Misallocation of talent, human capital formation, and development in Brazil*

Fernando Barros Jr† FEARP/USP
Marcos J Ribeiro‡ FEARP/USP
Bruno Delalibera§ BTG Pactual
Luciano Nakabashi¶ FEARP/USP

August 30, 2021

Abstract

In this paper, we elaborate a General Equilibrium Model to study why the relative wage of teachers is higher in low and middle income regions in Brazil. We define a Roy model where there is an externality in the occupational occupational choice because teachers' human capital is important to the formation of human capital of all workers. Then, when people with more idiosyncratic ability choose to become a teacher, all the workforce has benefits. However, in our model individuals' occupational decisions are distorted by barriers in the labor and educational goods markets, which in turn cause a misallocation of talent and harm economic growth. After calibrating the model using the Method of Moments and data from Brazil, we find that distortions are an important factor to explain differences in relative wages. Moreover, we performed a series of counterfactual exercises that show that reducing distortions in the labor and educational market would increase Brazilian GDP by a value between 9.19% and 43.49%.

Keywords: Misallocation; Frictions; Human Capital; Externalities; Development;

*We thank Fábio Gomes and Luiz Brotherhood for their comments. This research was carried out with the support of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Financing Code 001.

†Correspondence address: Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, Avenida Bandeirantes, 3900 - Vila Monte Alegre, Ribeirão Preto - SP, 14040-905. E-mail: fabarrosjr@usp.br.

‡Email: mjribeiro@usp.br.

§Email: brunodelalibera@gmail.com.

¶Email: luciano.nakabashi@gmail.com.

1 Introduction

Much effort has been made to try to find the causes and barriers to economic growth, and one of the approaches that has stood out is that of misallocation. This approach is used to measure the distortions of the economy, to understand the causes of misallocation of resources and to assess losses in TPF and output. Several papers on misallocation emphasize that there may be losses in productivity and in the aggregate output in an economy that presents misallocation of resources. Misallocation of capital, credit, and talent has been pointed out as possible barriers to growth (Banerjee and Duflo, 2005; Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Hsieh et al., 2019). Our main subject of study in this paper is the misallocation of talent in an economy where individuals face an occupational choice problem distorted due to market frictions.

Misallocation of talent across occupations, and sectors may be a consequence of race and gender discrimination, social norms and culture and barriers in human capital formation (Restuccia and Rogerson, 2017). As can be seen in Hnatkovska et al. (2012) the misallocation of talent in India comes from the caste system. And this system causes a social and economic gap between the privileged and non-privileged castes. On the other hand, Café (2018) shows that there is overqualification of workers in the public sector in relation to the private sector in Brazil. This is due to the fact that public tenders cause misallocation of talent, especially those in which the evaluation is not related to the work that will be performed.

Barros and Delalilbera (2018) identified an inverse relationship between the relative wage of teachers and the level of development of Brazilian states. They point out that the occupational choice of workers with multiple skills is driven by labor market incentives (net wage) and the costs of investing in specific human capital (education acquisition). Furthermore, in the model they develop, the human capital of teachers is a source of positive externalities for the human capital of all workers. Therefore, in regions where there is a higher cost to become a teacher and the relative wage is lower, more talented people can choose other occupations. Which, in turn, can cause misallocation of talent and harm the formation of human capital in other occupations.

In US economy Hsieh et al. (2019) identify that in 1960, 94% of doctors and lawyers were white men and in 2010, the fraction was 62%. So they built a general equilibrium model to study the effects of these changes on American economic development. They suppose that distribution of talent for each occupation is identical for whites, blacks, men and women, and introduce two main elements that will cause individuals to choose occupations where they do not comparative advantage: discrimination in the labor market and barriers to forming human capital. The main finding is that between 20% and 40% of growth in market GDP per person over the last five decades can be traced to declining occupational barriers. This decline in occupational barriers caused women and blacks to occupy highly qualified positions over time. And when the worker has the specific skill for the occupation, this reallocation represents a better allocation of talent, which in turn had a positive effect on GDP.

Based on Hsieh et al. (2019) and mainly on Barros and Delalilbera (2018) we build a general equilibrium model to discuss the misallocation of the labor force's abilities across the Brazilian States in the context of occupational choice. Our work differs from Barros and Delalilbera (2018) in two senses. First, in addition to taking to account that teachers' human capital is a source of positive externalities for the entire workforce, we explicitly model the importance of the number of teachers in the formation of human capital of the workers. Second, we use our model to show that a decrease in market frictions can explain the convergent of income across Brazilian states.

We introduce two distortions that influences individuals' occupational choices into the model. Those distortions are a source of misallocation of talent and are specified at the occupation and regional levels. First we consider a friction in the labor market, which can be interpreted as the relative difficulty of finding a job in a given occupation and region. Also, this distortion can be a result of social status or some form of discrimination. The second distortion appears in the educational market, which in turn affects human capital formation. This distortion can be understood as the cost of human capital formation in a given region and occupation.

The importance of this is clear when we take into account the fact that high distortions in a specific occupation reduce the proportion of workers in that occupation. Therefore, the number of talented people with high human capital choosing this occupation could decrease. And in the specific case of teachers, who are essential inputs for the formation of human capital, this would harm the entire workforce and development.

In Hsieh et al. (2019) it is clear the importance of human capital formation for economic growth. Since the fall in human capital barriers alone would have accounted for 36% of market growth in GDP per person. Human capital is crucial for economic growth and development by facilitating innovation and diffusion of technology as in Romer (1990), Mankiw et al. (1992), Borensztein et al. (1998) and Benhabib and Spiegel (2005). We contribute to this literature by showing how regional disparities in labor and educational markets can distort the formation of human capital generating a misallocation of talent.

The more recent literature has emphasized the relevance of human capital quality in economic growth. Hanushek and Woessmann (2012) argue that the Latin American countries lagged behind over the period 1960-2000 because of their students' poor performance in educational achievement. Latin American countries have been classified as the poorest in score tests across several developed and developing countries, which have been critical to their poor economic performance (Hanushek and Woessmann, 2012).

Going further, many studies point to the relevance of teachers in the students' learning process (Woessmann, 2016; Barros and Delalilbera, 2018; Hanushek et al., 2019). Relying on three different strategies and controlling for several variables that are related to students performance as family background, school inputs, institutional features of the school systems, and cross-country differences in educational inputs, Hanushek et al. (2019) find a robust and positive relationship between teacher cognitive skills and student performance measure by the Programme for International Student Assessment (PISA) scores. The cognitive skills of teachers are even more important to students' performance than the cognitive skills of their parents (Hanushek et al., 2019). We rely on those findings and explicitly model the importance of teacher's quality in the formation of human capital of all workers. Thus, occupational choice generates an externatily: if gifted people decide to become teachers, all workers benefit from that decision. On the other hand, if low-skilled people choose the teacher occupation, then the formation of human capital of the whole workforce can be compromised.

Using the student's mathematics test score of the PISA, Woessmann (2016), through an educational production function, points to the relevance of teachers' quality measured by the relative wage and education of teachers on students' performance. Woessmann (2016) argues that higher teacher's wages have a positive influence on recruiting higher ability individuals into teaching. For Brazil, Menezes-Filho and Pazello (2007) find evidence that the proficiency of public school students is positively impacted by the relative wages of teachers.

Machado and Scorzafave (2016) draws attention to the fact that wages may affect the decision of the most talented individuals to become teachers. In addition, since the individual becomes teacher, the wages can affect the effort that will be exercised in the classroom and encourage or discourage the turnover of these professionals in the schools. Several other studies also point to the fact that the ability of teachers is related to the salary offered to them, for example Stoddard (2003), Lakdawalla (2006) and Bacolod (2007). In our model, friction in the labor and educational markets induces the occupational choice of individual. In that sense, we show that distortions related to the teachers occupation have a higher impact on the economy due to the externality generated by teachers human capital and the amount of teachers to the formation of human capital of all workers.

To estimate these distortions we calibrated the model using the Method of Moments and 2015 Brazilian states data. We then ran a series of counterfactual exercises to show how sensitive the economy is to changes in teacher occupation distortions. Furthermore, we calibrated the model with 2003 data and discussed how the reduction of distortions over time were responsible for the absolute convergence of income that we observed.

Our results show that reducing distortions in the educational market has a greater positive impact on growth than reducing distortions in the labor market. Furthermore, we showed that there is a misallocation of talent in the Brazilian economy, and that allocating highly qualified people into the teaching occupation has a potential to generate a significant increase in the Brazilian GDP (between 9.19% and 43.49%).

To accomplish our goals, this article is divided into four more sections. In section 2 we present our general equilibrium model. In section 3 we explain how this model was calibrated using data from the Brazilian economy from 2015. The results of the calibration, some stylized facts, counterfactual exercises, and robustness check, are presented in section 4, and finally, the article is completed in section 5.

2 Model

We build a Roy model based on Hsieh et al. (2019) and Barros and Delalilbera (2018) to study how frictions in labor and educational markets affect the formation of human capital in an economy where teachers have a central role. Therefore, in this section we will discuss the behavior of firms and workers, within the model, its main implications, and finally define the competitive equilibrium.

2.1 Firms

We consider a country divided into $R \in \mathbb{N}$ independent regions (states). In each region, there is a continuum of workers that choose one of the $N \in \mathbb{N}$ occupations in the economy. A person born in region r can not work in a different region.[1] There is a large number of homogeneous competitive firms, so firms take prices as given. Firms hire workers in all regions and occupations to produce a single good. The production function used by the firms is given by

$$Y = \sum_{r=1}^{R} \sum_{i=1}^{N} A_r H_{ir}, \tag{1}$$

where Y is output, A_r is Total Factor Productivity (TFP) of region r and H_{ir} is the aggregate human capital of people working in occupation i at region r. Output can be consumed or used as an educational good, as shall become clear when we present the worker's problem. Thus, the firms problem is choosing labor in terms of efficient units (aggregate human capital) to maximize profit, taking wages (w_{ir}) of each occupation in each region as given.

$$\max_{H_{ir} \geq 0} \left[\sum_{r=1}^{R} \sum_{i=1}^{N} A_r H_{ir} - \sum_{r=1}^{R} \sum_{i=1}^{N} w_{ir} H_{ir} \right] \tag{2}$$

The solution of the problem described above is simple. The demand for human capital is given by:

$$H_{ir}^{d} = \begin{cases} 0 & \text{if } A_r < w_{ir} \\ x \in \mathbb{R}_+ & \text{if } A_r = w_{ir} \\ \infty & \text{if } A_r > w_{ir} \end{cases} . \tag{3}$$

[1] In Appendix D we discuss about migration.

2.2 Workers

Workers have idiosyncratic abilities for each occupation. In a world with multiple occupations, some workers can have a high talent for many occupations, some for only one, while others may lack the skills for any occupation in the economy. Individuals value consumption and leisure, which we model as the time not spent in school. Each worker is endowed with one unit of time to study or to consume as leisure. The utility of a person is given by

$$U(c, s) = c^{\beta}(1 - s) \tag{4}$$

where c represents consumption, s is time spent at school, and β is a parameter giving the importance of consumption in relation to leisure.

We follow Hsieh et al. (2019) and introduce two distortions in our model. People working in occupation i at region r is paid a net wage of $(1 - \tau_{ir}^{w})w_{ir}$ where τ_{ir}^{w} is a distortion specific for occupation i and location r. One can interpret τ_{ir}^{w} as an unobserved cost (or benefit) of occupation i at region r. For example, it can represent a social status or a difficult in finding a job in a given occupation and region.

Educational choices are also distorted due to a 'tax' on educational goods, e. For each good invested in education, a person pays τ_{ir}^{h} as a 'tax'. We can think of this distortion as representing forces that affect the cost of acquiring human capital in different occupations and regions. For example, τ_{ir}^{h} might indicate the difficulty of finding training to work in a specific occupation, or it can represent investments to develop math and science skills required by certain occupations.

Restuccia and Rogerson (2013) emphasize that distortions in this context create misallocation in the economy and can be thought of as generators of wedges in the first order conditions of workers' optimization problems. That is, wedges distance the optimal choices in the distorted economy from the original optimal choices creating misallocation. Based on Banerjee and Moll (2010) misallocation can be defined as:

Definition 1 (Misallocation of resources)**.** *There is a misallocation of resources in an economy if it is possible to redistribute factors $f1$ and $f2$ from one agent to another, and, with this, increase the aggregate product of that economy.*

As in Barros and Delalilbera (2018), we assume that human capital formation of workers depends on teachers' quality. That is, the quality of teachers is a source of positive externalities in the formation of human capital. As a contribution of the present paper, we incorporate the quantity of teachers as another element to workers' human capital formation. Therefore, the workers' human capital in each region is given by:

$$h_{ir}(e,s) = T_r^{\varphi} s_i^{\phi_i} e_{ir}^{\eta} \tag{5}$$

where e represents consumption of educational goods, s is time spent in school, η is elasticity of human capital function with respect consumption of educational goods, and $\phi_i > 0$ is the elasticity of human capital concerning time in school. Notice that this parameter varies among occupations and generates differences in schooling. Finally, T_r represents the role of teachers in the workers' human capital formation. We set $T_r = p_{tr}^{\alpha} H_{tr}^{1-\alpha}$ where $\alpha \in (0,1)$, p_{tr} is the fraction of people working as teachers, and H_{tr} is the teachers' aggregate human capital. We use this functional form to incorporate the quality and quantity of teachers on the human capital formation of workers.[2]

Following McFadden (1974), Eaton and Kortum (2002), and Hsieh et al. (2019), abilities dispersion is modeled as a multivariate Fréchet distribution. Let ϵ_i be the ability of an individual in occupation i, then the distribution of abilities across occupations is:

$$F(\epsilon_1, \ldots, \epsilon_N) = \exp\left[-\sum_{i=1}^{N} \epsilon_i^{-\theta}\right] \tag{6}$$

where θ governs the skill dispersion.

The worker's problem can be split into two steps. First, given the occupational choice i, for which the individual has an idiosyncratic ability ϵ_i, and taking wage w_{ir} as given, each worker chooses c, e, and s to solve the following problem:

$$\max_{c,s,e} c^{\beta}(1-s) \tag{7}$$

$$\text{s.t. } c = (1-\tau_{ir}^{w}) w_{ir} h_{ir}(e_{ir}, s_i)\epsilon_i - (1+\tau_{ir}^{h}) e_{ir}$$

Solving the problem above, we find the amount of time and goods spent on human capital

[2]See Krueger (2003) and Lakdawalla (2006) for a discussion on teachers' quality and quantity.

accumulation:[3]

$$
\begin{aligned}
s_i^* &= \left(1 + \frac{1-\eta}{\beta\phi_i}\right)^{-1} && (8)\\
e_{ir}^*(\epsilon) &= \left[\eta\left(\frac{1-\tau_{ir}^w}{1+\tau_{ir}^h} w_{ir}\right)(p_{tr}^\alpha H_{tr}^{1-\alpha})^\varphi \left(1+\frac{1-\eta}{\beta\phi_i}\right)^{-\phi_i}\epsilon_i\right]^\kappa && (9)
\end{aligned}
$$

where $\kappa = 1/(1-\eta)$.

For a given occupation, a higher elasticity of human capital concerning time leads to more time allocated to human capital accumulation. Individuals in high ϕ_i occupations acquire more schooling and have higher wages as compensation. Note that the wage and distortions do not affect schooling because they have the same effect on the return and cost of time.[4] However, they can change the returns of investments in human capital goods relative to the costs, with an elasticity that is increasing in η.

Second, we substitute the expressions in equations (8) and (9) and the budget constraint into the utility function to get the following expression for the indirect utility function of occupation i:

$$
D_{ir} = \left[\bar{\eta}\left(\frac{1-\tau_{ir}^w}{(1+\tau_{ir}^h)^\eta} w_{ir}\right)(p_{tr}^\alpha H_{tr}^{1-\alpha})^\varphi s_i^{\phi_i}(1-s_i)^{\frac{1}{\beta\kappa}}\epsilon_i\right]^{\beta\kappa} \tag{10}
$$

where $\bar{\eta} = \eta^\eta(1-\eta)^{1-\eta}$.

Therefore, the occupational choice problem reduces to picking the occupation that delivers the highest value of D_{ir}. Since talent is drawn from an extreme value distribution, the highest utility can also be characterized by an extreme value distribution (McFadden, 1974). Proposition 1 states that the overall occupational share can be obtained by aggregating the individual optimal choice.

Proposition 1 (Occupational choice)**.** *Aggregating across workers, the solution of individual's occupational choice problem is:*

$$
p_{ir} = \frac{\tilde{w}_{ir}^\theta}{\sum_{j=1}^N \tilde{w}_{jr}^\theta} \tag{11}
$$

[3]The complete solution of model can be view in Online Appendix.

[4]Barros and Delalilbera (2018) have evidence that average schooling is similar across Brazilian states giving an occupation. In accord to the authors, the amplitude (max - min) of the mean of years of schooling across states is at most 2.3 years, except for agricultural worker, where this statistic amounts to 3.3 years.

where p_{ir} is the fraction of workers in occupation i in region r, and:

$$\tilde{w}_{ir} = \bar{\eta}\left(\frac{1-\tau^w_{ir}}{(1+\tau^h_{ir})^\eta}w_{ir}\right)(p^\alpha_{tr}H^{1-\alpha}_{tr})^\varphi s_i^{\phi_i}(1-s_i)^{\frac{1}{\beta\kappa}}$$

Proof. Let:

$$\tilde{w}_{ir} = \bar{\eta}\left(\frac{1-\tau^w_{ir}}{(1+\tau^h_{ir})^\eta}w_{ir}\right)(p^\alpha_{tr}H^{1-\alpha}_{tr})^\varphi s_i^{\phi_i}(1-s_i)^{\frac{1}{\beta\kappa}}$$

Then, we can rewrite equation (10) as:

$$D_{ir} = [\tilde{w}_{ir}\epsilon_i]^{\beta\kappa}$$

Therefore, the problem solution of individual i living in region r involves picking the occupation with the highest value of $\tilde{w}_{ir}\epsilon_i$. Without loss of generality, consider the probability of an individual choosing occupation 1:

$$\begin{aligned}
p_{ir} &= Pr(\tilde{w}_{1r}\epsilon_1 > \tilde{w}_{ir}\epsilon_i) \quad \forall i \neq 1 \\
&= Pr\left(\epsilon_i < \frac{\tilde{w}_{1r}}{\tilde{w}_{ir}}\epsilon_1\right) \quad \forall i \neq 1 \\
&= \int F_1(\alpha_1\epsilon, \alpha_2\epsilon, ..., \alpha_N\epsilon)d\epsilon \qquad (12)
\end{aligned}$$

where F_1 represents the derivative of equation (6) with respect to its first argument and $\alpha_i = \tilde{w}_{1r}/\tilde{w}_{ir}$ for $i \in \{1, 2, ...N\}$. Taking the derivative of equation (6) with respect to ϵ_1, and evaluating in ϵ:

$$\begin{aligned}
F_1 &= \theta\epsilon_1^{-\theta-1}\exp\left(-\epsilon_1\hat{Z}\right) \\
F_1(\epsilon) &= \theta\epsilon^{-\theta-1}\exp\left(-\epsilon\hat{Z}\right)
\end{aligned}$$

where $\hat{Z} = \sum_{i=1}^n \alpha_i^{-\theta}$. Then, equation (12) can be written as:

$$\begin{aligned}
p_{1r} &= \int \frac{\hat{Z}}{\hat{Z}}\theta\epsilon^{-\theta-1}\exp\left(-\epsilon^{-\theta}\hat{Z}\right)d\epsilon \\
&= \frac{1}{\hat{Z}}\int \hat{Z}\theta\epsilon^{-\theta-1}\exp\left(-\epsilon^{-\theta}\hat{Z}\right)d\epsilon
\end{aligned}$$

This expression is the derivative of equation (6) with respect to ϵ. Hence, we have:

$$\begin{aligned} p_{1r} &= \frac{1}{\hat{Z}} \int dF(\epsilon) \\ &= \frac{1}{\hat{Z}} \\ &= \frac{\tilde{w}_{1r}^{\theta}}{\sum_{i=1}^{N} \tilde{w}_{ir}^{\theta}} \end{aligned}$$

□

We can interpret $\tilde{w}_{ir}$ as a net reward of a person from region r and occupation i with mean ability. Therefore, $\tilde{w}_{ir}$ is composed of wage per efficiency unit, schooling, teacher's human capital, and distortions. In this context, occupations with high w_i will attract more workers in all regions r. On the other hand, differences in the occupational choice of individuals are driven by distortions in the labor market and in the market for educational goods. Therefore, the fraction of individuals who choose the sector i is low when these individuals encounter barriers in the formation of human capital (τ^h is high) and barriers in the labor market (τ^w is high). In the next proposition we defines the workers' human capital in each occupation in a given region.

Proposition 2 (Average quality of workers)**.** *For a given region, the human capital of workers in occupation i is:*

$$H_{ir} = p_{ir}\mathbb{E}[h(e_{ir}, s_i)\epsilon_i | person\ choices\ i], \tag{13}$$

The average quality of workers is:

$$\mathbb{E}[h(e_{ir}, s_i)\epsilon_i | person\ choices\ i] = \bar{\Gamma}\left[\left(\frac{1-\tau_{ir}^{w}}{1+\tau_{ir}^{h}} w_{ir}\right)^{\eta} \tilde{h}_{ir} p_{ir}^{-\frac{1}{\theta}}\right]^{\kappa} \tag{14}$$

where $\bar{\Gamma} = \Gamma(1-\kappa/\theta)$ is related to the mean of the Fréchet distribution for abilities, $\tilde{h}_{ir} = [(p_{tr}^{\alpha} H_{tr}^{1-\alpha})^{\varphi} s_i^{\phi_i} \eta^{\eta}]^{\kappa}$ and $\kappa = 1/(1-\eta)$.

Proof. We have:

$$h(e_{ir}, s_i)\epsilon_i = (p_{tr}^{\alpha} H_{tr}^{1-\alpha})^{\varphi}\left[\eta\left(\frac{1-\tau_{ir}^{w}}{1+\tau_{ir}^{h}} w_{ir}\right)(p_{tr}^{\alpha} H_{tr}^{1-\alpha})^{\varphi} s_i^{\phi}\epsilon_i\right]^{\eta\kappa} s_i^{\phi_i}\epsilon_i \tag{15}$$

H_{ir} is the total efficiency units of labor supplied to occupation i in region r. Then,

$$\begin{aligned}
H_{ir} &= p_{ir}\mathbb{E}\left\{ (p_{tr}^{\alpha}H_{tr}^{1-\alpha})^{\varphi}\left[\eta\left(\frac{1-\tau_{ir}^{w}}{1+\tau_{ir}^{h}}w_{ir}\right)(p_{tr}^{\alpha}H_{tr}^{1-\alpha})^{\varphi}s_i^{\phi_i}\epsilon_i\right]^{\eta\kappa} s_i^{\phi_i}\epsilon_i \middle| \text{person choices } i\right\} \\
&= p_{ir}\left\{ (p_{tr}^{\alpha}H_{tr}^{1-\alpha})^{\varphi}\left[\left(\frac{1-\tau_{ir}^{w}}{1+\tau_{ir}^{h}}w_{ir}\right)\eta(p_{tr}^{\alpha}H_{tr}^{1-\alpha})^{\varphi}s_i^{\phi_i}\right]^{\eta\kappa} s_i^{\phi_i}\mathbb{E}\left[\epsilon_i^{\kappa}\middle| \text{person choices } i\right]\right\} \\
&= p_{ir}\tilde{h}_{ir}\left(\frac{1-\tau_{ir}^{w}}{1+\tau_{ir}^{h}}w_{ir}\right)^{\eta\kappa}\mathbb{E}\left[\epsilon_i^{\kappa}\middle| \text{person choices } i\right] \qquad (16)
\end{aligned}$$

To calculate this last conditional expectation, we use the Fréchet distribution. For now, we suppress the region index r because this calculation is similar in all regions. Let $y_i = \tilde{w}_i\epsilon_i$. Since we are maximizing y_i, it inherits the extreme value distribution:

$$\begin{aligned}
\mathbf{Pr}\left(\operatorname*{Max}_i y_i < z\right) &= \mathbf{Pr}(\epsilon_i < z/\tilde{w}_i) \quad \forall i \\
&= F(z/\tilde{w}_1, ..., z/\tilde{w}_N) \\
&= \exp\left[-\sum_{i=1}^{N}(z/\tilde{w}_i)^{-\theta}\right] \\
&= \exp\left[-kz^{-\theta}\right]
\end{aligned}$$

where $k = \sum_i^N \tilde{w}_i^{\theta}$.

The extreme value also has a Fréchet distribution. After some algebraic manipulations it can be concluded that the distribution of ϵ^*, the workers' ability in their chosen occupation, has a Fréchet distribution:

$$G(x) = \mathbf{Pr}(\epsilon^* < x) = \exp\left[-k^* x^{-\theta}\right] \qquad (17)$$

where $k^* = \sum_{i=1}^{N}(\tilde{w}_i/\tilde{w}^*)^{\theta} = 1/p^*$.

Finally, we can calculate the expectation of equation (16). Let i be the occupation the individual chooses, and λ some positive exponent.

$$\begin{aligned}
\mathbb{E}(\epsilon_i^{\lambda}) &= \int_0^{\infty}\epsilon_i^{\lambda}dG(\epsilon) \\
&= \int_0^{\infty}\theta\left(\frac{1}{p^*}\right)\epsilon^{(\lambda-\theta-1)}\exp\left[\left(\frac{1}{p^*}\right)\epsilon^{-\theta}\right]d\epsilon
\end{aligned}$$

We can set $x = \left(\frac{1}{p^*}\right)\epsilon^{-\theta}$ and rewrite the last expression as:

$$
\begin{aligned}
\mathbb{E}(\epsilon_i^{\lambda}) &= \left(\frac{1}{p^*}\right)^{\frac{\lambda}{\theta}} \int_0^{\infty} x^{-\frac{\lambda}{\theta}} \exp(-x) dx \\
&= \left(\frac{1}{p^*}\right)^{\frac{\lambda}{\theta}} \Gamma\left(1 - \frac{\lambda}{\theta}\right)
\end{aligned}
$$

Using this result in equation (16) completes the proof. □

This result points to a selection effect in the economy. The average quality in equation (14) is inversely related to the share of workers in occupation p_{ir}. If the distortion is high in occupation i and region r, only the most qualified workers are selected for that occupation. For example, in a region where it is easy to become a teacher, their average human capital will be small (intensive margin). On the other hand, holding the average constant, a higher share of workers in an occupation will result in higher aggregate human capital (extensive margin). The net effect depends on the parameters' relative size. If $\theta(1-\eta) > 1$, then the extensive margin dominates. Otherwise, the intensive margin dominates. Next, we solve the model for the average wage in occupation i and region r.

Corollary 1 (Gross average wages)**.** *Let W_{ir} be the gross average wage in occupation i in region r. Then:*

$$
W_{ir} = w_{ir}\mathbb{E}[h(e_{ir}, s_i)\epsilon_i] = \bar{\Gamma}\eta \frac{(1-s_i)^{-1/\beta}}{(1-\tau_{ir}^{w})} \left(\sum_{i=1}^{N} \tilde{w}_{ir}^{\theta}\right)^{\frac{\kappa}{\theta}} \tag{18}
$$

This result is a consequence of Proposition 2. Equation (18) shows that gross average wages in a given region differ among occupations due to schooling and labor market distortions. Occupations with higher workers' schooling or labor market distortions have greater gross average wages. In addition, equation (18) is essential to understand the average wages differences across regions. From equation (3), we conclude that in equilibrium $A_r = w_{ir}$. Then, $\tilde{w}_{ir}$ is a function of A_r, and consequently, W_{ir} is a function of regional TFP. Therefore, distortions, schooling, and TFP are important sources of average wage variation across states. Finally, we use a standard definition of a competitive equilibrium.

2.3 Equilibrium

Definition 2 (Competitive equilibrium)**.** *A competitive equilibrium in this economy consists of individual choices of* $\{c, e, s\}$*, an occupational choice by workers, total human capital in each occupation and region* H_{ir}*, final output* Y*, and efficiency wages* w_{ir} *for each occupation and region.*

(i) Given an occupational choice, w_{ir}*, and the idiosyncratic ability* ϵ*, each worker chooses* c*,* e*,* s *to maximize utility in equation* (7)*.*

(ii) Given market friction, w_{ir}*,* H_{it}*, and* ϵ*, a worker chooses the occupation that maximizes* D_{ir}*.*

(iii) A representative firm hires H_{ir} *to maximize profits.*

(iv) The occupational wage, w_{ir}*, clears the labor market in each occupation and region.*

(v) Total output is given by the production function in equation (1)*.*

3 Empirical Investigation

Our goal is to use our model to empirically study the Brazilian economy. Then, in our baseline calibration of the model, we Brazilian data for the year of 2015. In this exercise, we can study how market frictions influence human capital formation and other matters in Brazil. In addition, we try to explore dynamic aspects of the Brazilian economy and we also calibrate our model using data from the year of 2003. That exercise is important because we can analyse the convergence of income across the Brazilian states.

Our calibration strategy involves choosing parameters values of our model such that the competitive equilibrium is consistent with the Brazilian states dataset for 2015. We use four state-level data from National Household Sample Survey (PNAD): years of schooling; work hours; gross earnings; and occupation. After some adjustments[5], we have a sample of 109038 individuals distributed among eight big groups of occupation: 1) managers (except public

[5] We drop individuals with no occupation and those whose wage was less than 60% of the minimum wage. Therefore, we drop individuals that receive considerably less than the minimum wage. We also selected individuals between 25 and 65 years old. Concerning the occupations, we drop individuals with not well-defined occupations and those in the army. There is a code in the Brazilian Occupation Code (CBO) with workers classified in not well-defined occupations.

sector); 2) professionals of sciences and arts; 3) middle-level technicians; 4) administrative service; 5) service-sector; 6) sellers and service providers; 7) agriculture; 8) goods and industrial production, services, and repairs-maintenance. We aggregate groups 4, 5, and 6 into the service-sector workers. Finally, we separate those working as teachers. Thus, we have the following categories of occupation:

1. Managers (except public sector);
2. Professionals of sciences and arts (except teachers);
3. Middle-level technicians (except teachers);
4. Service-sector;
5. Agriculture;
6. Goods and industrial production, services and repairs-maintenance;
7. Teachers.

We consider each one of the regions in our model as one of the 26 Brazilian states[6] and the Federal District (DF) of Brazil. Thus, in our calibrated model presents $N = 7$ and $R = 27$.

Based on our model, we split the parameters into three groups, constant parameters (η, θ, φ, β, α), the elasticity of human capital concerning time in school (ϕ_i), distortions (τ_{ir}^{w} and τ_{ir}^{h}) and TPF (w_{ir}).

3.1 Constant Parameters

To estimate skill dispersion θ and the elasticity of spending on educational goods in the human capital formation function η to Brazil, we follow Hsieh et al. (2019) and assumption that wages within an occupation for a given region follow a Fréchet distribution with shape parameter $\theta(1-\eta)$. So, wage dispersion depends on the dispersion of talent (given by $1/\theta$) and amplification from accumulating human capital via spending (given by $1/(1-\eta)$). Therefore, the coefficient of variation (CV) of wages within the occupation and region can be written as:

[6] Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE), Tocantins (TO).

$$CV = \frac{\Gamma\left(1 - \frac{2}{\theta(1-\eta)}\right)}{\left(\Gamma\left(1 - \frac{1}{\theta(1-\eta)}\right)\right)^2} - 1 \tag{19}$$

The next step is take the residuals of cross-sectional regression of log worker wages hourly on 6×26 occupation-state dummies. So, we calculate the mean and variance of the exponent of these wage residuals and solve equation (19) for $\theta(1-\eta)$ using a root-finding algorithm. The result for 2003 is 2.39 and for 2015 it is 2.00, and the average across years it is 2.19.

The parameter η is equal to the fraction of output spent on human capital accumulation. So, η can be calculated as a ratio between spent on education as a share in GDP and share of labour compensation in GDP. Public and private expenditures on education as a proportion of GDP for the Brazilian economy, estimated for 2003 are 0.0642 and for 2015 are 0.0793 and the average is 0.0717.[7] Labor compensation is 0.53 and 0.58, on average 0.5571.[8] Therefore, the estimated η is 0.129. Since we have $\theta(1-\eta)$ and η values it is easy to find the θ value which is 2.52.

The other constant parameters of the model are specified according to the Table 1. It is noteworthy that we chose $\alpha = 0.6$ as our baseline value, but we do not have information about this parameter. Furthermore, we follow Hsieh et al. (2019) and use $\beta = 0.231$. Therefore in subsection 4.5, we will explore the robustness of our results by alternating the α, β, θ and η values.

Table 1: Baseline constant parameters

Parametes	Value	Description	Source
η	0.129	Elasticity of educational goods in the human capital function	Estimated using data from PNAD 2015 and 2003
φ	0.129	Elasticity of teacher's human capital in the human capital function	Assumption that $\varphi = \eta$
θ	2.52	Dispersion of skills	Estimated using data from PNAD 2015 and 2003
α	0.60	Weight of the share of teachers in T_r	Assumption
β	0.231	Consumption preference	Hsieh et al. (2019)

Source: Elaborated by authors.

3.2 Estimation of ϕ_i

The ϕ's compose the second group. We use equation (8) and years of schooling to estimate each occupation's ϕ. First, we compute the average years of schooling of each occupation that can be view in Table 2. Then, we calculate the effective time spent in education[9]. We assume

[7] In Appendix C we provide more details about estimation of spent on education in Brazil.
[8] Labour compensation as share of GDP for the Brazilian economy can be view in Penn World Table 10.0.
[9] An agent can consume leisure in the education step.

that the individual studies six hours a day on weekdays, so he studies $252 \times 6 = 1512$ hours a year. Therefore, of the 8760 hours available in a year, he spends 17.26% of his time studying. As the schooling period is composed of the first 25 years of the life cycle, we divide the average years of schooling by 25 and multiply it by the time available to education (0.1726). The calculated ϕ_i can be view in Table 3.

Table 2: Descriptive statistics of years of schooling among occupations

Occupation	Mean	1° Quartile	Median	3° Quartile	Variance
Managers	12.77	12	12	16	3.36
Sciences and arts	14.98	16	16	16	2.39
Middle-level technicians	12.67	12	12	15	2.58
Service-sector	9.91	7	12	12	3.79
Agriculture	6.07	3	5	9	3.92
Industrial production and services	8.75	6	9	12	3.69
Teachers	15.13	15	16	16	1.69

Source: Elaborated by the authors with data from PNAD 2015.

Table 3: Elasticity of human capital concerning time in school ϕ_i

Ocuppation	ϕ
Managers	0.31
Sciences and arts	0.37
Middle-level technicians	0.31
Service-sector	0.24
Agriculture	0.14
Industrial production and services	0.21
Teachers	0.38

Source: Elaborated by the authors with data from PNAD 2015.

3.3 Calibration of $\tau's$ and $A's$

The remaining parameters, τ's and A's, are calibrated using the Method of Moments by minimizing the distance between the statistics of our simulated model and the Brazilian states' data. In this group, we have $2NR + R$ parameters. We use two statistics for each occupation and region: the share of workers; and the average gross wage.

We calculate the average hourly wage based on the hours worked present in the PNAD microdata and we log this. Both, the average hourly wage and the share of workers, by occupation and region, can be seen in Appendix A. In our model those statistics are described in equations

(11) and (18). Also, we use the FOC's of firm's problem, where $w_{i,r} = A_r \, \forall \, i, r$, to recover the equilibrium wage rate.

Each occupation's share of workers sum to one in each region, $\sum_{i=1}^{N} p_{ir} = 1$. This means that we only have $(N-1)R$ independent statistics in each region. Thus, we assume that $\tau_{1r}^{h} = 0 \, \forall \, r$. Beside, we assume that $\tau_{1r}^{w} = \tau_{1}^{w} \quad \forall r$, i.e, that frictions in occupation 1 are equal across regions. Also, we fix A_R, the last region's TFP.[10] Thus, we have the same number of statistics and parameters to be fitted $(2(N-1)R + R)$.

We define the following objective function to our numerical routine:

$$\mathcal{M} = \sum_{i=1,r=1}^{N,R} \left(\frac{W_{ir}^{M} - W_{ir}^{D}}{W_{ir}^{D}} \right)^2 + \sum_{i=1,r=1}^{N-1,R} \left(\frac{p_{ir}^{M} - p_{ir}^{D}}{p_{ir}^{D}} \right)^2 \tag{20}$$

where the superscripts M and D indicate model and targets statistics, respectively.[11]

To minimize equation (20) we use the Nelder-Mead algorithm that finds $\mathcal{M} = 0.00092$, which we consider it a small number, because we have 378 different targets. Figures 1 (a) and (b) presents the adjustment of the average wage hourly of our model to the data and the adjustment of the share of workers of our model to the data, respectively. Note that the model and data have a good adjust, because the points very close to the 45° line in both figures. The calibrated values to τ_{ir}^{w}, τ_{ir}^{h} and w_r can be view in Appendix B.

[10]Since TFP is equal to a price in our model, we select A_R from a grid. We run the Method of Moments for each values of A_R in a grid between 10 and 30. Then, we select the value to A_R which implies the minimum distance between data and our simulated model. Our algorithm selects $A_R = 25$. This is the TFP of Amapá (AP). We can think that all other TFPs are relative to this value.

[11]We apply the logarithm in equation (20) to improve the numerical stability of the problem.

Figure 1: Model adjustment to data - wages and share of workers

4 Results

In this section, we discuss the main results of our model calibrated with 2015 data. We first compare our results with a set of stylized facts. We then do a series of counterfactual exercises to see how sensitive GDP is to changes in the labor market and educational distortions. In addition, we calibrated the model for 2003 and compared it with the 2015 model to study absolute income convergence across Brazilian states. Finally, we proceed with many robustness checks to verify how sensitive are our results to changes in constant parameters.

4.1 Comparing model results with a set of stylized facts

Our calibrated model has an excellent fit to GDP per worker as can be seen in Figure 2 (a), recall that we do not use GDP as target. It can also be seen in Figure 2 (b) that our model suggests a positive relationship between GDP and TFP.

Figure 2: GDP - Model

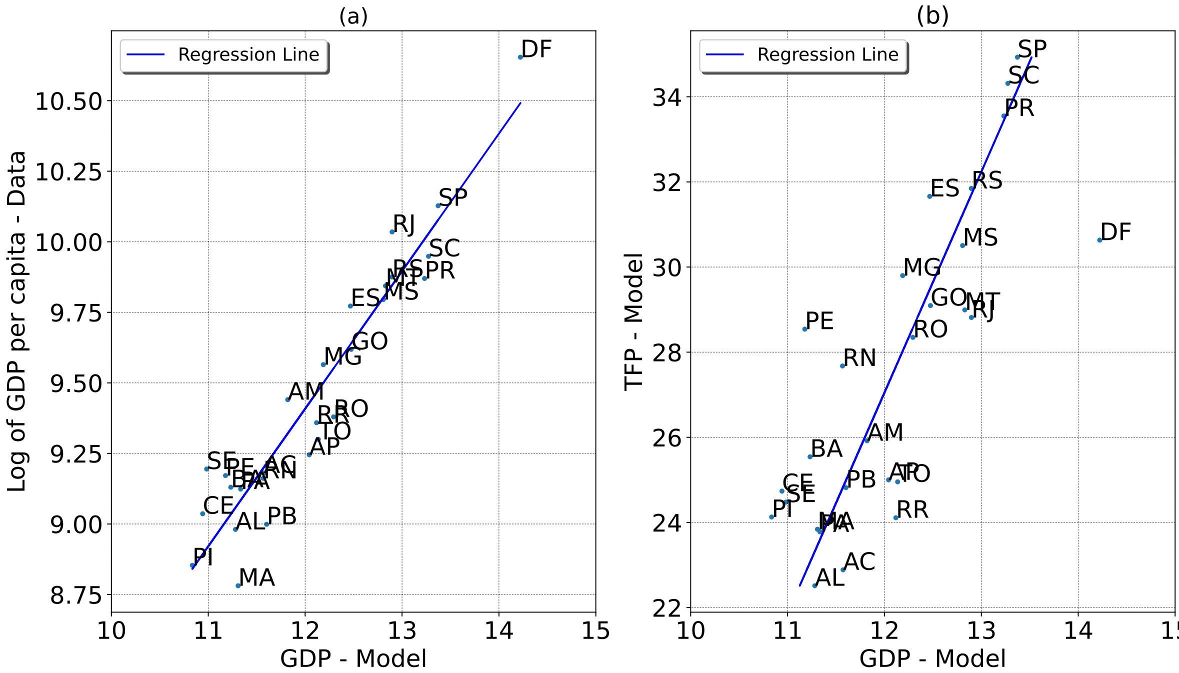
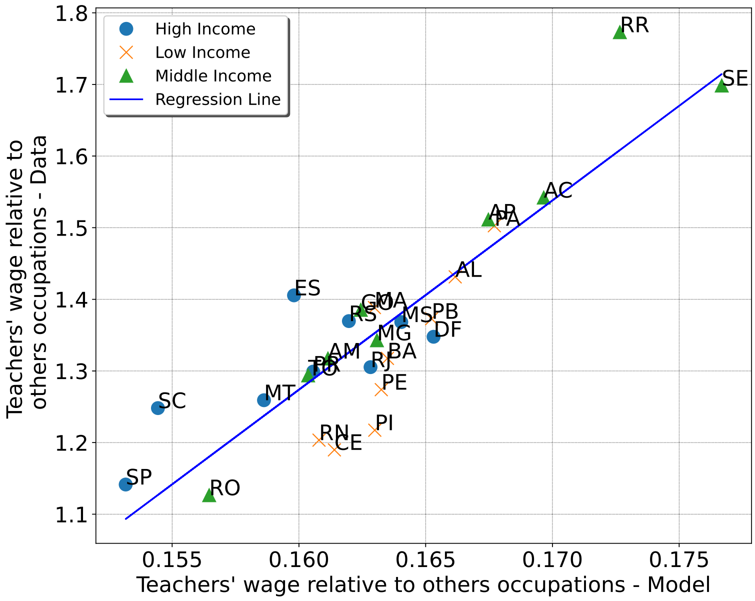


Figure 3 shows teacher's wage relative to occupations for model and data.[12] Note that our model also displays a good fit for relative wages. Furthermore, the Figure shows that, on average, in low-income and middle-income states, teachers have a higher relative wage than in high income states.[13] Barros and Delalilbera (2018) argue that one of the reasons why the relative wage of teachers is higher in the poorest states is that the occupation of teachers is labor-intensive and is not much affected by technological and structural changes. Therefore, in states where more advanced technologies are present, the relative salary of teachers is lower than in less developed states. It is noteworthy that the higher the relative wage of teachers, the greater will be the share of people who will choose this occupation. The consequence of this can be seen in the Figure 4.

[12]In Brazil, Law N 11.738 of 2008 regulates the national minimum wage for public teaching professionals in basic education. It is noteworthy that, when we look at the data, we notice that there are professionals in the area who receive amounts higher than the minimum amount stipulated by law. Therefore, the average hourly wage of teachers varies between states, as can be seen in Table A1 from Appendix A.

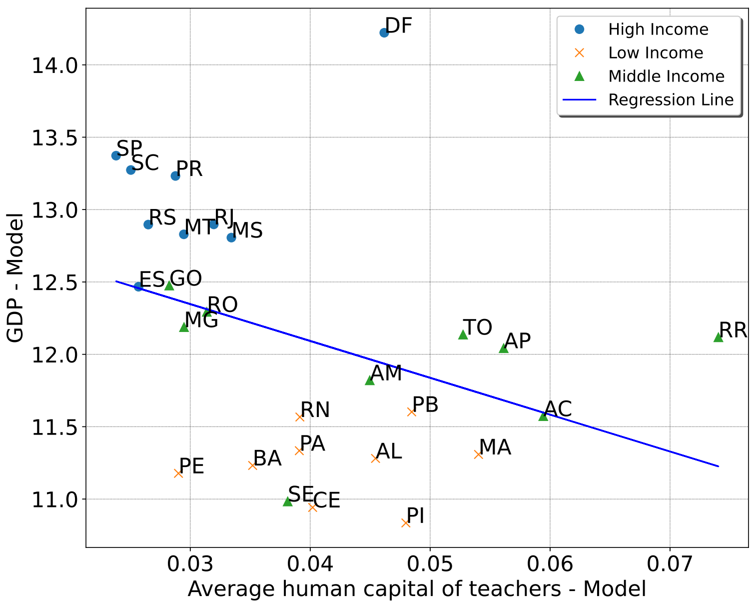
[13]We rank the 27 Brazilian states using 2015 GDP *per capita* data. The first nine are considered high income, the middle nine are middle income, and the last nine are low income.

Figure 3: Teachers' relative wage



Since relative wage affect individual decisions. The low and middle income states are able to attract more talented people to the teacher occupation compared to the high income states. This fact compromises the formation of human capital and consequently the output of high-income states. It is worth mentioning that, according to Machado and Scorzafave (2016), teachers' wages can also affect their efforts in the classroom and discourage teacher turnover in schools.

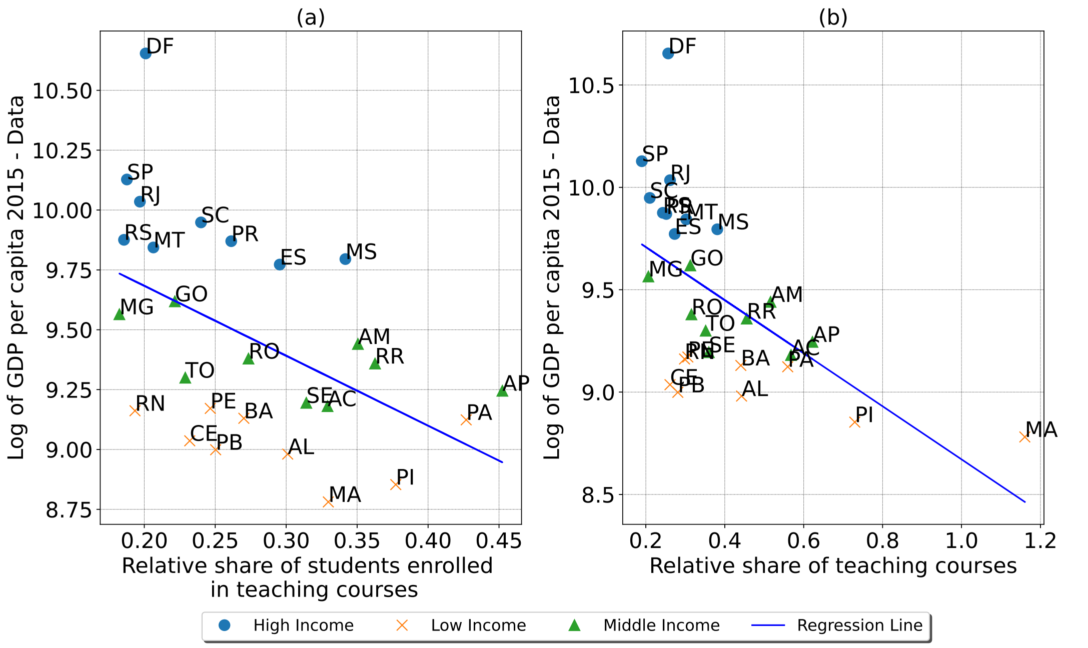
Figure 4: GDP per worker and average human capital



Both Figures (3 and 4) capture the effects of the distortions in the labor (τ^w) and educational market (τ^h) of the model. In poorer states, the teaching occupation has fewer barriers in the labor and educational market. Then, in those states it is easier to become a teacher compared to richer states. When we look at the data from the Brazilian Higher Education Census we find evidence of this.

Figure 5 (a) shows that, on average, in the poorest states there are more students enrolled in teaching courses than in the richest states. This can be explained by the fact that in the poorest states there is a greater offer of these courses, as can be seen in the Figure 5 (b). This distortion can be caused by the government itself, since it can offer more teaching courses in the poorest states, given that its budget constraint is limited and teaching courses have a lower cost when compared to engineering and science courses.[14] Therefore, individuals choose the occupation according to those incentives.

Figure 5: Share of students in teaching courses, and share of teaching courses offered



Note: This figure was elaborated using data from the Higher Education Census of the 2015, provided by the National Institute of Educational Studies and Research Anísio Teixeira (INEP).

We also highlight that the agricultural occupation a presents the greatest average distortion in the educational market, as can be seen in Appendix B. We do not report it here, but when we look at the 2015 PNAD data we notice that most workers in this occupation reside in the countryside and in these places there may be difficulties in finding and attending schools.

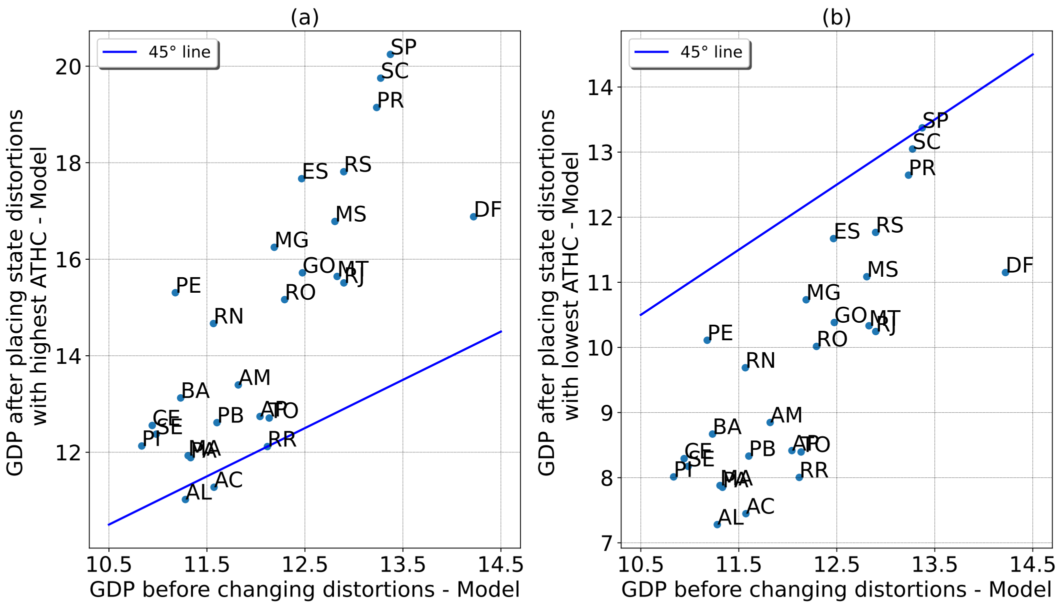
[14] Magalhães et al. (2010) using data from the Federal University of Viçosa provide evidence that the cost per student is higher in agricultural sciences and biological and health sciences courses than in courses in the humanities, letters and arts.

Furthermore, the quality of rural schools may be lower than that of urban ones[15]. Therefore, this high distortion may be evidence of the barriers that agricultural workers face when trying to accumulate human capital.

4.2 GDP and productivity gains

To check how sensitive the economy is to changes in teacher occupation distortions and to analyze what the best allocation in the economy would be, we conduct some counterfactual exercises. In the first exercise, we input the distortions, τ^w and τ^h, of the states with highest and lowest Average Teachers Human Capital (ATHC) in all other states. Figure 6 (a) shows that if all the Brazilian states have distortions of Roraima (RR), the state with highest ATHC, there would be a generalized growth in the GDP of all states. In this case, the Brazilian GDP would increase 21.26%. It is noteworthy that, by doing this, the relative wage of teachers in all states would become equal to the relative wages in the Roraima state. On the other hand, if the states have the same distortions of São Paulo (SP), the state with lowest ATHC, the GDP of all states would decreases(Figure 6 (b)), and the Brazilian GDP would decrease 24.85%.

Figure 6: GDP before and after placing state distortions with the highest and lowest ATHC in all states



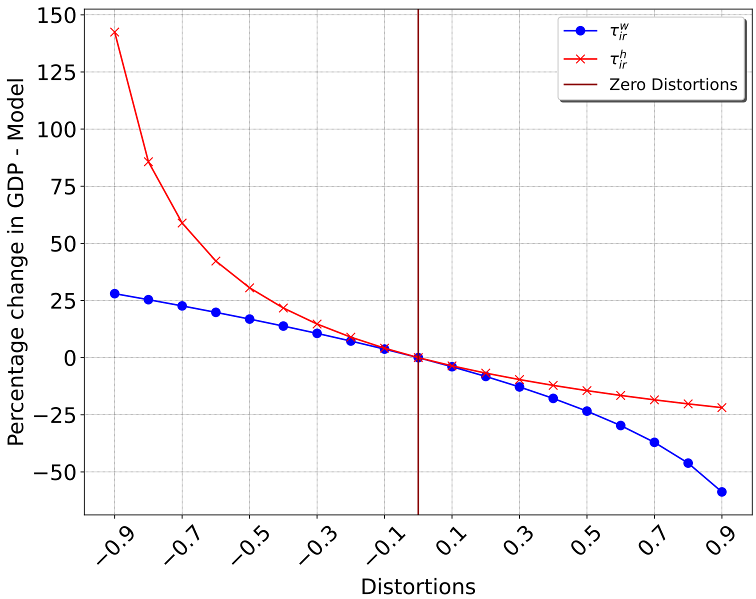
From this exercise, we can infer that there is a problem of misallocation of work in the Brazilian economy. When we reduce distortions related to the teacher occupation, the relative

[15]See Williams (2005) and Zhang (2006).

wage of teachers increases, since incentives affect individuals' occupational choices, there is a reallocation of talent in occupations. Thus, since there is an externality in the choice to be a teacher, a change in the distortions that induces more talented people choose this occupation, has a big effect on state GDP because all workers benefit from the increase in the teachers quality and quantity.

In the next exercise, we look at how changes in market frictions in all occupations affect GDP per capita. First, we calculate GDP per capita considering no distortions in labor and educational markets as a reference value. Then fixing the distortion in the educational goods market at zero, we vary the labor market distortion from -0.9 to 0.9, and then we calculate the percentage change in the GDP per capita in relation to the reference value. Next, we conduct a similar exercise, where we fix the frictions in the labor market to zero and vary the distortion in the educational goods market. The results can be seen in Figure 7. It is noted that when they become negative, distortions in the educational goods market are more efficient in increasing GDP. On the other hand, labor market distortions have a greater negative effect when they are positive.

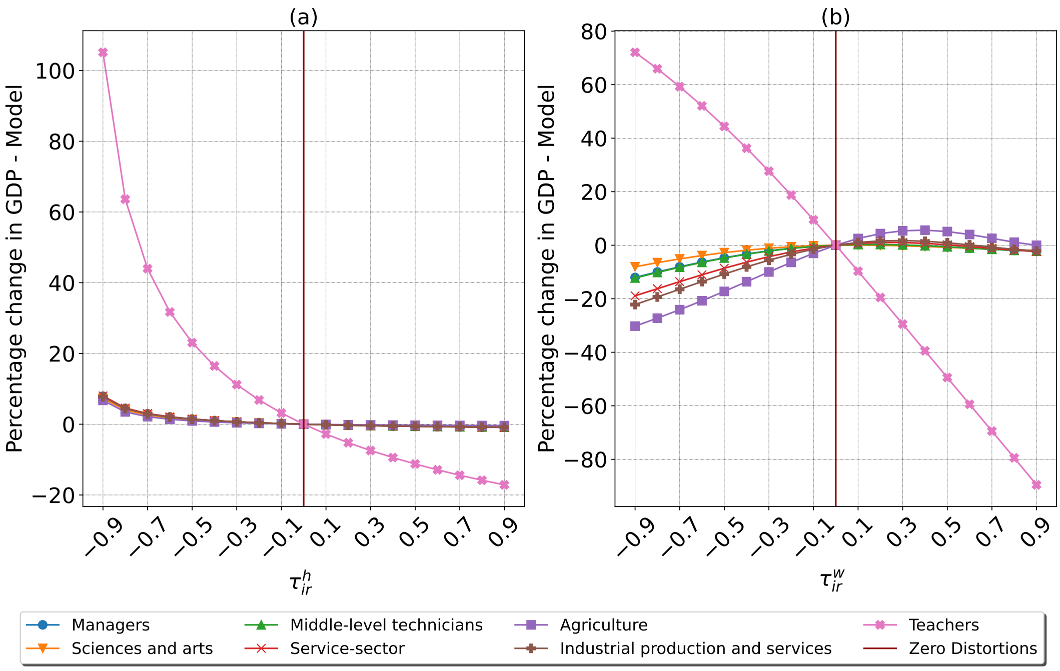
Figure 7: Increases in the distortions of all occupations and the percentage effects on GDP - Model



4.3 Teacher's human capital

In the next exercise, instead of varying the distortions of all occupations simultaneously, we vary the distortions of each occupation keeping the others at zero. The Figures 8 (a) and (b) shows that GPD is more sensitive to the changes in the distortions related to occupation 7, the teacher occupation. In other words, the greater the distortion of teachers, the lower the GDP. Again, it is noted that negative distortions in the educational goods market, which become incentives, are more efficient in increasing GDP. On the other hand, labor market distortions have a greater negative effect when they are positive. This leads us to the conclusion that in order to promote GDP growth, a public policy should create incentives to more qualified people to become a teacher.

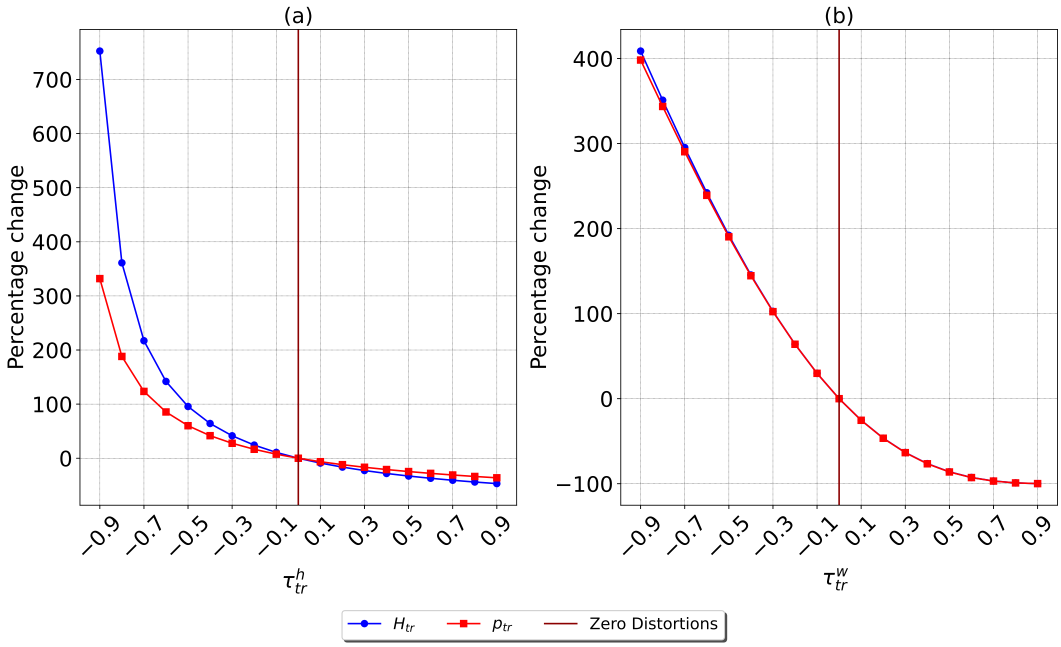
Figure 8: Increases in the distortions of each occupation and the percentage effects on GDP



One of the possible causes of this result can be seen in Figure 9. In this exercise, we did something similar to the exercise in Figure 8, but instead of measuring the percentage change in GDP, we verified the percentage change in the human capital of teachers and in the proportion of teachers. It is noted that, increases in the distortions cause a decrease in the human capital of teachers and in the proportion of employed teachers, and a decrease in the distortions has the opposite effect.

When we reduce distortions, the teacher's human capital increases. It is as if we are removing barriers to their human capital formation and entry into the labor market. By lowering barriers, being a teacher becomes more attractive, which in turn makes a greater proportion (p_{ir}) of people want to be a teacher. However, the effect of the ratio on human capital is ambiguous, as more people becoming teachers can reduce average human capital. On the other hand, it can make more talented people become teachers, which increases the average human capital. [16]

Figure 9: Increases in distortions and percentual effects on H_{tr} and p_{tr}



Note: H_{tr} = Human capital of teachers, p_{tr} = proportion of workers in teacher occupation.

4.4 Conditional convergence

Absolute income convergence, according to Barro and Sala-i Martin (1992), is the fact that poorer economies tend to grow faster than richer economies in per capita terms. Hence, the income gap between poor and rich economies tends to narrow over time. In this sense, using data from the model calibrated for 2015 and 2003, we verified whether there is income convergence between the Brazilian states by estimating the following equation via OLS:

$$\frac{1}{T}\log\left(\frac{Y_{r,2015}}{Y_{r,2003}}\right) = a + b\log(Y_{r,2003}) + \epsilon_r \tag{21}$$

where $Y_{r,2015}$ and $Y_{r,2003}$ are the 2015 and 2003 GDPs of the region r, T is the number of

[16] Remember the equation ref eq23.

periods, a and b are constant, and ϵ_r is the error term. The existence of a negative b supports the convergence hypothesis.

From the results reported in Table 4, it can be said that there is absolute convergence of income between Brazilian states within our model since b is negative and significant. From these results we can calculate the speed of convergence of this economy, which is $\beta_s = 7.58\%$[17]. This result can be better interpreted using the concept of half life, which is the time required for the income gap between the poorest and richest economies to be reduced by half. The half life is given by $HL = \log(2)/\beta_s$, and for our economy it is 9.14.

Table 4: OLS to verify the absolute convergence of income between Brazilian states between 2003 and 2015

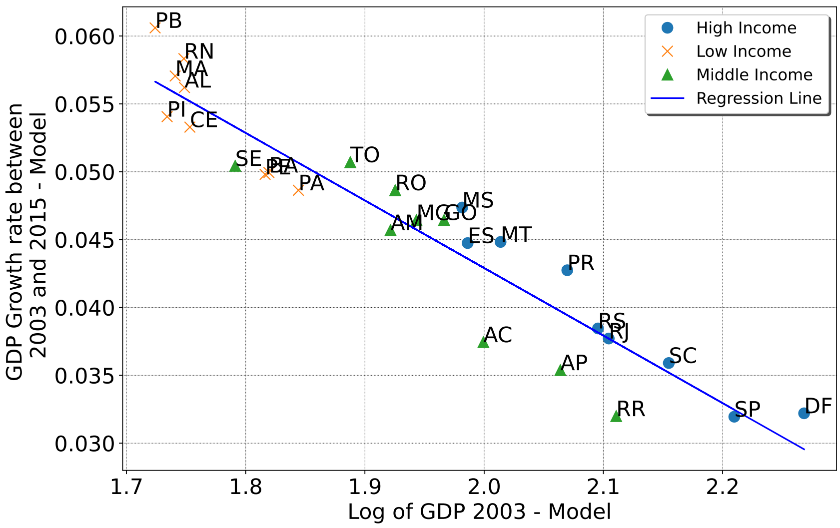
	$\frac{1}{T}\log\left(\frac{Y_{r,2015}}{Y_{r,2003}}\right)$
a	0.1424***
	(0.0066)
b	-0.0498***
	(0.0034)
R-squared	0.8964
R-squared Adj.	0.8922

Source: Search results.
Notes: Standard errors in parentheses. Single (*), double (**) and triple (***) asterisk denote statistical significance at 10%, 5% and 1%, respectively.

In Figure 10 the result of absolute income convergence becomes clearer. It is noted that low-income states such as Paraíba (PB), Rio Grande do Norte (RN), Maranhão (MA), Alagoas (AL), Piauí (PI) and Ceará (CE) were the ones that grew the most in this period. On the other hand, high-income states such as the Federal District (DF), São Paulo (SP), Santa Catarina (SC), Rio de Janeiro (RJ) and Rio Grande do Sul (RS) showed lower growth rates.

[17] Speed of convergence is given by: $\beta_s = -\frac{\log(Tb+1)}{T}$.

Figure 10: Growth rate from 2003 to 2015 and Log of GDP 2003



The absolute convergence of income that we show can be explained, above all, by the reduction in educational market distortions and the increase in the TFP in 2015 compared to 2003[18]. We note that distortions in the educational market, on average, have reduced more sharply in the poorest states and in Rio Janeiro (RJ). As for occupations, there was a decrease, on average, in the occupation of teachers, however, in agriculture it was the one with the greatest reduction. This fact may be associated with the increase in the level of average years of education of Brazilian workers during this period.

We also noticed when comparing 2003 with 2015 that, on average, there was a small increase in labor market distortions. However, as we saw in Figure 7, increases in labor market distortions have a smaller effect on GDP change when compared to increases in educational market distortions.

4.5 Robustness check

Table 5 presents robustness check for constant parameters of our model. In this robustness check, we repeat the counterfactual exercise presented in section 4.2, where we put the distortions, τ^w and τ^h, of the states with highest and lowest Average Teachers Human Capital (ATHC) in all states to verify the effects on GDP. The benchmark line consists of the GDP variation found using baseline values.

[18]Calibrated distortions and the TFP for 2003 can be seen in Appendix B.

In our counterfactual exercise, by varying the elasticity of educational goods in the human capital function η using the distortions of the state with the highest ATHC (Roraima) in all states, the variation in GDP increases substantially from 9.19% with $\eta = 0.05$, to 21.26% with our baseline $\eta = 0.129$, to 43.39% with $\eta = 0.20$. When we use the distortions of the state with the lowest ATHC (São Paulo), the variation in GDP showed a similar behavior, but with the opposite sign. Note that, among all constant parameters, GDP variation is more sensitive to variations in η.

Table 5: Robustness check for constant parameters

Parameter	GDP variation (Largest ATHC)	GDP variation (Lowest ATHC)
$\eta = 0.25$	43.49%	-43.73%
$\eta = 0.05$	9.19%	-9.36%
$\theta = 2.0$	14.93%	-21.99%
$\theta = 3.0$	21.77%	-25.10%
$\beta = 0.1$	17.41%	-17.55%
$\beta = 0.3$	21.86%	-24.99%
$\alpha = 0.3$	26.46%	-30.04%
$\alpha = 0.8$	19.01%	-22.10%
Benchmark	21.26%	-24.85%

Source: Search results.
Notes: ATHC is Average teacher human capital.
Recall that our baseline values are $\eta = 0.129$, $\beta = 0.231$, $\theta = 2.52$ and $\alpha = 0.6$.

In the next two lines of Table 5 we vary the value of the skill dispersion θ. In the second column we see that when skill dispersion is $\theta = 2$, the variation in GDP is 14.93%, and 21.26% with our baseline value and 21.77% when $\theta = 3$. In the third column, it can be noted that the change in GDP is a little more sensitive to changes in θ. In other words, in an economy with greater distortion, the dispersion of skills has a greater impact on GDP variation. By changing the β we notice that the change in GDP is similar to the case of the θ, that is, values lower than β are associated with a lower change in GDP.

Finally, we vary the weight of teacher participation in T_r. Note that by increasing α we are increasing the proportion of teachers and decreasing their average human capital, within the model[19]. As can be seen in the second column, when $\alpha = 0.3$ the GDP variation is 26.46%, against 21.26% using the baseline value and 19.01% when $\alpha = 0.8$.

[19]Recall that $T_r = p_{tr}^{\alpha} H_{tr}^{(1-\alpha)}$. Where H_{tr} is the average of human capital of teachers and p_{tr} is the proportion of teachers, in region r.

5 Conclusion

In this paper, we develop a General Equilibrium model to analyze the negative relationship between the relative wage of teachers and regional economic development across the Brazilian States. In this model, teachers' human capital is a source of positive externalities for the workforce and individuals' occupational choices are distorted by barriers in the labor and educational market. These barriers cause individuals to choose an occupation in which they do not have a comparative advantage, consequently, they can cause less talented people to choose to be a teacher, which in turn harms the entire workforce and economic development.

We calibrated this model using the Method of Moments and Brazilian data from 2015. Our counterfactual exercises show that a reduction in the distortions would increase Brazilian GDP by a value between 9.19% and 43.49%. This is mainly because, by reducing distortions, the relative wage of teachers increases. Since occupational choices are affected by incentives, this increase in relative wages causes a reallocation of talent in the economy.

We also show that increasing distortions in teacher occupation have a more negative impact on GDP per capita than increasing distortions in other occupations. This is because it reduces the human capital of teachers and the proportion of individuals in this occupation, therefore, the human capital of the workforce is harmed. This is in line with the evidence we find in the literature that analyzes the effects of teachers' human capital endowment on students.

Furthermore, when we compare the calibrated model for the years of 2015, we note that the absolute convergence of income was mainly due to the reduction in educational market distortions, which in turn increased the average human capital and productivity. In this sense, policy makers should focus on increasing incentives to enter the teaching profession so that it can attract more talented people.

References

Bacolod, Marigee P, "Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality," *The Review of Economics and Statistics*, 2007, *89* (4), 737–751.

Banerjee, Abhijit V and Benjamin Moll, "Why does misallocation persist?," *American Economic Journal: Macroeconomics*, 2010, *2* (1), 189–206.

_ **and Esther Duflo**, "Growth theory through the lens of development economics," *Handbook of economic growth*, 2005, *1*, 473–552.

Barro, Robert J and Xavier Sala i Martin, "Convergence," *Journal of political Economy*, 1992, *100* (2), 223–251.

Barros, Fernando Jr. and Bruno R. Delalilbera, "Market frictions, misallocation of talent and development," *Economics Bulletin*, 2018, *38* (4), 2410–2430.

Benhabib, Jess and Mark M Spiegel, "Human capital and technology diffusion," *Handbook of economic growth*, 2005, *1*, 935–966.

Borensztein, Eduardo, Jose De Gregorio, and Jong-Wha Lee, "How does foreign direct investment affect economic growth?," *Journal of international Economics*, 1998, *45* (1), 115–135.

Café, Renata Motta, "A Alocação de talentos no setor público brasileiro." PhD dissertation 2018.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Hanushek, Eric A and Ludger Woessmann, "Schooling, educational achievement, and the Latin American growth puzzle," *Journal of Development Economics*, 2012, *99* (2), 497–512.

_ **, Marc Piopiunik, and Simon Wiederhold**, "The value of smarter teachers international evidence on teacher cognitive skills and student performance," *Journal of Human Resources*, 2019, *54* (4), 857–899.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh Paul, "Castes and labor mobility," *American Economic Journal: Applied Economics*, 2012, *4* (2), 274–307.

Hsieh, Chang-Tai and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, *124* (4), 1403–1448.

_ , **Erik Hurst, Charles I. Jones, and Peter J. Klenow**, "The allocation of talent and US economic growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Krueger, Alan B, "Economic considerations and class size," *The Economic Journal*, 2003, *113* (485), F34–F63.

Lakdawalla, Darius, "The economics of teacher quality," *The Journal of Law and Economics*, 2006, *49* (1), 285–329.

Machado, Laura Muller and Luiz Guilherme Dácar da Silva Scorzafave, "Distribuição de salários de professores e outras ocupações: uma análise para graduados em carreiras tipicamente ligadas à docência," *Revista Brasileira de Economia*, 2016, *70* (2), 203–220.

Magalhães, Elizete Aparecida, Suely de Fátima Ramos Silveira, Luiz Antônio Abrantes, Marco Aurélio Marques Ferreira, and Vasconcelos Reis Wakim, "Custo do ensino de graduação em instituições federais de ensino superior: o caso da Universidade Federal de Viçosa," *Revista de Administração Pública*, 2010, *44*, 637–666.

Mankiw, N Gregory, David Romer, and David N Weil, "A contribution to the empirics of economic growth," *The quarterly journal of economics*, 1992, *107* (2), 407–437.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., *Frontiers of Econometrics*, New York, NY: Academic Press, 1974, pp. 105 – 142.

Menezes-Filho, Naércio and Elaine Pazello, "Do teachers' wages matter for proficiency? Evidence from a funding reform in Brazil," *Economics of Education Review*, 2007, *26* (6), 660–672.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic dynamics*, 2008, *11* (4), 707–720.

_ **and** _ , "Misallocation and productivity," 2013.

_ **and** _ , "The causes and costs of misallocation," *Journal of Economic Perspectives*, 2017, *31* (3), 151–74.

Romer, Paul M, "Endogenous technological change," *Journal of political Economy*, 1990, *98* (5, Part 2), S71–S102.

Stoddard, Christiana, "Why has the number of teachers per student risen while teacher quality has declined?: The role of changes in the labor market for women," *Journal of Urban Economics*, 2003, *53* (3), 458–481.

Williams, JH, "Cross-national variations in rural mathematics achievement," *Journal of Research in Rural Education*, 2005, *20* (5), 20–5.

Woessmann, Ludger, "The importance of school systems: Evidence from international differences in student achievement," *Journal of Economic Perspectives*, 2016, *30* (3), 3–32.

Zhang, Yanhong, "Urban-rural literacy gaps in Sub-Saharan Africa: The roles of socioeconomic status and school quality," *Comparative Education Review*, 2006, *50* (4), 581–602.

Appendix A Descriptive Statistics

Table A1: Descriptive statistics of teachers' hourly wages by state

State	Relative Wage	Mean	1° Quartile	Median	3° Quartile	Variance	Income Group
AC	1.57	17.53	9.72	14.29	23.81	10.46	Middle Income
AL	1.45	16.04	9.38	13.91	20.37	9.65	Low Income
AM	1.30	16.13	9.52	14.29	20.24	8.63	Middle Income
AP	1.48	19.77	13.17	17.80	23.53	10.03	Middle Income
BA	1.33	15.40	8.33	11.90	17.86	11.97	Low Income
CE	1.21	14.04	8.33	11.90	15.87	11.17	Low Income
DF	1.49	27.93	13.69	23.81	35.71	17.42	High Income
ES	1.29	18.16	9.68	15.01	21.33	12.99	High Income
GO	1.42	19.30	10.19	14.29	22.55	15.18	Middle Income
MA	1.30	15.93	8.93	11.90	20.40	12.50	Low Income
MG	1.37	18.42	9.52	14.29	21.65	13.78	Middle Income
MS	1.45	21.84	11.11	17.86	26.19	16.22	High Income
MT	1.31	18.93	11.90	17.06	21.43	10.58	High Income
PA	1.50	17.81	9.38	14.29	21.71	13.82	Low Income
PB	1.37	17.13	9.04	12.50	21.60	12.55	Low Income
PE	1.29	14.99	7.28	11.43	19.05	11.57	Low Income
PI	1.28	14.24	9.52	13.10	15.67	7.72	Low Income
PR	1.36	21.04	11.90	17.27	23.81	14.86	High Income
RJ	1.33	20.00	9.52	15.87	23.81	15.05	High Income
RN	1.24	15.53	7.37	11.90	18.45	13.25	Low Income
RO	1.25	16.21	10.39	14.07	17.86	10.30	Middle Income
RR	1.62	22.34	9.72	20.22	29.17	14.21	Middle Income
RS	1.38	20.40	10.84	15.16	23.81	15.12	High Income
SC	1.21	18.15	11.90	14.88	20.83	11.36	High Income
SE	1.70	19.61	9.40	16.67	26.19	13.52	Middle Income
SP	1.13	18.61	9.52	14.88	23.15	14.07	High Income
TO	1.30	17.32	9.38	14.58	19.05	13.09	Middle Income

Source: Elaborated by the authors with data from PNAD 2015.
Notes: Relative wage is the average hourly wage of teachers divided by the average hourly wage of other six occupations. Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal(DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE), Tocantins (TO).

Table A2: Logarithm of average hourly wages by occupation and state

	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers
AC	2.87	2.86	2.19	2.00	2.04	2.06	2.86
AL	2.68	2.99	2.29	1.95	2.06	1.95	2.77
AM	2.97	3.05	2.46	2.02	1.91	2.07	2.78
AP	3.11	3.13	2.50	2.08	1.86	2.10	2.98
BA	2.81	3.10	2.37	1.92	1.78	2.02	2.73
CE	2.83	3.13	2.44	1.92	1.53	1.87	2.64
DF	3.42	3.47	2.97	2.36	2.24	2.36	3.33
ES	2.92	3.21	2.74	2.06	2.03	2.27	2.90
GO	2.97	3.02	2.64	2.14	2.24	2.26	2.96
MA	3.10	2.99	2.44	1.99	1.75	1.95	2.77
MG	2.92	3.20	2.63	2.04	2.00	2.17	2.91
MS	3.05	3.27	2.70	2.14	2.29	2.26	3.08
MT	2.93	3.22	2.56	2.16	2.37	2.36	2.94
PA	2.94	2.92	2.48	1.96	2.00	1.99	2.88
PB	2.80	3.24	2.45	1.95	1.91	2.01	2.84
PE	2.91	3.07	2.36	1.89	1.72	1.94	2.71
PI	2.92	3.00	2.29	1.88	1.65	1.92	2.66
PR	3.10	3.23	2.79	2.23	2.26	2.32	3.05
RJ	3.03	3.42	2.65	2.15	1.88	2.27	3.00
RN	2.97	3.13	2.51	2.00	1.77	1.92	2.74
RO	2.89	3.00	2.56	2.08	2.20	2.28	2.79
RR	3.04	3.26	2.66	2.03	1.80	2.11	3.11
RS	3.05	3.24	2.68	2.19	2.22	2.22	3.02
SC	3.01	3.18	2.73	2.30	2.33	2.34	2.90
SE	2.92	3.02	2.39	1.91	1.65	1.97	2.98
SP	3.25	3.30	2.83	2.22	2.22	2.35	2.92
TO	2.90	3.20	2.54	2.10	2.01	2.19	2.85

Source: Elaborated by the authors with data from PNAD 2015.

Table A3: Average years of schooling by occupation and state

	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers
AC	10.82	11.82	11.69	9.71	5.26	7.79	15.40
AL	11.33	14.24	12.91	8.83	4.70	7.20	15.13
AM	12.48	14.48	12.24	10.09	5.39	9.40	15.20
AP	11.82	14.50	12.84	9.88	5.92	8.38	14.95
BA	12.09	14.51	12.27	9.80	4.52	8.19	14.69
CE	11.57	14.45	12.46	9.65	4.69	8.44	15.19
DF	13.76	15.18	12.88	10.55	6.61	8.90	15.47
ES	12.07	14.85	12.83	9.82	6.82	8.91	15.59
GO	12.69	14.43	12.47	9.77	6.85	8.73	15.45
MA	12.35	14.62	12.27	9.60	5.32	8.07	14.36
MG	12.57	14.97	12.54	9.62	6.02	8.41	15.11
MS	12.71	14.98	12.48	9.58	6.39	8.42	15.11
MT	12.08	14.79	12.37	10.14	6.71	8.54	15.39
PA	11.82	13.58	11.68	9.63	5.05	8.01	15.09
PB	12.88	14.82	11.81	9.51	4.41	7.29	15.32
PE	12.09	15.06	12.50	9.58	5.38	8.02	15.13
PI	12.03	14.35	12.37	9.21	5.05	7.13	15.27
PR	12.93	14.81	12.79	10.01	7.32	9.02	15.29
RJ	12.99	15.17	12.76	10.02	6.29	9.16	15.02
RN	11.47	14.57	12.13	9.85	4.67	8.09	15.03
RO	11.08	15.03	11.69	9.78	6.45	8.03	15.28
RR	11.83	13.95	13.05	10.34	5.83	8.23	15.16
RS	12.66	14.92	12.73	9.98	7.10	8.75	15.42
SC	12.60	14.69	12.57	10.20	7.57	9.18	15.37
SE	12.07	14.68	12.11	9.53	4.27	7.30	15.20
SP	13.48	15.27	13.02	10.18	7.46	9.38	15.09
TO	11.40	14.34	12.34	9.98	6.13	9.05	15.20

Source: Elaborated by the authors with data from PNAD 2015.

Table A4: Share of workers in each occupation by state

	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers
AC	0.04	0.03	0.05	0.42	0.13	0.25	0.09
AL	0.04	0.04	0.07	0.41	0.12	0.25	0.07
AM	0.05	0.05	0.08	0.38	0.07	0.29	0.08
AP	0.04	0.03	0.07	0.44	0.07	0.25	0.10
BA	0.05	0.04	0.06	0.45	0.08	0.26	0.06
CE	0.05	0.03	0.06	0.44	0.05	0.31	0.07
DF	0.06	0.10	0.10	0.48	0.01	0.17	0.08
ES	0.07	0.06	0.07	0.35	0.12	0.29	0.05
GO	0.05	0.05	0.06	0.42	0.08	0.29	0.05
MA	0.04	0.04	0.06	0.37	0.14	0.27	0.09
MG	0.06	0.06	0.06	0.39	0.10	0.28	0.06
MS	0.06	0.06	0.05	0.38	0.12	0.27	0.06
MT	0.05	0.05	0.06	0.35	0.16	0.28	0.05
PA	0.03	0.04	0.05	0.44	0.11	0.27	0.06
PB	0.05	0.05	0.07	0.42	0.08	0.25	0.08
PE	0.05	0.06	0.07	0.46	0.05	0.25	0.06
PI	0.04	0.03	0.05	0.40	0.10	0.30	0.08
PR	0.08	0.07	0.07	0.36	0.07	0.29	0.06
RJ	0.05	0.08	0.08	0.46	0.01	0.25	0.06
RN	0.06	0.05	0.07	0.42	0.06	0.26	0.07
RO	0.06	0.04	0.05	0.35	0.16	0.28	0.06
RR	0.05	0.03	0.07	0.39	0.11	0.24	0.11
RS	0.06	0.08	0.08	0.39	0.06	0.28	0.05
SC	0.08	0.06	0.07	0.32	0.09	0.31	0.06
SE	0.04	0.04	0.05	0.43	0.14	0.25	0.06
SP	0.07	0.09	0.08	0.41	0.03	0.27	0.05
TO	0.05	0.04	0.05	0.33	0.21	0.23	0.09

Source: Elaborated by the authors with data from PNAD 2015.

Appendix B Calibrated $\tau's$ and $A's$

Table B1: Labor market distortions τ_{ir}^{w} - 2015

State	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers	**Mean distortion by state**	Income Level
AC	0.95	0.95	0.94	0.94	0.95	0.94	0.95	0.95	Middle Income
AL	0.95	0.96	0.95	0.94	0.95	0.94	0.95	0.95	Low Income
AM	0.95	0.95	0.94	0.94	0.94	0.94	0.95	0.95	Middle Income
AP	0.95	0.95	0.94	0.94	0.94	0.94	0.95	0.94	Middle Income
BA	0.95	0.95	0.94	0.94	0.94	0.94	0.95	0.95	Low Income
CE	0.95	0.95	0.95	0.94	0.93	0.94	0.95	0.94	Low Income
DF	0.95	0.95	0.95	0.94	0.94	0.94	0.95	0.95	High Income
ES	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	High Income
GO	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	Middle Income
MA	0.95	0.95	0.94	0.93	0.93	0.94	0.94	0.94	Low Income
MG	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	Middle Income
MS	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	High Income
MT	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	High Income
PA	0.95	0.95	0.95	0.94	0.94	0.94	0.95	0.95	Low Income
PB	0.95	0.96	0.95	0.94	0.95	0.94	0.95	0.95	Low Income
PE	0.95	0.95	0.94	0.93	0.94	0.94	0.95	0.94	Low Income
PI	0.95	0.95	0.94	0.93	0.93	0.94	0.94	0.94	Low Income
PR	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	High Income
RJ	0.95	0.96	0.95	0.94	0.94	0.95	0.95	0.95	High Income
RN	0.95	0.95	0.95	0.94	0.94	0.94	0.95	0.94	Low Income
RO	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	Middle Income
RR	0.95	0.95	0.95	0.94	0.94	0.94	0.95	0.95	Middle Income
RS	0.95	0.95	0.95	0.94	0.95	0.94	0.95	0.95	High Income
SC	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	High Income
SE	0.95	0.95	0.94	0.94	0.93	0.94	0.95	0.94	Middle Income
SP	0.95	0.95	0.95	0.94	0.94	0.94	0.94	0.95	High Income
TO	0.95	0.95	0.95	0.94	0.95	0.95	0.95	0.95	Middle Income
Mean by occupation	0.95	0.95	0.95	0.94	0.94	0.94	0.95	–	–

Source: Search results. Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal(DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE), Tocantins (TO).

Table B2: Education market distortions τ^h_{ir} - 2015

State	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers	**Mean distortion by state**	Income Level
AC	0	0.11	1.84	−0.71	1.57	−0.37	−0.72	0.25	Middle Income
AL	0	−0.51	−0.12	−0.78	0.81	−0.49	−0.69	−0.26	Low Income
AM	0	−0.21	0.21	−0.56	9.36	−0.32	−0.47	1.15	Middle Income
AP	0	0.24	0.54	−0.66	12.54	−0.14	−0.63	1.70	Middle Income
BA	0	−0.45	0.29	−0.70	6.99	−0.36	−0.44	0.76	Low Income
CE	0	−0.24	0.32	−0.69	31.72	−0.31	−0.46	4.33	Low Income
DF	0	−0.61	−0.08	−0.60	166.13	0.93	−0.42	23.62	High Income
ES	0	−0.42	0.33	−0.36	4.54	−0.33	0.07	0.55	High Income
GO	0	−0.15	0.36	−0.68	3.62	−0.49	−0.25	0.34	Middle Income
MA	0	−0.02	0.59	−0.58	4.17	−0.17	−0.60	0.48	Low Income
MG	0	−0.50	0.43	−0.50	5.74	−0.27	−0.18	0.68	Middle Income
MS	0	−0.31	1.10	−0.46	2.68	−0.18	−0.26	0.37	High Income
MT	0	−0.37	0.70	−0.60	0.53	−0.54	−0.26	−0.08	High Income
PA	0	−0.26	0.12	−0.77	1.77	−0.48	−0.62	−0.04	Low Income
PB	0	−0.56	0.19	−0.63	6.46	−0.22	−0.60	0.66	Low Income
PE	0	−0.43	0.56	−0.55	24.81	0.10	−0.09	3.49	Low Income
PI	0	0.19	1.18	−0.56	8.54	−0.30	−0.53	1.22	Low Income
PR	0	−0.25	0.58	−0.32	9.18	−0.08	0.12	1.32	High Income
RJ	0	−0.76	0.04	−0.70	125.09	−0.35	−0.32	17.57	High Income
RN	0	−0.23	0.51	−0.53	19.04	0.27	−0.27	2.68	Low Income
RO	0	0.11	0.79	−0.55	0.99	−0.50	−0.21	0.09	Middle Income
RR	0	0.18	0.22	−0.53	7.53	−0.06	−0.75	0.94	Middle Income
RS	0	−0.55	0.20	−0.54	10.88	−0.19	−0.01	1.40	High Income
SC	0	−0.12	0.79	−0.34	5.74	−0.23	0.45	0.90	High Income
SE	0	−0.36	0.36	−0.71	3.68	−0.34	−0.63	0.29	Middle Income
SP	0	−0.48	0.46	−0.40	48.78	0.05	0.65	7.01	High Income
TO	0	−0.28	0.79	−0.55	0.82	−0.29	−0.60	−0.02	Middle Income
Mean by occupation	0	−0.27	0.49	−0.58	19.40	−0.21	−0.32	–	–

Source: Search results.

Table B3: Total productivity factors - 2015

State	A_r	State	A_r	State	A_r
AC	22.89	MA	23.84	RJ	28.81
AL	22.51	MG	29.80	RN	27.68
AM	25.92	MS	30.50	RO	28.35
AP	25.00	MT	28.99	RR	24.11
BA	25.54	PA	23.78	RS	31.85
CE	24.74	PB	24.82	SC	34.31
DF	30.63	PE	28.54	SE	24.48
ES	31.66	PI	24.13	SP	34.93
GO	29.09	PR	33.55	TO	24.95

Source: Search results.
Notes: Recall that in our model TFP is equal across occupations.
The average of TFP is 27.61.

Table B4: Labor market distortions τ_{ir}^{w} - 2003

State	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers	**Mean distortion by state**	Income Level
AC	0.90	0.90	0.87	0.80	0.72	0.80	0.86	0.84	Middle Income
AL	0.90	0.91	0.85	0.76	0.68	0.81	0.85	0.83	Low Income
AM	0.90	0.90	0.81	0.78	0.73	0.78	0.84	0.82	Middle Income
AP	0.90	0.90	0.86	0.80	0.77	0.83	0.87	0.85	Middle Income
BA	0.90	0.90	0.88	0.79	0.75	0.81	0.85	0.84	Low Income
CE	0.90	0.90	0.88	0.80	0.68	0.79	0.86	0.83	Low Income
DF	0.90	0.91	0.88	0.83	0.89	0.83	0.89	0.88	High Income
ES	0.90	0.89	0.88	0.81	0.82	0.82	0.87	0.86	High Income
GO	0.90	0.90	0.89	0.81	0.84	0.83	0.87	0.86	Middle Income
MA	0.90	0.90	0.83	0.78	0.76	0.79	0.86	0.83	Low Income
MG	0.90	0.90	0.88	0.81	0.81	0.84	0.89	0.86	Middle Income
MS	0.90	0.91	0.88	0.83	0.89	0.82	0.87	0.87	High Income
MT	0.90	0.90	0.86	0.80	0.84	0.83	0.86	0.86	High Income
PA	0.90	0.91	0.87	0.80	0.85	0.82	0.88	0.86	Low Income
PB	0.90	0.90	0.87	0.77	0.71	0.76	0.86	0.83	Low Income
PE	0.90	0.90	0.86	0.79	0.70	0.79	0.85	0.83	Low Income
PI	0.90	0.92	0.86	0.78	0.69	0.73	0.84	0.82	Low Income
PR	0.90	0.90	0.88	0.82	0.86	0.84	0.88	0.87	High Income
RJ	0.90	0.90	0.88	0.82	0.70	0.84	0.89	0.85	High Income
RN	0.90	0.91	0.88	0.80	0.67	0.83	0.88	0.84	Low Income
RO	0.90	0.91	0.89	0.81	0.87	0.84	0.89	0.87	Middle Income
RR	0.90	0.88	0.87	0.80	0.79	0.86	0.87	0.85	Middle Income
RS	0.90	0.90	0.88	0.83	0.85	0.84	0.88	0.87	High Income
SC	0.90	0.90	0.90	0.85	0.88	0.86	0.88	0.88	High Income
SE	0.90	0.92	0.86	0.80	0.73	0.80	0.85	0.84	Middle Income
SP	0.90	0.89	0.88	0.83	0.82	0.84	0.87	0.86	High Income
TO	0.90	0.91	0.88	0.79	0.82	0.83	0.86	0.86	Middle Income
Mean by occupation	0.90	0.90	0.87	0.80	0.78	0.82	0.87	–	–

Source: Search results.

Table B5: Education market distortions τ_{ir}^{h} - 2003

State	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers	**Mean distortion by state**	Income Level
AC	0	0.43	3.93	1.91	723.61	10.96	0.31	105.88	Middle Income
AL	0	−0.15	4.05	5.90	122.74	6.19	0.46	19.88	Low Income
AM	0	0.58	9.47	2.35	1007.36	6.27	1.90	146.85	Middle Income
AP	0	−0.47	1.33	1.12	288.02	0.66	−0.22	41.49	Middle Income
BA	0	0.43	1.67	3.31	77.23	6.87	2.00	13.07	Low Income
CE	0	−0.15	1.38	1.15	417.73	4.95	0.63	60.81	Low Income
DF	0	−0.62	0.77	1.20	168.84	7.63	−0.01	25.40	High Income
ES	0	0.33	2.59	2.74	28.81	4.79	1.67	5.85	High Income
GO	0	0.14	1.43	1.82	23.17	3.43	1.51	4.50	Middle Income
MA	0	0.41	6.78	3.31	40.01	6.04	0.20	8.11	Low Income
MG	0	−0.07	1.87	2.32	56.59	2.88	0.29	9.13	Middle Income
MS	0	0.70	5.02	1.81	8.21	6.72	1.96	3.49	High Income
MT	0	0.15	5.88	5.40	12.36	3.93	2.22	4.28	High Income
PA	0	−0.45	1.40	0.82	51.82	2.35	−0.01	7.99	Low Income
PB	0	0.49	2.52	4.50	190.48	15.56	0.22	30.54	Low Income
PE	0	−0.13	2.03	2.06	395.90	9.28	1.38	58.64	Low Income
PI	0	−0.51	2.19	2.71	187.22	27.75	0.15	31.36	Low Income
PR	0	0.06	1.54	2.54	27.04	3.26	1.24	5.10	High Income
RJ	0	−0.65	0.83	0.56	2523.83	1.71	−0.03	360.89	High Income
RN	0	−0.52	2.31	0.80	261.58	1.68	−0.49	37.91	Low Income
RO	0	0.91	1.39	1.38	12.10	1.48	−0.25	2.43	Middle Income
RR	0	0.57	7.78	2.97	185.01	1.97	−0.09	28.32	Middle Income
RS	0	−0.16	1.12	1.68	36.27	3.06	1.00	6.14	High Income
SC	0	0.57	1.23	1.80	12.79	1.96	1.61	2.85	High Income
SE	0	−0.45	2.22	1.65	101.71	4.57	0.31	15.72	Middle Income
SP	0	−0.19	1.81	1.58	208.67	3.00	1.95	30.97	High Income
TO	0	0.14	2.31	6.53	24.79	5.50	0.77	5.72	Middle Income
Mean by occupation	0	0.05	2.85	2.44	266.44	5.72	0.77	–	–

Source: Search results.

Table B6: Total productivity factors - 2003

State	A_r	State	A_r	State	A_r
AC	20.26	MA	16.77	RJ	20.64
AL	17.41	MG	20.40	RN	14.40
AM	21.78	MS	22.50	RO	17.67
AP	17.50	MT	22.94	RR	18.81
BA	20.31	PA	17.56	RS	23.56
CE	17.34	PB	17.10	SC	24.52
DF	22.40	PE	19.10	SE	16.08
ES	22.85	PI	15.49	SP	26.22
GO	21.98	PR	23.14	TO	18.83

Source: Search results.
Notes: Recall that in our model TFP is equal across occupations. The average of TFP is 19.92.

Appendix C Public and private spending on education

Using table 49 from the Family Budget Survey (POF)[20] we estimate private expenditures on education in Brazil for 2003, 2009 and 2018. In 2003 these expenditures were around R$ 32.4 billion, in 2009 around R$ 40.5 billion and in 2018 about R$ 145.4 billion. As a percentage of GDP, private spending on education is 1.8, 1.2 and 2.0, respectively, on average 1.7.

Public spending on education as a percentage of GDP is provided by the National Institute of Educational Studies and Research Anísio Teixeira (INEP). In 2003 it was 4.6 and in 2015, 6.2. So we have that public and private spending on education, as a share of GDP, in Brazil, is 6.4 in 2003 and 7.9 in 2015.

Appendix D Migration between states

Using PNAD microdata, we found that 58% of people said they were born in the state where they lived. In Table D1 we show the share of people who were born and not born in the state where they live by occupation. It is noted that for all occupations, most people answered that they were born in the state where they live.

[20]Details about the POF can be seen on the Brazilian Institute of Geography and Statistics (IBGE) website.

Table D1: Proportion of people who were born and not born in the state in which they lived by occupation

	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers
Was born in this state	0.59	0.60	0.59	0.56	0.57	0.57	0.64
Was not born in this state	0.41	0.40	0.41	0.44	0.43	0.43	0.36

Source: Elaborated by the authors with data from PNAD 2015.

Also we verify that 89.13% of people did not lived in another Brazilian state or in other countries. In Table D2 we show the proportion of people that lived and did not lived in other states or countries by occupation. As can be seen, most workers have never lived in another Brazilian state or in other countries.

Table D2: Share of workers who have lived and have not lived in other brasilian states by occupation

	Managers	Sciences and arts	Middle-level technicians	Service sector	Agriculture	Industrial production and services	Teachers
Lived in other state	0.14	0.14	0.12	0.10	0.11	0.11	0.11
Did not live in other state	0.86	0.86	0.88	0.90	0.89	0.89	0.89

Source: Elaborated by the authors with data from PNAD 2015.

When we cross this information we find that 14.5% of the people who were born in the state in which they live, lived in another state or country. On the other hand, 85.5% of people who were born in the state they live in did not live in another state or country.