Clive and John Turner**, “Strategies to Reduce Contribution Evasion in Social Security Financing,” *World Development*, 2001, 29 (2), 385 – 393.
- Berniell, Lucila**, “Occupational Choice and Investment in Human Capital in Informal Economies,” *Working Paper, CAF Development Bank Of Latinamerica*, 2020. [2](#), [5](#), [8.1](#)
- Bick, Alexander, Nicola Fuchs-Schündeln, David Lagakos, and Hitosi Tsujiyama**, “Why are Average Hours Worked Lower in Richer Countries?,” *Working Paper*, 2020. [5.1](#)
- Bobba, Matteo, Luca Flabbi, and Santiago Levy**, “Labor Market Search, Informality and Schooling Investments,” *Toulouse School of Economics Working Paper*, January 2018. [2](#)
- Bovenberg, A. Lans and Bas Jacobs**, “Redistribution and Education Subsidies are Siamese Twins,” *Journal of Public Economics*, 2005, 89 (11), 2005 – 2035. [2](#)
- Busato, Francesco and Bruno Chiarini**, “Market and Underground Activities in a Two-sector Dynamic Equilibrium Model,” *Economic Theory*, 2004, 23 (4), 831–861. [4.1](#), [4.2](#)
- Chong, Alberto and Mark Gradstein**, “Inequality and Informality,” *Journal of Public Economics*, 2007, 91 (1-2), 159–179.
- Conesa, Juan C. and Dirk Krueger**, “Social Security Reform with Heterogeneous Agents,” *Review of Economic Dynamics*, October 1999, 2 (4), 757–795. [5.1](#)

- De Fraja, Gianni**, “The Design of Optimal Education Policies,” *The Review of Economic Studies*, 2002, 69 (2), 437–466. [2](#)
- D’Erasmus, Pablo N., Hernan J. Moscoso Boedo, and Asli Senkal**, “Misallocation, Informality, and Human capital: Understanding the Role of Institutions,” *Journal of Economic Dynamics and Control*, 2014, 42 (C), 122–142. [2](#)
- Di Nola, Alessandro, Georgi Kocharkov, Almuth Scholl, and Anna-Mariia Tkhir**, “The Aggregate Consequences of Tax Evasion,” *Review of Economic Dynamics*, forthcoming, 2021. [7](#), [20](#)
- Fairris, David and Erik Jonasson**, “Determinants of Changing Informal Employment in Brazil, 2000–2010,” *MPRA Paper 71475, University Library of Munich, Germany.*, February 2016. [1](#), [2](#), [3](#)
- Fernandez-Bastidas, Rocio**, “Entrepreneurship and Tax Evasion,” *manuscript*, June 2018. [4.1](#)
- Filho, Fernando De Holanda Barbosa**, “An Estimation of the Underground Economy in Brazil,” *IBRE-FGV , FGV Projetos*, 2012. [1](#)
- Filho, Tito Nícias Teixeira Da Silva**, “Estimating Brazilian Potential Output: A Production Function Approach,” Working Paper, Central Bank of Brazil, Research Department, August 2002. [5.3](#)
- Garcia, Italo Lopez**, “Human Capital and Labor Informality in Chile: A Life-Cycle Approach,” *RAND Corporation Working Paper*, February 2015. [1](#), [2](#)
- Gobetti, Sérgio Wulff and Rodrigo Octávio Orair**, “Taxation and Distribution of Income in Brazil: New Evidence from Personal Income Tax Data,” *Brazilian Journal of Political Economy*, 2017, 37, 267–286. [5.1](#), [8.1](#)
- Guvenen, Fatih**, “An Empirical Investigation of Labor Income Processes,” *Review of Economic Dynamics*, 2009, 12 (1), 58–79. [5.1](#)
- , **Burhanettin Kuruscu, and Serdar Ozkan**, “Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis,” *The Review of Economic Studies*, 2014, 81 (2), 818–850. [2](#)
- Haanwinckel, Daniel and Rodrigo R. Soares**, “Fighting Employment Informality with Schooling,” *IZA World of Labor*, 10 2017, (394). [2](#)
- Heathcote, Jonathan, Fabrizio Perri, and Giovanni Violante**, “Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967–2006,” *Review of Economic Dynamics*, 2010, 13 (1), 15–51. [8.1](#)

- , **Kjetil Storesletten**, and **Giovanni L. Violante**, “Optimal Tax Progressivity: An Analytical Framework,” *The Quarterly Journal of Economics*, 06 2017, 132 (4), 1693–1754. [5.1](#), [5.1](#), [8.1](#)
- Huggett, Mark**, “Wealth Distribution in Life-Cycle Economies,” *Journal of Monetary Economics*, 1996, 38 (3), 469–494. [4](#)
- Ihrig, Jane and Karine S. Moe**, “Lurking in the Shadows: the Informal Sector and Government Policy,” *Journal of Development Economics*, 2004, 73 (2), 541–557. [4.2](#)
- Jung, Juergen and Chung Tran**, “The Extension of Social Security Coverage in Developing Countries,” *Journal of Development Economics*, 2012, 99 (2), 439 – 458. [5.1](#)
- Krueger, Dirk and Alexander Ludwig**, “Optimal Progressive Labor Income Taxation and Education Subsidies when Education Decisions and Intergenerational Transfers are Endogenous,” *American Economic Review*, 2013, 103 (3), 496–501. [2](#), [5.1](#)
- Maloney, William F.**, “Informality Revisited,” *World Development*, 2004, 32 (7), 1159–1178. [2](#)
- Meghir, Costas, Renata Narita, and Jean-Marc Robin**, “Wages and Informality in Developing Countries,” *American Economic Review*, 2015, 105 (4), 1509–46. [1](#), [2](#), [3](#)
- OECD**, “Pensions at a Glance 2015,” Organization for Economic Co-operation and Development 2016. [5.3](#)
- , “Education at a Glance 2018,” Organization for Economic Co-operation and Development 2018. [3](#)
- , “Education at a Glance 2019,” Organization for Economic Co-operation and Development 2019. [3](#)
- Porta, Rafael La and Andrei Shleifer**, “Informality and Development,” *Journal of Economic Perspectives*, 2014, 28 (3), 109–26. [4.2](#)
- Schneider, Friedrich, Andreas Buehn, and Claudio Montenegro**, “New Estimates for the Shadow Economies all over the World,” *International Economic Journal*, 2010, 24 (4), 443–461. [1](#), [3](#)
- Tauchen, George and Robert Hussey**, “Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models,” *Econometrica*, March 1991, 59 (2), 371–396. [5.1](#)
- Ulyssea, Gabriel**, “Firms, Informality, and Development: Theory and Evidence from Brazil,” *American Economic Review*, 2018, 108 (8), 2015–47. [5](#), [8.1](#)

**World Bank**, “Summary Note on Pension Reform in Brazil,” World Bank Other Operational Studies 26388, The World Bank 2017. [10](#)

## 8 Appendix

### 8.1 Data

#### Data Sample and Main Empirical Findings

My main source of data for this study is the Brazilian National Household Sampling Survey (PNAD) for a period from 2010 to 2015. PNAD is a survey conducted annually by the Brazilian National Statistics Bureau (IBGE) and investigates a wide range of population characteristics such as education, labor and income. PNAD is a cross-sectional data which I pool for calculating main calibration targets regarding educational attainment, occupation and income. For estimating a stochastic ability process, I rely on the labor force survey of Brazil, Pesquisa Mensal de Emprego (PME). The advantage of the latter is the panel dimension, however, it has a smaller coverage in terms of survey questions. Both data sets are conducted by the IBGE and are complementary to each other.

For the purpose of this paper, I focus on male individuals of age 25-65 who are household heads. Following the literature, I define an individual to be working informally if she is an employee but did not sign a working contract (working card) with the exception for public employees who do not have to sign it (Ulyssea 2018). I do not consider individuals who report to be unpaid workers, unemployed or self-employed. It is evident that a substantial share of the self-employed in Brazil operate informally. However, the effect of education on the entrepreneurial ability may be different from the effect of education on the productivity of employees (see, e.g., Berniell (2020)). Hence, I abstract from the self-employed in this study.<sup>20</sup> Another possible concern is that individuals may operate formally and informally simultaneously making it difficult to categorize an individual in either of a group. In my sample, less than three percent of individuals report to have more than one job making this concern negligible.

Table 13 shows that there are around 83% of formal and 17% of informal employees in the sample. Around 32% of total employees do not have a high-school degree, meaning that they have acquired only compulsory primary education. Around 52% have a high-school degree and around 16% have acquired tertiary education.

Table 13: Level of Informal Employment and Education Level

Formal, %	82.40
Informal, %	17.60
<hr/>	
Low ed., %	32.22
Mid. ed., %	52.13
High ed., %	15.65

---

<sup>20</sup>See Di Nola et al. (2021) analyzing tax evasion by self-employed.



Table 14 shows the results for a Probit regression for the pooled sample of formal and informal employees. The dependent variable is informal occupation which takes the value of one if an individual operates informally and zero else-wise. Controlling for age, race and region shows that the education level measured in years of education has a statistically significant negative effect on the probability of operating informally. Note that the value of the coefficient in the Probit regression does not have an economic interpretation.

Table 14: Probit

	(1)
	Informality
yedu	-0.0592*** (-51.18)
age	-0.0658*** (-17.45)
age2	0.000771*** (17.48)
race	0.00456** (2.71)
region	-0.0215*** (-29.76)
_cons	1.276*** (15.95)
<i>N</i>	107981

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 15 presents empirical evidence on the education premium for both formal and informal employees. The dependent variable is a logarithm of the gross hourly wage. I split the sample into formal and informal sector employees and control for age, race and region. The coefficients for education are positive and statistically significant for formal as well as informal workers. However, the education premium is larger in the formal sector.

### Estimating Labor Income Process

The estimation of the stochastic component of labor productivity follows closely the procedure described by Heathcote et al. (2010). First, I regress the logarithm of hourly gross labor earnings on observable individual characteristics such as education, race and age irrespective of individual's occupation.

$$\ln inc_{i,t} = \alpha_0 + \beta_0 educ_{i,t} + \beta_1 race_{i,t} + \beta_2 age_{i,t} + a_3 age_{i,t}^2 + \eta_{it}, \quad (15)$$

Table 15: Education Wage Premia

	(1)	(2)
	logw, Formal	logw, Informal
yedu	0.0750*** (157.41)	0.0694*** (62.20)
age	0.0370*** (24.01)	0.0407*** (11.06)
age2	-0.000306*** (-16.81)	-0.000376*** (-8.81)
race	-0.0228*** (-34.94)	-0.0234*** (-13.67)
region	0.0131*** (43.95)	0.0172*** (24.41)
_cons	5.431*** (166.89)	5.190*** (66.33)
$N$	88987	18988

$t$  statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

where  $i$  is individual and  $t$  is time index respectively. Then, I model the residual  $\eta$  as a first-order auto-regressive process:

$$\log \eta_{i,t+1} = \rho_\eta \log \eta_{i,t} + \zeta_{i,t+1}, \quad (16)$$

where  $\zeta_{i,t+1} \sim N(0, \sigma_\eta^2)$ . I estimate this process and obtain a persistence parameter  $\rho_{\eta,a} = 0.776$  and the dispersion parameter  $\sigma_{\eta,a} = 0.354$  on the annual basis. Since the model period is defined to be five years, I translate the obtained parameters into 5-year equivalents:

$$\rho_\eta = \rho_{\eta,a}^T, \quad (17)$$

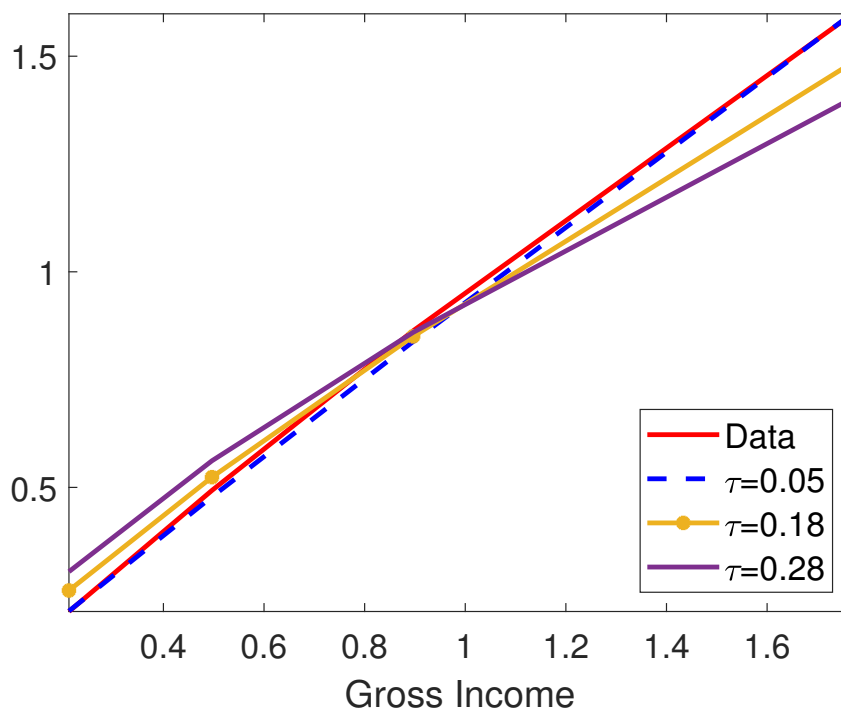
$$\sigma_\eta^2 = \sigma_{\eta,a}^2 \sum_{t=0}^{T-1} \rho_{\eta,a}^{2t}, \quad (18)$$

where  $T = 5$  is the number of years in the period. Consequently, the  $\rho_\eta = 0.282$  and  $\sigma_\eta = 0.539$ .

### Estimating Tax Progressivity

For the calibration of the progressive income taxation, I follow [Heathcote et al. \(2017\)](#) and consider a functional form:

Figure 7: Gross versus Net Income



$$\tau_l(y) = y - \lambda_{hsv} y^{1-\tau_{hsv}},$$

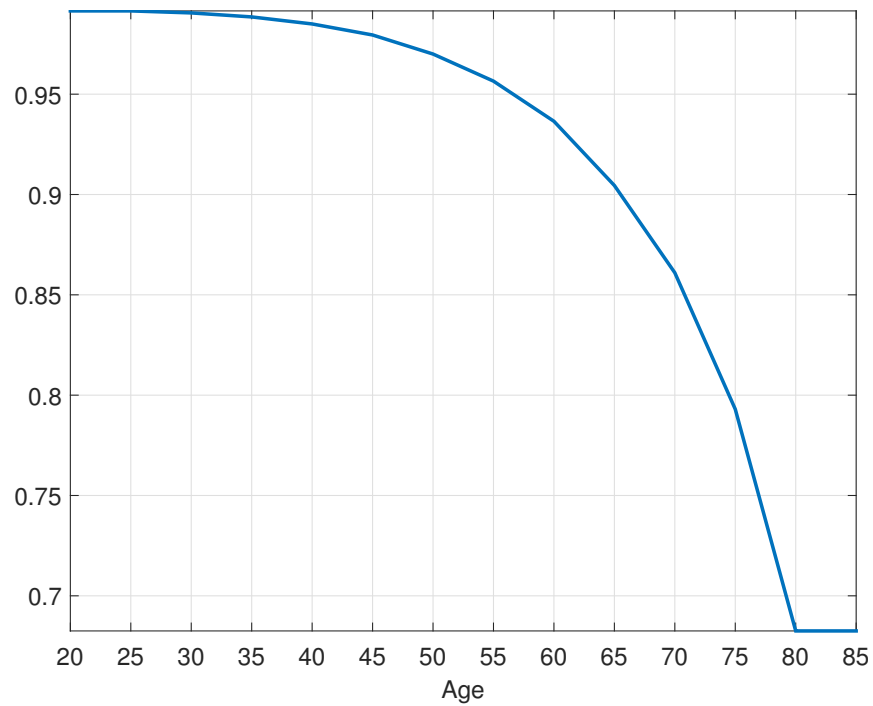
where  $\tau_l(y)$  is the tax rate at the income level  $y$ ,  $\tau_{hsv} > 0$  is a measure of progressivity of the tax schedule and  $\lambda_{hsv}$  is a parameter that governs the average tax rate. I use the data from [Gobetti and Orair \(2017\)](#) who report average tax rates by income brackets in Brazil to pin down the shape of the tax function. The procedure goes as follows: (1) given a gross income bracket and a corresponding tax rate, I calculate net income for each bracket in the data; (2) using a specified functional form, I calculate net income in the model:  $\tilde{y}^F = \lambda_{hsv} y_F^{1-\tau_{hsv}}$ . (3) I pin down a parameter  $\tau_{hsv}$  to minimize the distance between net to gross income ratio along income brackets in the model and in the data.

## 8.2 Computational Algorithm

In this subsection, I outline the computational algorithm employed for this project.

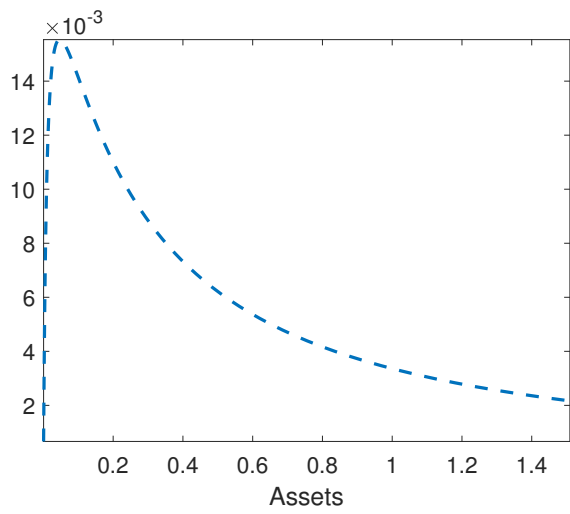
1. Given guesses for capital and labor input as well as tax rates, compute factor and consumer prices.
2. Given prices and public transfers, determine household policy functions by backward iteration.
3. Compute the distribution of households over the state space.

Figure 8: Surviving Probabilities

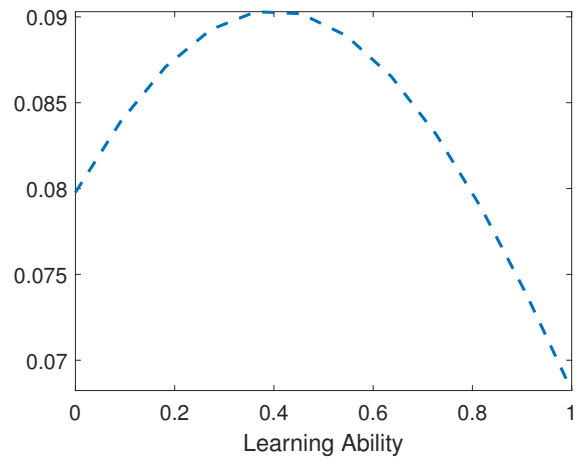


4. Aggregate household decisions.
5. Calculate the absolute value of the relative difference between demand and supply for capital and labor. If the difference is small enough the equilibrium is found and we can stop the iteration procedure. If not, start again at point 1.

### 8.3 Additional Figures and Tables



(a) Initial Asset Distribution



(b) Initial Learning Ability Distribution

Figure 10: Informality over Age

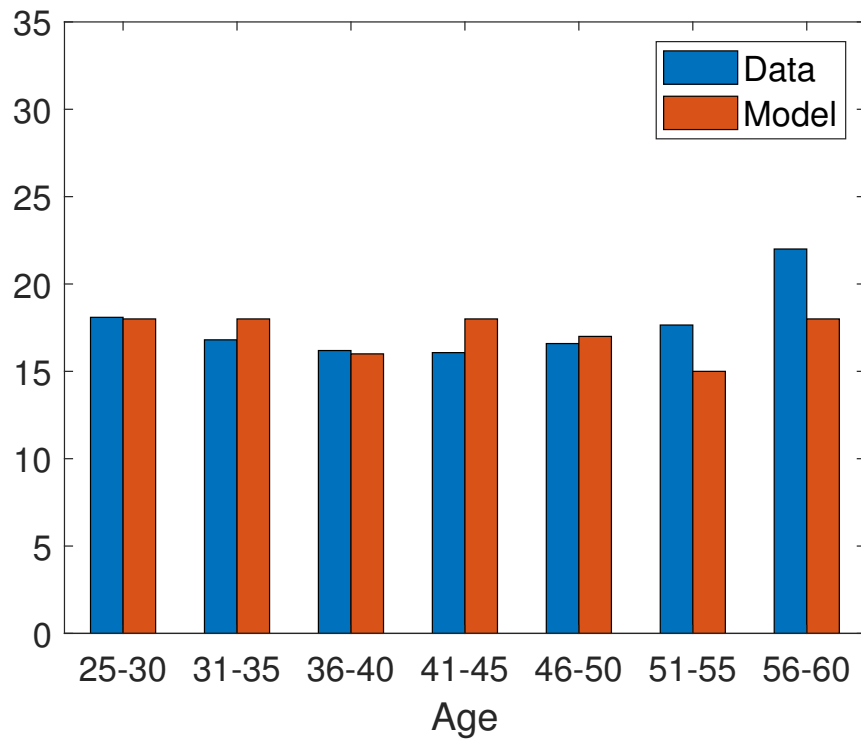
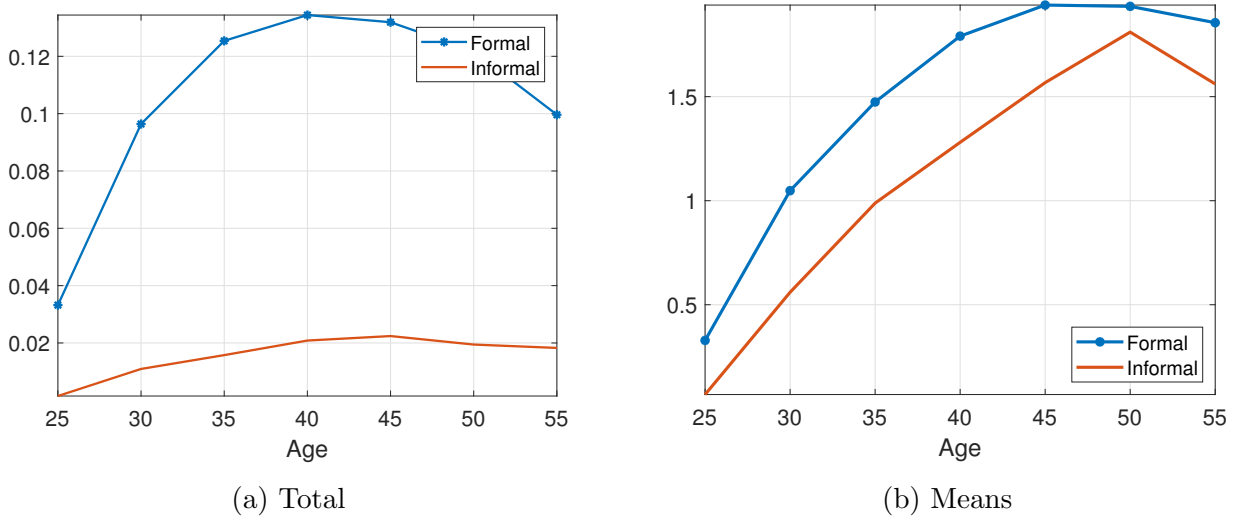
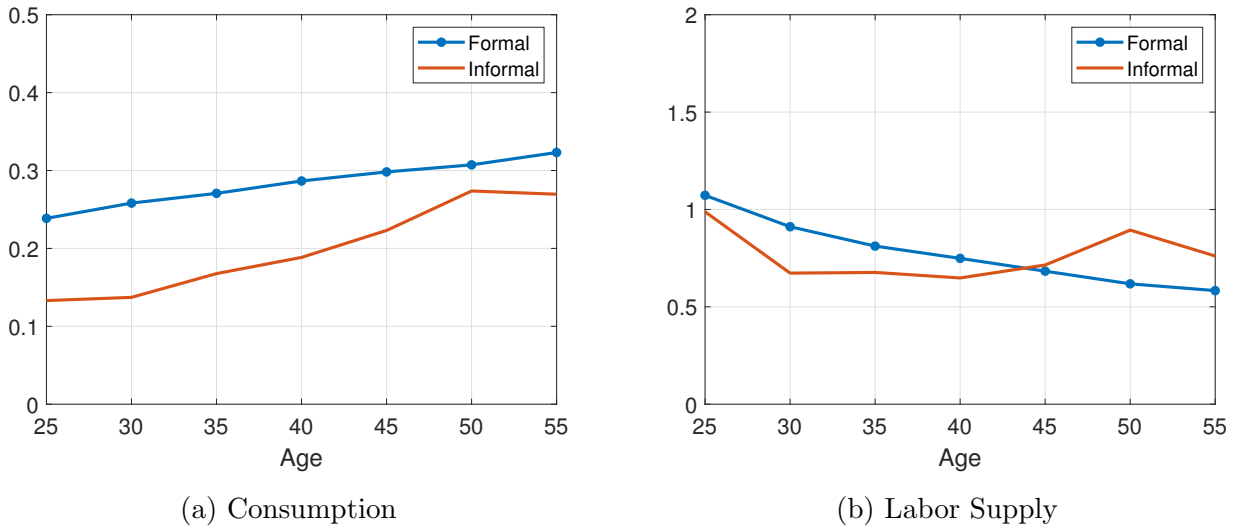


Figure 11: Asset Holdings Comparison



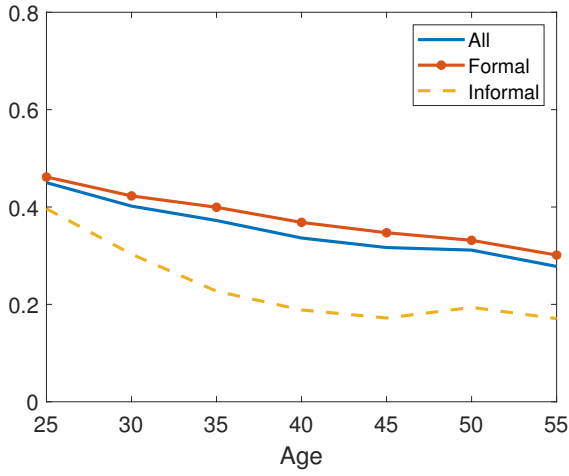
*Notes:* The left panel compares total asset holdings of formal and informal workers over the life-cycle whereas the right panel compares the respective means, measured in model units.

Figure 12: Consumption and Labor Supply Comparison

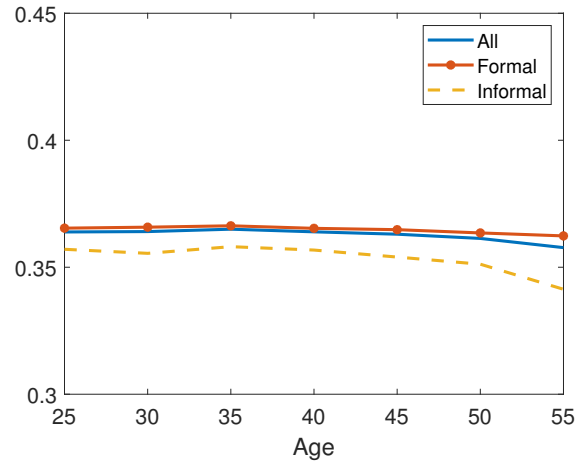


*Notes:* The left panel compares the mean consumption whereas the right panel compares the mean labor supply in efficiency units of formal and informal workers over the life-cycle.

Figure 13: Hours of Work over the Life-cycle

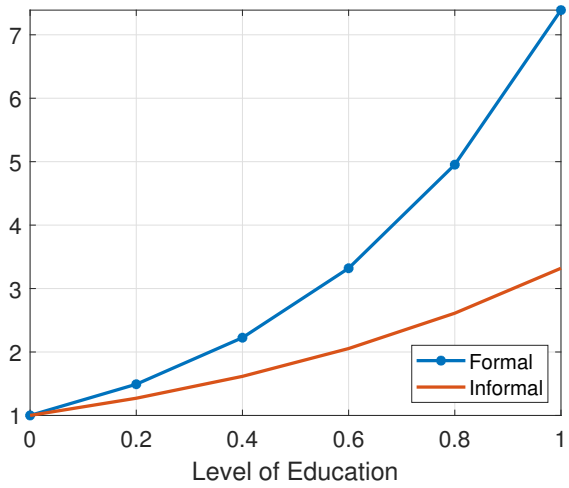


(a) Hours of Work, Model

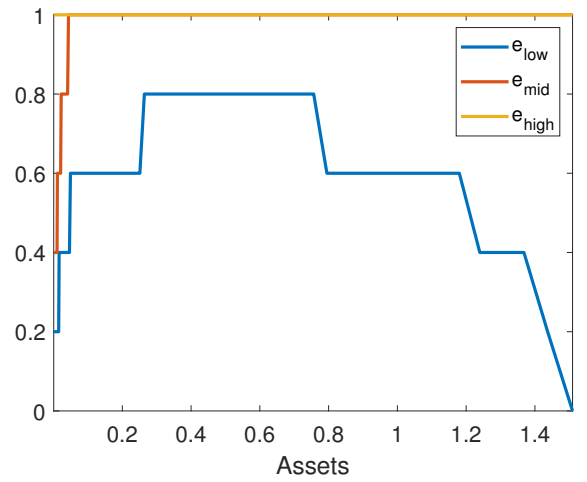


(b) Hours of Work, Data

Figure 14: Education Premia and Educational Policy for subsidy = 1



(a) Theoretical Education Premium



(b) Education Policy, subsidy = 1

*Notes:* The left panel of this figure depicts a theoretical education premium for the formal and the informal sector. Specifically, it shows  $\exp(\theta^F h)$  and  $\exp(\theta^I h)$ . The right panel shows educational choice as a function of assets and innate learning ability.

Table 16: Education Subsidy Change, Partial Equilibrium

Subsidy $s$	0.0	0.1	0.2	0.3	0.33	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Share of inf., %	21.05	20.64	18.45	16.96	17.02	16.34	13.65	12.77	9.80	9.13	8.71	8.73
Low educated, %	52.15	49.74	38.10	30.72	30.23	28.48	15.23	12.71	4.42	1.38	0.48	0.41
Mid. educated, %	36.59	37.74	46.97	52.41	52.20	50.39	60.51	58.48	59.69	54.58	43.54	30.58
Highly educated, %	11.26	12.52	14.93	16.88	17.56	21.13	24.27	28.81	35.89	44.04	55.97	69.01
Interest rate, %	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2
Wage ratio	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74	1.74
GDP	0.33	0.33	0.35	0.36	0.37	0.37	0.39	0.41	0.43	0.46	0.48	0.51
Hours work, %	38.10	37.77	37.25	36.61	36.36	35.88	35.14	34.27	33.50	32.19	30.50	28.39
Capital	0.20	0.20	0.21	0.22	0.22	0.23	0.24	0.25	0.26	0.28	0.29	0.31
Welfare, hev	-0.06	-0.05	-0.03	-0.01	0.00	0.01	0.06	0.09	0.14	0.20	0.25	0.33
Welfare, cev	-0.14	-0.12	-0.07	-0.01	0.00	0.03	0.14	0.22	0.38	0.53	0.72	0.99

*Notes:* This table presents the education policy effects in partial equilibrium such that the prices of labor and capital are fixed to their benchmark values. A gray shaded column refers to the benchmark economy with the subsidy level of 33%.

Table 17: Income Tax Level Change, Partial Equilibrium

lambda	0.60	0.65	0.70	0.75	0.80	0.85	0.90
Share of inf., %	31.28	26.84	23.71	17.02	13.78	12.23	11.34
Low educated, %	30.67	30.44	30.29	30.23	30.21	30.18	30.18
Mid. educated, %	49.89	50.41	51.58	52.20	52.74	53.72	54.11
Highly educated, %	19.44	19.15	18.13	17.56	17.04	16.10	15.71
Interest rate, %	5.14	5.14	5.14	5.20	5.14	5.14	5.14
Wage ratio	1.74	1.74	1.74	1.74	1.74	1.74	1.74
GDP	0.34	0.34	0.35	0.37	0.37	0.38	0.39
Hours work, %	35.97	35.50	35.34	36.36	36.90	37.30	37.60
Capital	0.19	0.20	0.21	0.22	0.23	0.25	0.26
Welfare, hev	-0.06	-0.04	-0.02	0.00	0.02	0.05	0.07
Welfare, cev	-0.14	-0.10	-0.06	0.00	0.05	0.11	0.17

*Notes:* This table shows the results of the tax level change policy in partial equilibrium. Labor and capital prices are fixed to the benchmark values. Parameter  $\lambda_{hsv}$  defines the level of income taxation in the economy. The higher the  $\lambda_{hsv}$ , the lower is the income tax level. A gray shaded column refers to the benchmark economy with  $\lambda_{hsv} = 0.75$ . All other tax rates remain unchanged.