Wages differences between foreigners and natives: Evidence in Brazil

Paulo Domingos da Silva Matos
Doctoral Student in Applied Economics, Universidade Federal da Paraíba, Brazil

Wallace Patrick Santos Farias de Souza
Department of Economics, Universidade Federal da Paraíba, Brazil

Cássio da Nóbrega Besarria
Department of Economics, Universidade Federal da Paraíba, Brazil

The present research investigates the existence of salary differences between foreigners and natives in the Brazilian labor market. For the conditional average, it is highlighted that foreigners constitute a positively selected group, that is, they tend to receive higher salaries than natives, but this difference has decreased over the decades. Considering the results the conditional distribution for 1991 and 2000, we have found that the salary gap exhibits an inverted-U shape, with the greatest salary difference being observed in intermediate occupations. For 2010, the inverted-U shape is disfigured. Additionally, there was a significant reduction in salary differences between individuals in low-skilled.

Keywords. Foreigners, Labor market, Wages, Quantile Regression.

JEL classification. C21, F22, F61, J21.

1. Introduction

International migratory movements can be explained by several factors,[1] among which the economic aspect stands out. As a result, the labor market has been widely used for understanding these flows,[2] as it allows for the assessment of the occupational position and socioeconomic status of immigrants in the destination society. Nevertheless, more than a superficial assessment, this investigation is relevant for the design of public policies aimed at welcoming and integrating these individuals into society, given that these prerogatives are guaranteed by Human Rights and by the Federal Constitution of Brazil itself,[3] which provides for, among other things, social, labor and productive inclusion.

Paulo Domingos da Silva Matos : first@somewhere.com
Wallace Patrick Santos Farias de Souza : second@somewhere.com
Cássio da Nóbrega Besarria : third@somewhere.com
This work was carried out with the support of the Cactvs Educa Program.

[1] Wars, environmental disasters, political and/or cultural persecutions, ethnic conflicts.

[2] In Brazil, foreign labor flows were significant until 1940, after that period, internal migrations became responsible for the redistribution of workers in the country, in a way that the society practically closed itself to international inflows (Brito, 1995).

[3] Act 13.445/2017 deals with the rights and duties of migrants.

This is an open access article distributed under the terms of the Creative Commons Attribution License.

In the 1980s, migrations were typically characterized by the departure of workers (emigration) from less developed regions[4] towards developed countries such as the United States and Western European locations (Sales, 1991). According to Petit (2002), for example, the number of Latin Americans registered in the United States had doubled in that period. Additionally, reports from the United Nations Secretariat[5] pointed out that approximately 110 million migrants had moved to the aforementioned countries. It is noteworthy that this configuration was not an event exclusive to specific peoples, but a global exodus from emerging countries towards developed nations (Castles, 1998).

However, given the tightening of migration control policies and the slowdown of the world economy, emerging countries like Brazil have become an alternative route for migrants. According to information provided by the Brazilian Institute of Geography and Statistics (IBGE), the 1991 Census reports that 68,549 foreigners had taken up residence in Brazil, of which 8,095 obtained Brazilian nationality. In the 2000 Census, the decade in which Brazil's opening of trade takes place, the survey showed a growth of almost 41.14% in relation to the previous period, totaling approximately 96,752 foreigners. In 2010, the immigrant population grew approximately 70% when compared to the year 2000 and 139% compared to 1991. Therefore, these records make it clear that the sum of these events (hardening of migratory policies in developed countries, slowdown of the world economy and the opening of trade in Brazil) are relevant to understanding the migratory inflows directed to the country.

The first studies on migratory flows sought to understand the motivations for which individuals decide to migrate. Due to salary differences between the regions, Sjaastad (1962) argues that at the heart of this decision is the economic motivation, in which individuals aim for better occupations/salaries and living conditions. However, in addition to considering opportunities, moving from one location to another involves monetary and/or non-monetary costs[6] that negatively affect the mobility of individuals. In summary, it is reasonable to think that the decision to migrate almost always involves a cost-benefit analysis.

In this sense, the first generations understand that immigrants constitute a positively selected group, that is, after a period of assimilation, they tend to receive salaries greater than the individuals who were born in the place of destination (Chiswick, 1978; Carliner, 1980; Freitas, 1981). Chiswick (1978), for example, explains this result based on the unobservable characteristics (motivation, skills, effort, among others), stating that these individuals have greater investment incentives and, for that reason, their earnings tend to grow at a rate greater than the wages of the natives.

Nevertheless, Borjas (1987) shows that the aforementioned positive selection depends on the wage distribution in the country of origin. Considering poor countries,

[4] In general, due to the economic instabilities that these countries were facing Soares et al. (2015).

[5] UNITED NATIONS SECRETARIATE. Trends in Total Migrant Stock: the 2003 Revision. New York: Department of Economic and Social Affairs, Population Division, 2004a.

[6] For the author, monetary costs are those that increase the productivity of human resources and non-monetary costs correspond to opportunity costs such as: earnings lost during travel time, research and the need to learn skills for the new job.

where returns to education and wage dispersion are relatively high, there will be negative selection, that is, immigrants would receive lower wages than natives. This is because individuals who decide to migrate tend to be those with below-average qualifications in their country of origin. In rich countries, where returns to education and wage dispersion are relatively low, there will be positive selection of immigrants, so that the individuals most likely to migrate are those with above-average skills. Under this aegis, Borjas (1990) and Borjas (1995), explained that the new inflows to the United States were coming from poor countries that, compared to previous years, earned lower salaries and required longer for there to be a reduction in the wage gap between them and the natives, thus refuting the evidence in Chiswick (1978).

Using population censuses from the United States and Mexico, Chiquiar and Hanson (2005) tested the hypothesis formulated by Borjas (1987) and found that, although Mexicans were less educated than Native Americans, migrants were on average more educated than non-migrant Mexicans. Additionally, if these individuals who migrated to the United States were paid according to the salary per educational level in Mexico, they would be occupying medium and high positions in the wage distribution of the country of origin. Such results, therefore, did not confirm the tested premise.

In Brazil there is little evidence about the wage gap between foreigners and natives. In short, investigations have been restricted to regional, interstate and intercity migratory movements. In addition, the empirical strategies adopted have not been adequate to obtain robust results. The works Santos Junior (2002), Vilela (2011), de Cerqueira Rondon et al. (2013), among others, for example, ignore the existence of biases in the earnings equation Mincer (1974) when estimated via OLS (Ordinary Least Squares). Although the entry of immigrants in Brazil is primarily exogenous, that is, due to the hardening of migration policies in developed countries, slowdown of the world economy and opening of trade in Brazil, it is possible to identify two biases in the estimations of the mincerian earnings equation.

As documented by Heckman et al. (2006), the first source of bias comes from the correlation between the education variable and the stochastic term. According to the authors, there are omitted factors that can generate wage returns that are not necessarily higher educational attainment, but individual skills not observed. Consequently, it is almost unlikely to guarantee that the hypothesis of strict exogeneity required by the OLS estimator is respected.

The second source of bias comes from the estimation of log-linearized models by OLS. According to Figueiredo et al. (2014) identification of the multiplicative model does not imply the identification of the log-linear model, especially if the data present heterocedasticity. This result is a direct consequence of Jensen's[7] inequality to the conditional mean, where the expected value of the log of a random variable is different from the logarithm of its expected value, i.e., $E[(\ln y|X))] \neq \ln[E(y|X)]$. In other words, this result implies that the conditional mean is not invariant to monotonic transformations. In order to overcome this issue, the authors suggest that the estimates of the log-linearized

[7] Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a given probability space. Additionally, let $g : \mathcal{R} \to \mathcal{R}$ be a convex function, and let $y(\cdot)$ and $g(\cdot))$ be such that $E(y|X) < \infty$ and $E(g(y)|X) < \infty$, then $E(g(y)|X) \geq g(E(y)|X)$.

models be performed by Quantile Regression due to the equivalence property of quantiles, since these are invariant to monotonic transformations.

In this context, this study aims to analyze the existence of salary differences between foreigners and natives. The information used comes from the Brazilian Institute of Geography and Statistics (IBGE) corresponding to the 1991, 2000 and 2010 Demographic Census. To address the above-mentioned endogeneity problem, the Instrumental Variable (IV) estimator, whose instrument is the number of individuals residing in the household, is used. This instrument was proposed by Croix and Doepke (2003), but it is also a relevant result of the theoretical model of Becker and Lewis (1973). Generally speaking, Becker and Lewis (1973) show that there is a trade off in optimizing the utility function of parents involving the amount of children they wish to have and the quality of life they intend to offer them. Because of resource constraints, parents who decide to have many children invest less in their offspring's quality of life and, consequently, invest less in education. Thus, it is possible to note that the instrument meets the conditions of relevance and exogeneity, since the number of individuals is correlated with the education variable, but is independent of the salary. Finally, as suggested by Figueiredo et al. (2014), to bypass the second identification bias mentioned above, the Quantile Regression method proposed by Koenker and Bassett Jr (1978) is used, given that the quantiles have the property of equivalence, that is, they are invariant to monotonic transformations $Q_\tau\{g(y|x)\} = g(Q_\tau\{y|x\})$.

In addition to this introduction, this survey is organized as follows. Section 2 discusses the empirical strategy. Section 3 presents the results and section 4 presents the final considerations.

2. Methods

In this section, the empirical models used in order to properly analyze the existence of a wage gap between foreigners and natives in the Brazilian labor market. Initially, the Instrumental Variable (IV) model is used to estimate Mincer's earnings equation. Then, the Quantile Regression method proposed by Koenker and Bassett Jr (1978) is employed to check the wage gap along the conditional distribution and bypass the identification problem present in linearized log models.

2.1 *Instrumental variable and quantile regression*

To verify the existence of salary differences between natives and foreigners the conventional earnings equation proposed by Mincer (1974) will be used, which establishes that an individual's wages depends on their level of education (schooling), experience level and other individual attributes such as gender, race and color, for example. To achieve the present work, a binary variable (dummy) is included in the wage equation that indicates whether the individual is foreign or not. This constitutes the variable of interest of this study and its inclusion does not produce specification bias, since immigration to Brazil is caused by exogenous issues. As already mentioned, the main motivation for foreigners to enter the country stems from exogenous events such as the slowdown in

the world economy, the opening of trade and mainly the tightening of migratory policies in developed countries. Furthermore, the 2020 report produced by the Getúlio Vargas Foundation, FGV (2020), ratifies that Brazil has been chosen as a destination, but it is not the first option for immigrants. After this caveat, log-linearized wage equation can be expressed in the following way:

$$\ln w_i = \beta_0 + \beta_1 foreign_i + \beta_2 schooling_i + \beta_3 age_i + \beta_4 age_i^2 + \gamma X_i + \varepsilon_i \quad (1)$$

Where i corresponds to the $i-th$ individual, $foreign_i$ is a dummy that assumes value 1 when the individual is a foreigner and 0 otherwise, $\ln w_i$ corresponds to the wage logarithm,[8] $schooling_i$ is education measured in years of study a ge_i corresponds to an individual's age,[9] X_i is a vector of attributes such as race, gender, region, religion and ε_i is the stochastic term.

In general, the inference of the equation (1) can be obtained by the **Ordinary Least Squares** (OLS)[10] method, however Heckman et al. (2006) point out that this method tends to produce biased results, as there is a correlation between the education variable and the stochastic term. There are two reasons for that: (i) unobservable skills bias and (ii) Measurement errors. Skill positively affects individuals' wages and education, so that part of the return to education would come from a higher level of skill and not schooling itself. As a consequence, the hypothesis of exogeneity of the OLS estimator is violated, and it is not possible to guarantee that the covariance between the explanatory variable ($schooling_i$) and the error term is equal to 0, that is, $\mathrm{Cov}(schooling_i, \varepsilon_i) = 0$, which characterizes the endogeneity problem. Measurement errors stem from the fact that people describe their level of education inaccurately, as they report the level of education above the correct by cause of rounding for an entire period or a year. Thus, measurement errors also cause bias in the estimation by OLS (Harmon and Walker, 1995).

Under these circumstances, the most suitable inference for the expression (1) can be obtained through the Instrumental Variable (IV) method. The method requires the existence of an instrument that satisfies two conditions, namely: Relevance and Validity. According to Stock and Watson (2015), the relevance condition establishes that the covariance between the instrumental variable and the endogenous regressor must be different from 0, that is, $\mathrm{Cov}(schooling_i, Z_i) \neq 0$. In other words, this condition requires a strong correlation between the instrument and the endogenous variable. The validity hypothesis, in turn, postulates that the covariance between the stochastic term, ε_i, and the instrument, Z_i, must be equal to 0, $\mathrm{Cov}(Z_i, \varepsilon_i) = 0$. That is, the instrument must also satisfy the condition of exogeneity.

The problem of implementing the instrumental variable method is associated with the difficulty in finding instruments that meet the above conditions. Mariano and de Albuquerque (2011), for example, argue that several works try to correct the endogeneity

[8] Income received in 1991, 2000 and 2010 were duly converted and corrected for December 2019 based on the methodology informed by the Central Bank.

[9] Santos Junior (2002) and Gheasi et al. (2017) suggest using age as a proxy for experience.

[10] See Mincer (1974), Gheasi et al. (2017), Fernandes (2019).

bias from the regressor education, however, they diverge as to the choice of instruments, either because of the inadequacy of the instrument or the lack of data. Nevertheless, the quality of the instrument is essential for a robust inference, as the use of weak (or inadequate) instruments can lead to biased estimates, which should produce erroneous results and, consequently, misinterpretations.

The instrument used in this study, Z_i, is the number of people living in the family of the $i-th$ individual. This variable was proposed by Croix and Doepke (2003). The authors elaborate a model where fertility and education are related decisions. They suggest that because of budget constraints, families that decide to have more children invest less in education. Furthermore, this hypothesis is also validated in the Becker and Lewis (1973) theoretical model, in which parents maximize a utility function that depends not only on the amounts of children they want to have, but also the quality of life they want to offer their offspring. Therefore, parents who prefer to have a smaller number of children want to offer a better quality of life (for example, investing more in education) and vice versa. In Brazil, this instrument was used by Suliano and Siqueira (2012) when estimating returns to education for the Southeast and Northeast regions.

Assuming that Z_i satisfies the relevance and validity conditions, the estimation of the expression (1) can be decomposed into two stages. In the first stage, a regression of the endogenous variable ($shooling_i$) against the instrumental variable (number of people in the family), being able to insert other control variables. The first stage equation (or equation in reduced form) is given by:

$$schooling_i = \beta_0 + \beta_1 Z_i + \gamma X_i + \nu_i, \tag{2}$$

where $schooling_i$ is the endogenous variable, Z_i corresponds to the instrument, X_i is a vector of individual attributes and ν_i is the error term. First stage regression is estimated via (OLS). After concluding, the estimated values are used ($\widehat{schooling_i}$) in the second stage regression, which is expressed as follows:

$$\ln w_i = \beta_0 + \beta_1 foreign_i + \beta_2 \widehat{schooling_i} + \beta_3 age_i + \beta_4 age_i^2 + \gamma X_i + \varepsilon_i. \tag{3}$$

However, according to Figueiredo et al. (2014), the expression estimation (3) still presents bias when estimated by OLS. According to the authors, this problem is a direct consequence of Jensen's inequality, in which the expected value of the logarithm of a random variable is different from the logarithm of its expected value, $E(y|X) \neq g(E(y|X))$. That is so because the conditional mean is not invariant to monotonic transformations. To illustrate this point, following Figueiredo, Lima and Schaur (2014), let us allow the following model in its multiplicative form:

$$W_i = \exp\{x_i\beta\}\phi_i, \tag{4}$$

where ϕ_i is a non-negative random variable and $E(\phi_i|x) = 1$. One of the hypotheses of the OLS estimator recommends that the model must be linear in the parameters. In this case, a monotonic transformation makes it possible to obtain the estimates by that method. Applying a logarithmic transformation in expression (4), we have:

$$\ln W_i = x_i\beta + \ln \phi_i. \tag{5}$$

To see how heteroscedasticity biases the coefficients, assume that $\phi_i = \exp[(x_i\delta)\varepsilon_i]$, where ε_i is independent and normally distributed, $i.i.d$ and $\varepsilon_i \sim \mathcal{N}(\mu, \sigma^2)$. Thus, ϕ_i has a log-normal distribution whose variance is a function of x_i. That said, the expression (5) can be rewritten as:

$$\ln W_i = x_i\beta + (x_i\delta)\varepsilon_i. \tag{6}$$

Equation (6) is also known as the *location-scale model*. The explanatory variables x_i affect not only the location of the conditional distribution of $\ln W_i$, but also the scale and the dispersion through the term $x_i\delta$. Due to the presence of heteroscedasticity and properties of the lognormal distribution, we have that $E(\ln\phi_i|x_i) = -\frac{1}{2}\sigma^2_{\phi_i}$ when the model is identified, i.e., $E(\phi_i|x) = 1$. In the model expressed in (5), the conditional mean of $\ln W_i$ is given by:

$$E(\ln W_i|x_i) = x_i\left(\beta - \frac{\sigma^2_{\phi_i}}{2}\right) \neq x_i\beta. \tag{7}$$

Therefore, the presence of heteroscedasticity leads to inconsistency in log-linear regression if $E(\phi_i|x) = 1$ and the multiplicative model is identified. Finally, still according to Figueiredo et al. (2014), suppose that $E(\varepsilon_i|x) = E(\varepsilon_i) = 0$. By (6) it is possible to observe that the log-linear model is identified, since $E(\ln W_i|x_i) = x_i\beta + x_i\delta E(\varepsilon_i|x) = x_i\beta$. Nonetheless, according to the properties of the log-normal distribution, if $E(\varepsilon_i) = 0 \implies E(\phi_i|x_i) \neq 1$. Thus, if the log-linear model was identified, the multiplicative model is not identified in view of that $E(W_i|x) \neq \exp(x_i\beta)$. Thus, it can be concluded that the identification of the model expressed in (4) does not lead to log-linear identification (5), the opposite being also true.

In order to overcome this issue, the authors recommend the use of Quantile Regression, for, among other things, the equivalence property of this method ensures that if $g(\cdot)$ is a non-decreasing function, so for any random variable W, $Q_\tau(g(W)) = g(Q_\tau(W))$, where $\tau \in (0,1)$ are the quantiles of the regression.

The Quantile Regression method was developed by Koenker and Bassett Jr (1978). According to them, it is possible to estimate the effects of covariates along the distribution of the dependent variable, allowing exploration of different parts of the conditional distribution. In addition, the method has other desirable characteristics, among which the following stand out: (i) Robust to the presence of heteroscedasticity, (ii) flexibility regarding the distribution of the stochastic term, which may not be normal, (iii) allows controlling the existing heterogeneity in the sample. Furthermore, the use of this method allows characterizing the existence of positive (or negative) selection through the quantiles of the wage distribution. According to Sassen (1990) chances of negative selection occurring are greater in the lower quantile, $(\tau = 0.25)$, while at the higher end, $(\tau = 0.75)$, there is a greater probability of positive selection occurring.

In general, the estimators are obtained by minimizing the weighted sum of absolute deviations, which is formally written as:

$$\underset{\beta}{\arg\min} \quad \frac{1}{n}\left[\sum_{i:\ln w_i \geq X_i'\beta_\theta} \theta|\ln w_i - X_i'\beta| + \sum_{i:\ln w_i \leq X_i'\beta_\theta} (1-\theta)|\ln w_i - X_i'\beta_\theta|\right], \tag{8}$$

where $\ln W_i$ corresponds to the logarithm of the $i-th$ individual, X_i is a $K \times 1$ vector that contains the variables described in Table 1 and $\theta \in (0,1)$ corresponds to the sample quantile. Expression (8) can be rewritten as:

$$\underset{\beta}{\arg\min} \quad \frac{1}{n}\sum_{i=1}^{N}\rho_\theta(u_{i\theta}), \tag{9}$$

where $\rho_\theta(\cdot)$ is the Check function, defined as follows:

$$\rho_\theta(\cdot) = \begin{cases} \theta u_{\theta i}, & \text{if } u_{\theta i} \geq 0 \\ (\theta-1)u_{\theta i}, & \text{if } u_{\theta i} < 0 \end{cases} \tag{10}$$

The quantile regression coefficients are obtained through the partial derivative of the conditional quantile (θ) in relation to each explanatory variable. Thus, the variation in the $\theta - th$ conditional quantile given a change in exogenous variables.

In summary, in the first stage **equation** (2) is estimated in reduced form by OLS to obtain the exogenous variation of the variable (predicted values) education, which will be used in the second moment. Then, due to the issues already discussed, the second stage is estimated by OLS, but also by Quantile Regression. It is worth emphasizing the importance of analyzing the specification tests by Wu (1974) and Hausman (1978) reported at the end of the estimation, as it will indicate whether the instrument used is adequate and if it is really necessary to estimate the equation of wages through the instrumental variables model, since using IV when it is not accurate tends to produce inconsistent results.

2.2 *Data*

The data used in this study come from Demographic Censuses[11] for the years 1991, 2000 and 2010, produced by the Brazilian Institute of Geography and Statistics (IBGE). Although the information set is not recent, the analysis period is within a context of significant economic changes such as the slowdown in the world economy, the tightening of migration policies by developed countries, Brazil's economic stabilization plan, integration into the global economy due to the process of commercial opening, among others. These events not only accentuated the circulation of goods, but also increased the flow of foreigners in the national territory, justifying then the analysis of the mentioned years, which is still little explored in the literature (Vilela, 2011).

The Demographic Census has a relevant mapping of the living conditions of international immigrants; however, it should be noted that these records can be underestimated. This is because there is a non-negligible portion of immigrants who enter/reside in the country illegally, which tends to distort reality in official records. What is known, for example, is that the number of Latin Americans is higher than those presented in

[11] The Demographic Census is one of the most important surveys on the living conditions of the Brazilian population at an individual level. Information is collected through the application of questionnaires that include a survey of the characteristics of residents and households. In addition to having national coverage, the survey is carried out periodically every ten years (decade).

the official statistics Silva (2006). For this reason, the analyzes carried out in this work must be interpreted with reservations, considering that the data may not be consistent with reality.

For the present study, international immigrants (foreigners) are those individuals who were born in another country or became naturalized Brazilian.[12] In the model, a dummy variable is inserted that measures the existence of a wage gap between natives and foreigners. If it is positive (and significant), it indicates that immigrants are positively selected, that is, their wages are superior to those of the natives. As discussed earlier, Chiswick (1978) argues that this salary premium would be explained in terms of unobserved characteristics such as commitment, motivation, skills, among others. Otherwise, if negative (and significant), it indicates the existence of negative selection, that is, foreigners tend to receive lower salaries than the natives as pointed out by Borjas (1987).

Analogously to Vilela (2011), the natives were separated into two groups,[13] those who did not migrate and those who carried out migratory movements internally.[14] The reason for that begins with the assumption that these individuals contain characteristics that differentiate them from non-migrants. In fact, not considering this aspect may not isolate the impact of international migration on wages from the effects of internal migration Thus, the estimates made compare immigrants from another nationality with Brazilians who have not migrated and with individuals who migrated internally, so that the wage effect of foreigners may be properly isolated.

Due to methodological changes in the 2010 Census, the education variable had to be grouped into four levels of instruction, namely: 1 - No education and incomplete Elementary; 2 - Complete elementary school and incomplete high school; 3 - Complete High School and Incomplete Higher Education; 4 - Complete Higher Education. In other years (Census 1991 and 2000), it was measured in years of study. It should be noted that this is a limitation relevant to the data and, therefore, beyond the scope of the researcher.

According to Borjas (1987), pensioners, members of the armed forces, civil servants, unpaid workers, residents in collective rooms and/or accommodation and individuals aged[15] under 20 and over 60 are excluded from the sample. In addition to those, individuals who were self-employed were also discarded, that is, independent workers. Finally, the variables used in this work are reported in Table 1.

[12] In the questionnaire of the Demographic Census applied by the IBGE, three categories are identified: (i) Native Brazilian; (ii) Naturalized Brazilian; (iii) Foreigner.

[13] Individuals residing in the Federation Unit (State) of birth were considered non-migrants. And individuals who resided in a State other than their place of birth was considered a migrant.

[14] For the State of Minas Gerais, Gama and Hermeto (2017) find that internal migrants earn salary benefits when deciding to migrate.

[15] Borjas (1987) and Vilela (2011) propose to use a minimum age of 25 years. The authors argue that these individuals are more likely to have completed their studies and, therefore, to be inserted in the labor market, however, by making this cut, a significant portion of immigrants is lost. For this reason, we chose to use the minimum age of 20 years.

Table 1. Description of used variables

Variables	Description	Format
Wages	Main work's wages	Continuous
Schooling	Measurement of years of schooling	Discreet
Foreign	Takes on value 1 if an individual is foreign/naturalized Brazilian, 0 otherwise.	Dummy
Age	Refere-se a idade do indivíduo	Discreet
Control Variables		
Formal employment	Takes on value 1 if an individual has a formal employment contract, 0 otherwise	Dummy
Worked hours	Number of hours worked weekly	Discreet
White	Takes on value 1 if an individual is white skinned, 0 otherwise	Dummy
Male	Takes on value 1 if an individual is a man, 0 otherwise.	Dummy
Catholic	Takes on value 1 if an individual is a Catholic, 0 otherwise	Dummy
Number of people	Number of individuals who live in a household	Discreet
Home condition	Takes on value 1 if an individual is the Boss, 0 otherwise	Dummy

2.3 *Descriptive data analysis*

Table 2 shows the descriptive statistics of the data used.[16] In the first place, it is observed that the average salary of the sample increased from $1{,}964.054$, in 1991, to $2{,}562.524$ in 2000, and reduced in 2010 to $2{,}395.281$, which represents an increase of approximately 22% over two decades. Despite this, the standard deviation in the year 2000 is almost 3 times higher than in 1991, which suggests a greater wage dispersion (greater inequality).

The average of years of schooling also increased, increasing from 5.36, in 1991, to 6.82 in 2000, a variation equivalent to 27.1%. Unlike salaries, the educational dispersion practically did not present significant variations, remaining in around 4 standard deviation. We also highlight the increase in the average number of immigrants who entered the country over the decade whose average variation was approximately 62.5%, indicating a significant increase in foreign flows in the Brazilian territory.

Table 3 contains the average salary of non-migrant natives, migrants and foreigners by gender and for the Censuses. As can be seen, for all intervals the average salary of foreigners is higher than that of non-migrant natives, and the disparity is greater in the upper ranges for men in the year 2000. In the range between the quantiles 75-90, the average salary of foreigners is almost 6.6 times higher than the average salary of natives, while the average salary of foreign women is approximately 4 times higher than the average of native women's salaries, which is lower when compared to men for that year. Furthermore, wage inequality among women had a relevant reduction over the decade in all ranges. Comparing foreigners and native migrants, it is also possible to infer that there is a salary difference between the groups, however this difference is slightly smaller when compared to non-migrant natives, especially to those situated in the upper ranges.

This evidence seems to confirm the existence of positive selection, in accordance with results by Chiswick (1978). However, an important aspect should be highlighted: along the decades there has been a reduction in the wage gap between foreigners and migrant and non-migrant natives for both genders, especially in the lower ranges 15-30, 30-45. In a preliminary analysis, it is possible to emphasize that part of the reduction

[16] To keep the text display more organized, the results of the control variables were omitted.

Table 2.: Descriptive Statistics – 1991, 2000 and 2010 Census

	1991 Census			2000 Census			2010 Census		
Variables	N	mean	std.dev.	N	mean	std.dev.	N	mean	std.dev.
Adjusted salary	1,440,015	1,964.05	3,679.51	276,391	2,562.52	9,534.54	330,669	2,330.31	5,283.84
Log salary	1,440,015	7.08	0.91	276,391	7.27	0.89	330,669	7.34	0.79
Schooling	1,440,015	5.36	4.43	276,391	6.82	4.38	330,669	2.13	1.07
Foreign	1,440,015	0.01	0.09	276,391	0.01	0.11	330,669	0.01	0.07
Age	1,440,015	34.02	11.27	276,391	30.89	9.55	330,669	33.60	9.68
Formal employment	1,440,015	0.68	0.47	276,391	0.51	0.50	330,669	0.69	0.46
Worked hours	1,440,015	45.26	10.76	276,391	46.66	13.65	330,669	43.16	13.82
White	1,440,015	0.53	0.50	276,391	0.53	0.50	330,669	0.46	0.50
Male	1,440,015	0.66	0.47	276,391	0.63	0.48	330,669	0.62	0.48
Catholic	1,440,015	0.83	0.38	276,391	0.72	0.45	330,669	0.62	0.49
Number of people	1,440,015	4.37	1.99	276,391	4.15	2.05	329,879	3.40	1.63
Condition in the family	1,440,015	0.55	0.50	276,391	0.52	0.50	330,669	0.51	0.50

Table 3.: Average wages of non-migrant natives, migrants, and foreigners by gender - 1991, 2000 and 2010 Censuses

Average wages: Non-migrant Natives vs. Foreigners												
	1991				2000				2010			
	Natives		Foreigners		Natives		Foreigners		Natives		Foreigners	
Quantile Range	Man	Woman	Man	Woman	Man	Woman	Man	Woman	Man	Woman	Man	Woman
15-30	605.2	435.0	2,000.6	1,131.0	908.5	632.6	1,808.0	893.8	1,033.7	890.2	1,297.4	906.7
30-45	887.7	585.6	3,384.6	1,815.4	1,188.5	844.0	3,537.5	1,463.5	1,335.8	946.9	1,989.7	1,124.8
45-60	1,253.4	814.9	5,190.6	2,699.9	1,475.2	1,093.1	6,457.2	2,428.5	1,621.8	1,140.7	3,388.6	1,605.8
60-75	1,804.2	1,106.8	8,415.9	4,228.8	1,857.0	1,325.3	10,703.6	4,411.4	2,037.4	1,418.3	6,336.7	2,771.0
75-90	3,000.0	1,735.6	14,627.1	7,272.2	2,889.9	1,953.0	19,213.9	8,418.1	3,249.4	2,206.0	12,123.5	5,969.2
Average wages: Migrant Natives vs. Foreigners												
	1991				2000				2010			
	Natives		Foreigners		Natives		Foreigners		Natives		Foreigners	
Quantile Range	Man	Woman	Man	Woman	Man	Woman	Man	Woman	Man	Woman	Man	Woman
15-30	605.2	435.0	2,000.6	1,131.0	908.5	632.6	1,808.0	893.8	1,033.7	890.2	1,297.4	906.7
30-45	887.7	585.6	3,384.6	1,815.4	1,188.5	844.0	3,537.5	1,463.5	1,335.8	946.9	1,989.7	1,124.8
45-60	1,253.4	814.9	5,190.6	2,699.9	1,475.2	1,093.1	6,457.2	2,428.5	1,621.8	1,140.7	3,388.6	1,605.8
60-75	1,804.2	1,106.8	8,415.9	4,228.8	1,857.0	1,325.3	10,703.6	4,411.4	2,037.4	1,418.3	6,336.7	2,771.0
75-90	3,000.0	1,735.6	14,627.1	7,272.2	2,889.9	1,953.0	19,213.9	8,418.1	3,249.4	2,206.0	12,123.5	5,969.2
15-30	790.0	531.0	2,000.6	1,131.0	942.9	615.5	1,808.0	893.8	1,060.7	871.6	1,297.4	906.7
30-45	1,113.8	778.9	3,384.6	1,815.4	1,246.7	830.1	3,537.5	1,463.5	1,371.6	940.5	1,989.7	1,124.8
45-60	1,534.2	986.3	5,190.6	2,699.9	1,622.4	1,101.8	6,457.2	2,428.5	1,712.1	1,145.7	3,388.6	1,605.8
60-75	2,102.9	1,339.5	8,415.9	4,228.8	2,267.4	1,424.9	10,703.6	4,411.4	2,285.6	1,474.2	6,336.7	2,771.0
75-90	3,292.0	1,963.8	14,627.1	7,272.2	4,499.4	256,661.0	19,213.9	8,418.1	3,748.5	2,379.8	12,123.5	5,969.2

in the differential is explained by the growth of the average salary of non-migrant and migrant native men and women, and another part results from a strong reduction in the average wages of foreigners, mainly between 2000 and 2010.

The information presented also suggests the existence of intragroup gender inequality, being more pronounced among foreigners for the year 2000. In the range 45-60, for example, men have an average salary approximately 2.65 higher than the average salary for women. Among non-migrant and migrant natives, differences by gender remain constant over the period and at all ranges.

Since schooling is an important determinant of wages, Table 4 shows the school-average age of non-migrants, migrants, and foreigners for both sexes. It is evident that on average foreigners are more educated than non-migrant and migrant natives, which may justify part of the wage gap between the groups. Considering intra-group differences for non-migrant natives and foreigners in 1991, it appears that the average schooling of women is superior to that of men, or the difference is very small and yet they receive a lower average remuneration, thus reinforcing the prevalence of the discriminatory component by gender. The same result can be seen in the year 2000 for men and non-migrant native women, except among foreigners whose average schooling of men the schooling of women is higher. A similar result is observed when comparing native foreign migrants.

In order to explore the data visually, Figure 1 shows the wage densities of both groups segmented by Gender. Figures 1a and 1c make the inequality between the wage distributions of foreign and native men for both years visibly clear, in a way that foreigners obtain superior earnings in practically all deciles of the distribution. Among foreigners, there is a strong increase in inequality in intra-group distribution in 2000, especially in occupations above the median indicated by a higher dispersion, while the differences between natives show a slight reduction in salary dispersion. For natives, the greatest weight in the lower tail of distributions in 1991 and 2000, suggest a greater participation of these individuals in occupations with lower remuneration. However, over the decade it is possible to identify a reduction in the wage gap between the groups in the upper tail.

Among women, the wage distribution of foreign women is shifted to the right when compared to that of native women and exhibit a more symmetrical pattern in both years, as can be seen in Figures 1b and 1d. Furthermore, it is possible to visualize an increase in the intra-group difference in that foreign women present a greater wage dispersion in the year 2000. In addition, there is a major reduction in wage inequalities between groups at the lower end, so that in the lower deciles there seem to be negative selection[17] traits for foreign women. Analyzing the upper end for the year 2000, the greatest weight of native women in this tail suggest a more accentuated participation in higher occupations, while the participation of foreign women in lower occupations increases significantly between years.

Analyzing Figures 1e and 1f (2010 Census) it is possible to observe a significant reduction in wage differences between foreigners and natives, especially for women.

[17] According to Borjas (1987), negative selection consists of the lower income that immigrants receive in relation to natives in similar occupations.

Table 4.: Average schooling of non-migrant natives, migrants, and foreigners by gender - 1991, 2000 and 2010 Censuses

Average schooling: Non-migrant Natives vs. Foreigners												
	1991				2000				2010			
	Natives		Foreigners		Natives		Foreigners		Natives		Foreigners	
Quantile Range	Man	Woman	Man	Woman	Man	Woman	Man	Woman	Man	Woman	Man	Woman
15-30	2.03	2.70	5.00	5.61	2.91	3.87	9.88	7.83	1.00	1.00	2.05	1.52
30-45	3.71	4.00	8.96	9.61	4.15	5.55	11.31	10.97	1.00	1.85	3.00	2.61
45-60	4.19	4.96	11.00	11.00	5.73	7.70	14.37	12.55	1.97	2.68	3.17	3.00
60-75	6.47	7.77	13.57	12.66	7.88	10.06	15.32	14.93	2.71	3.00	4.00	3.80
75-90	9.41	10.65	15.33	15.00	10.56	11.02	16.41	16.01	3.00	3.38	4.00	4.00
Average schooling: Migrant Natives vs. Foreigners												
	1991				2000				2010			
	Natives		Foreigners		Natives		Foreigners		Natives		Foreigners	
Quantile Range	Man	Woman	Man	Woman	Man	Woman	Man	Woman	Man	Woman	Man	Woman
15-30	1.31	2.13	5.00	5.61	3.02	3.84	9.88	7.83	1.00	1.00	2.05	1.52
30-45	3.41	3.80	8.96	9.61	4.26	5.73	11.31	10.97	1.00	1.70	3.00	2.61
45-60	4.00	4.47	11.00	11.00	6.26	8.02	14.37	12.55	1.90	2.55	3.17	3.00
60-75	5.66	7.22	13.57	12.66	8.74	10.59	15.32	14.93	2.71	3.00	4.00	3.80
75-90	8.86	9.97	15.33	15.00	11.21	11.71	16.41	16.01	3.03	3.43	4.00	4.00

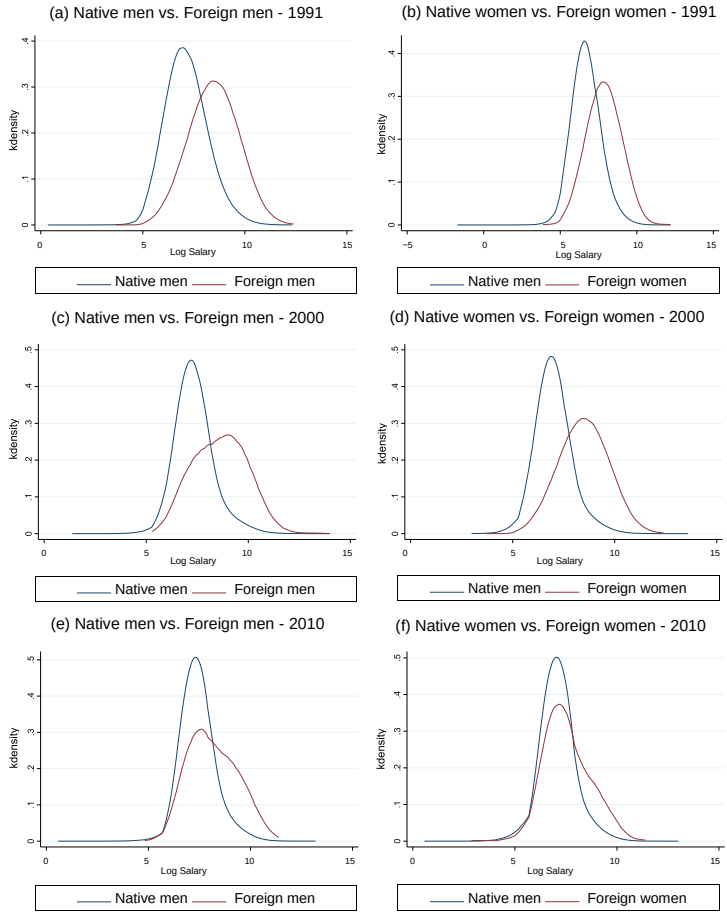

Figure 1. Wage density for 1991, 2000 and 2010 by gender

Over the decades, the tails of the immigrants' wage distribution become denser at the lower extremities, which may reflect an increase in less-skilled foreigners.

In summary, densities preliminarily indicate the possible existence of positive selection in relation to international immigrants for both sexes, except in the lower deciles in Figure 1d. Furthermore, it is worth emphasizing the prevalence of a significant heterogeneous component, which justifies the use of quantile regression, since the results for the mean do not capture this effect. Therefore, this analysis presents some relevant insights, however it is necessary to investigate other characteristics that affect the wage distribution of each group.

3. Results and discussion

Initially, it is advisable to evaluate the specification tests listed at the end of Tables 5, 6 and 7. For testing weak instruments, the p-value was significant in all years and specifications, which implies rejection of the null hypothesis and indicates that the instrument

Table 5. Instrumental Variable and OLS Results - 1991 Census

	Dependent variable: Log salary					
	Complete sample		Non-migrants vs. Foreigners		Migrants vs. Foreigners	
	(1) - OLS	(2) - IV	(3) - OLS	(4) - IV	(5) - OLS	(6) - IV
Foreigner	0.549***	0.247***	0.556***	0.291***	0.539***	0.190***
	(0.006)	(0.010)	(0.007)	(0.011)	(0.007)	(0.014)
Schooling	0.080***	0.142***	0.084***	0.143***	0.076***	0.144***
	(0.0001)	(0.001)	(0.0002)	(0.002)	(0.0002)	(0.002)
Age	0.062***	0.062***	0.060***	0.060***	0.063***	0.064***
	(0.0003)	(0.0003)	(0.0004)	(0.0005)	(0.0005)	(0.001)
Age²	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.00000)	(0.00000)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Constant	4.938***	4.433***	5.134***	4.605***	4.970***	4.397***
	(0.008)	(0.014)	(0.066)	(0.072)	(0.011)	(0.023)
Number of obs.	1,440,015	1,440,015	815,211	815,211	636,145	636,145
Control variables	yes	yes	yes	yes	yes	yes
R^2	0.442	0.365	0.474	0.409	0.405	0.305
RMSE	0.680	0.726	0.679	0.720	0.683	0.738
Weak instruments		p < 0.00		p < 0.00		p < 0.00
Hausman test		p < 0.00		p < 0.00		p < 0.00

Notes: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

used is robust. Additionally, the specification test Hausman (1978) also presented a significant p-value, which suggests the rejection of the null hypothesis and points out that the results estimated by Instrumental variables are consistent.

The results for the 1991 census are reported in Table 5. Columns (1) and (2) correspond the estimates for the complete sample, (3) and (4) compare non-migrant individuals with foreigners and columns (5) and (6) compare national migrants with international migrants. In all specifications, it is inferred that foreigners constitute a positively selected group in the Brazilian labor market.

As discussed in section 2.1, the estimates obtained via OLS tend to be biased and inconsistent when endogeneity is present. As can be seen, for all years and specifications, the results indicate that the OLS estimator overestimates the coefficient of the variable of interest (Foreigner). Considering the complete sample, for example, the coefficient estimated by OLS is 54.9%, however, after treating the problem, the result obtained by instrumental variable is approximately twice smaller, 24.7%. For this reason, the analysis of the results will be made based on the estimates obtained by instrumental variables.

When comparing non-migrants and foreigners, the results indicate that immigrants earn on average 29.1% more than non-migrants. This result can be explained by two perspectives: first, Bonacich (1973) argues that foreigners are inserted in occupations within a market to the dual[18] one, which, in a way, protects them from the competition with the general market of the host society. For that reason, they tend to occupy intermediate or superior positions to the natives. Second, Chiswick (1978) explains that after a period of adaptation, foreigners have a higher wage growth rate than the salary rate of the natives, as they are more skillful, motivated and committed.

[18] According to Piore and Safford (2007), the labor market is divided into privileged occupations (primary employment sector) and discontinuous occupations (secondary employment sector). In the first sector, jobs are stable, well paid and with the possibility of social ascension. In the second sector, jobs are disqualified, with low and unstable pay.

Table 6. Instrumental Variable and OLS Results - 2000 Census

	Dependent variable: Log salary					
	Complete sample		Non-migrants vs. Foreigners		Migrants vs. Foreigners	
	(1) - OLS	(2) - IV	(3) - OLS	(4) - IV	(5) - OLS	(6) - IV
Foreigner	0.395***	0.147***	0.456***	0.168***	0.280***	0.113***
	(0.011)	(0.016)	(0.011)	(0.017)	(0.013)	(0.023)
Schooling	0.105***	0.151***	0.098***	0.150***	0.118***	0.152***
	(0.0003)	(0.002)	(0.0004)	(0.002)	(0.001)	(0.004)
Age	0.074***	0.069***	0.073***	0.068***	0.077***	0.072***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Age²	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00002)	(0.00002)
Constant	4.418***	4.077***	4.492***	4.106***	4.274***	4.026***
	(0.016)	(0.021)	(0.018)	(0.025)	(0.033)	(0.043)
Number of obs.	276,391	276,391	202,804	202,804	77,095	77,095
Control variables	yes	yes	yes	yes	yes	yes
R^2	0.447	0.403	0.438	0.379	0.495	0.474
RMSE	0.658	0.684	0.634	0.666	0.726	0.740
Weak instruments		p < 0.00		p < 0.00		p < 0.00
Hausman test		p < 0.00		*p* < 0.00		p < 0.00

Notes: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Comparing internal and foreign migrants, the results reported in column (6) suggest that there is still a positive wage gap in favor of foreigners, (19%), but the difference between them is relatively smaller when compared to non-migrants, which corroborates the hypothesis that internal migrants have similar characteristics to foreigners Vilela (2011).

Returns to schooling, in turn, showed an expected sign. Again, it's possible to observe that the OLS estimator is biased when the exogeneity hypothesis is violated. For the education variable, the effects are underestimated. Therefore, the results found suggest that for each additional year of study, ceteris paribus, the salary increases by average 14.3%. While not directly comparable research, evidence from Harmon and Walker (1995) and Rahman and Al-Hasan (2018) corroborate the findings of this research.

The results for the year 2000 are explored in Table 6. As provided in column (4), it appears that foreigners are positively selected, receiving an average salary 16.1% higher than non-migrant natives. When comparing the estimates in Table 5 with those in Table 6, it appears that the wage differences between these groups had a significant reduction over the decade, around 42.3%. In addition, the salary difference between internal and foreign migrants also reduced the equivalent of 40.53%. Such results ratify the analyzes previously made. In relation to educational returns, the coefficients practically did not change.

Estimates for the year 2010 are shown in Table 7. As can be observed in columns (4) and (6), the wage gap between the groups persists, with a small reduction between migrants and foreigners when compared to the year 2000. Educational returns, in turn, showed expected signs. It is important to highlight that this variable was aggregated in four levels due to IBGE's methodological change. In this case the marginal effect does not refer to an extra year of study, but rather to a change in the education category.

Despite having addressed the problem of endogeneity, there is still an important bias in log-linearized models, since the conditional mean is not invariant to monotonic

Table 7. Instrumental Variable and OLS Results - 2010 Census

	Dependent variable: Log salary					
	Complete sample		Non-migrants vs. Foreigners		Migrants vs. Foreigners	
	(1) - OLS	(2) - IV	(3) - OLS	(4) - IV	(5) - OLS	(6) - IV
Foreigner	0.233***	0.151***	0.261***	0.175***	0.196***	0.109***
	(0.015)	(0.016)	(0.015)	(0.017)	(0.016)	(0.018)
Schooling	0.356***	0.464***	0.333***	0.442***	0.389***	0.505***
	(0.001)	(0.007)	(0.001)	(0.008)	(0.002)	(0.011)
Age	0.060***	0.060***	0.057***	0.057***	0.063***	0.064***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age²	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00002)	(0.00002)
Constant	4.723***	4.426***	4.843***	4.553***	4.598***	4.245***
	(0.016)	(0.025)	(0.020)	(0.030)	(0.029)	(0.044)
Number of obs.	330,669	329,879	196,284	196,284	135,240	135,240
Control variables	yes	yes	yes	yes	yes	yes
R^2	0.386	0.367	0.372	0.351	0.410	0.390
RMSE	0.618	0.628	0.597	0.607	0.649	0.659
Weak instruments		p < 0.00		p < 0.00		p < 0.00
Hausman test		p < 0.00		p < 0.00		p < 0.00

Notes: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figueiredo et al. (2014). In order to overcome this issue, we carried out new inferences using the Quantile Regression method. Also, given the existence of wage discrimination by gender, estimates were obtained stratified by gender.

The results estimated by Quantile Regression are shown in Figures 2, 3 and 4 for the years 1991, 2000 and 2010, respectively. The solid red line indicates the coefficient estimated by OLS and the dashed lines corresponds to the confidence interval for a level of 5% significance. Meanwhile, the estimated points denote the estimates by quantile and gray shaded area refers to the confidence interval.

In general, in all years and specifications, the existence of heterogeneity between individuals is observed, which again justifies the use of Quantile Regression. Analyzing the estimates for the year 1991, it is possible to observe a similar behavior of the wage gap for men and women. Both have an inverted U-shape, being more evident among foreigners and internal migrants. Additionally, the evidence of positive selection, which becomes greater in intermediate occupations.

For the year 2000, estimates report a similar pattern, however the existence of negative selection for women when comparing foreigners and non-migrants, Figure 3b. Considering foreigners and internal migrants, Figure 3d shows that in the quantile lower there are no wage differentials for women. In this case, the reductions in differences mentioned above are related to low-skilled occupations. Particularly, this decade is important for the Brazilian economy due to the economic stabilization plan and, above all, trade liberalization. The Heckscher-Ohlin-Samuelson (HOS) model advocates that the opening of trade causes a change in relative prices, favoring industries intensive in low-skilled work. These price variations in addition to changing the composition of employment, provide a reduction in wage inequalities (effect denoted such as Stolper-Samuelson), which corroborates the evidence of this research.

Looking at the results for the 2010 census, it appears that the inverted-U shape practically disfigures itself. It is possible to identify that the wage gap for the comparison groups substantially reduced in low-skilled occupations. Such evidence reinforces

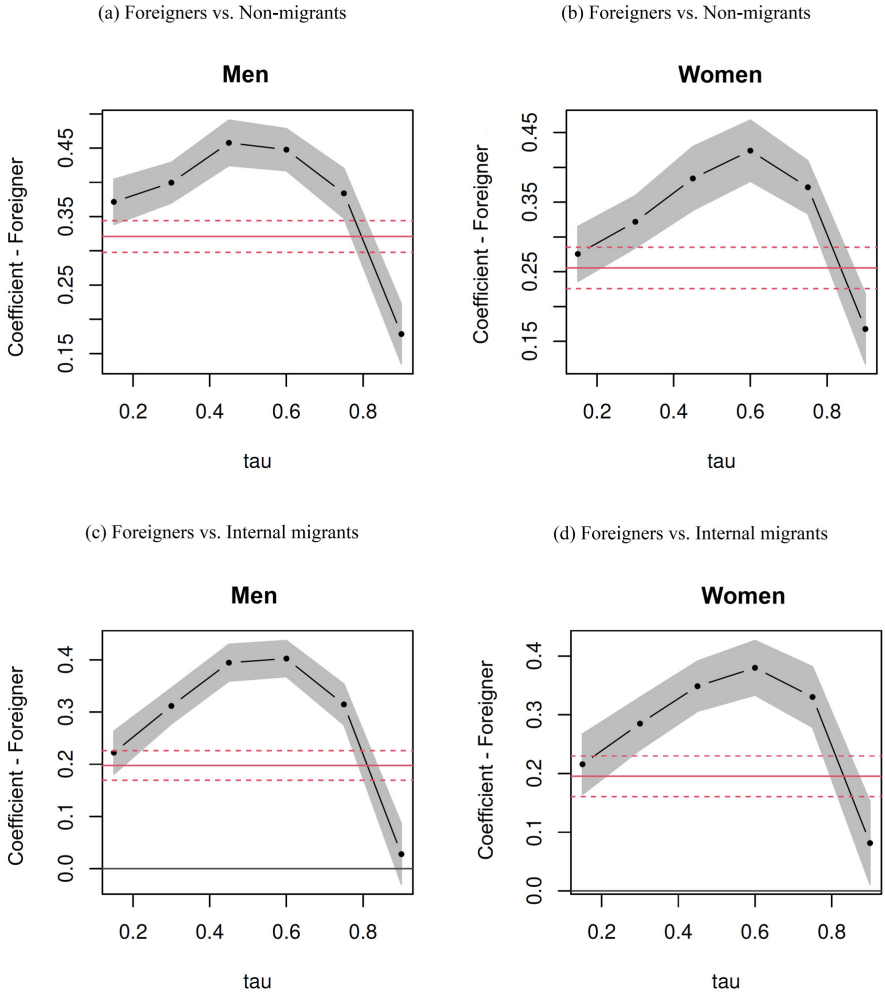

Figure 2. Quantile Regression results by gender - 1991 Census

the idea of increasing low-skilled immigration, especially Haitians and Venezuelans for example. Furthermore, it is possible to identify the replacement of native by foreign labor in the lower quantiles, whose salaries paid are relatively cheaper. At the upper end of the conditional distribution, the salary differences for men and women still persist, considering that the most qualified foreign workforce is certainly more productive when compared to Brazilian skilled labor.

In short, the results obtained in 1991 indicate that foreigners constitute a positively selected group. However, due to the economic stabilization and opening up plan from 2000 onwards, a reduction in the wage gap is observed, mainly for women and in low-skilled occupations. In 2010, evidence of replacement of native labor by foreigners becomes stronger, mainly due to increased migration from underdeveloped countries. It should be noted that the results presented must be viewed with caution, as there is a significant portion of immigrants residing illegally in the country and accessing the labor

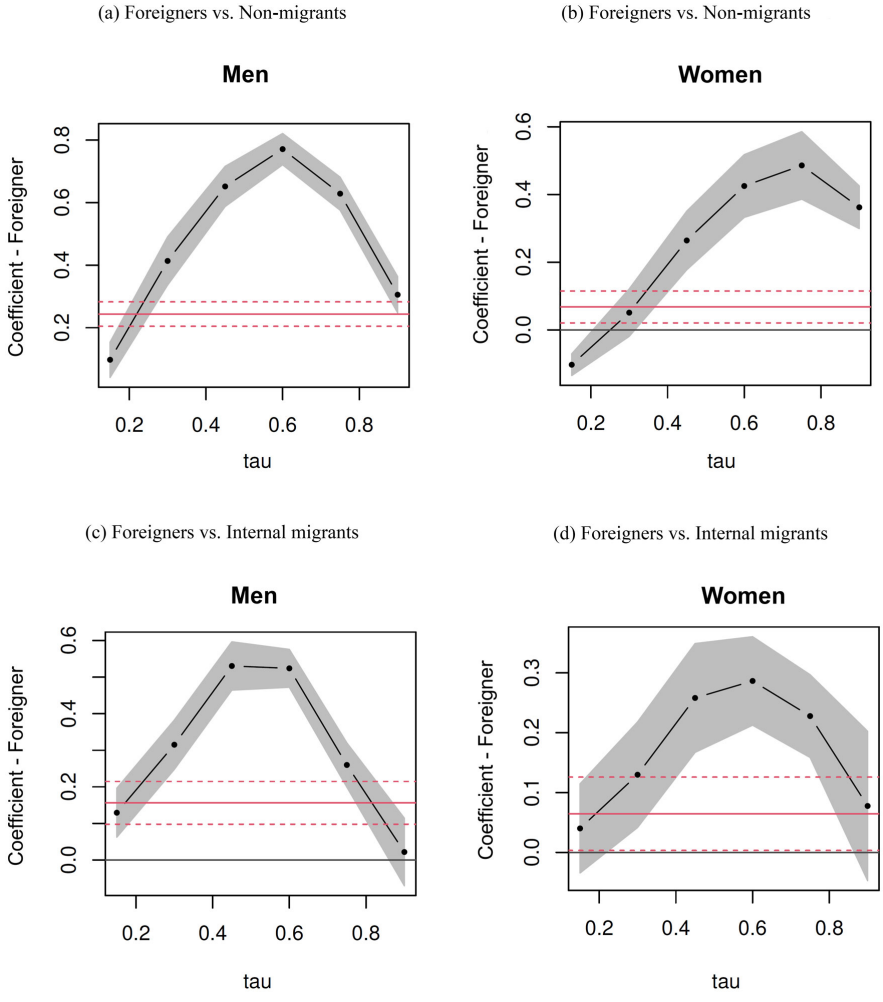

Figure 3. Quantile Regression results by gender - 2000 Census

market through informality, which makes it difficult to obtain more realistic information even when controlling for that sector.

4. Conclusions

This work contributes to the economic literature on migratory flows by deepening the investigation into the wage gap between foreigners and natives in the labor market, since much of the literature has been dedicated to studying the internal migratory movements. The difference between this study and previous research consists of using census data, which provides a large amount of information on immigration and the use of a robust strategy, the identification of linearized log models.

Considering the findings for the conditional mean, we found that the OLS estimator overestimates the coefficient that captures the wage gap between foreigners and na-

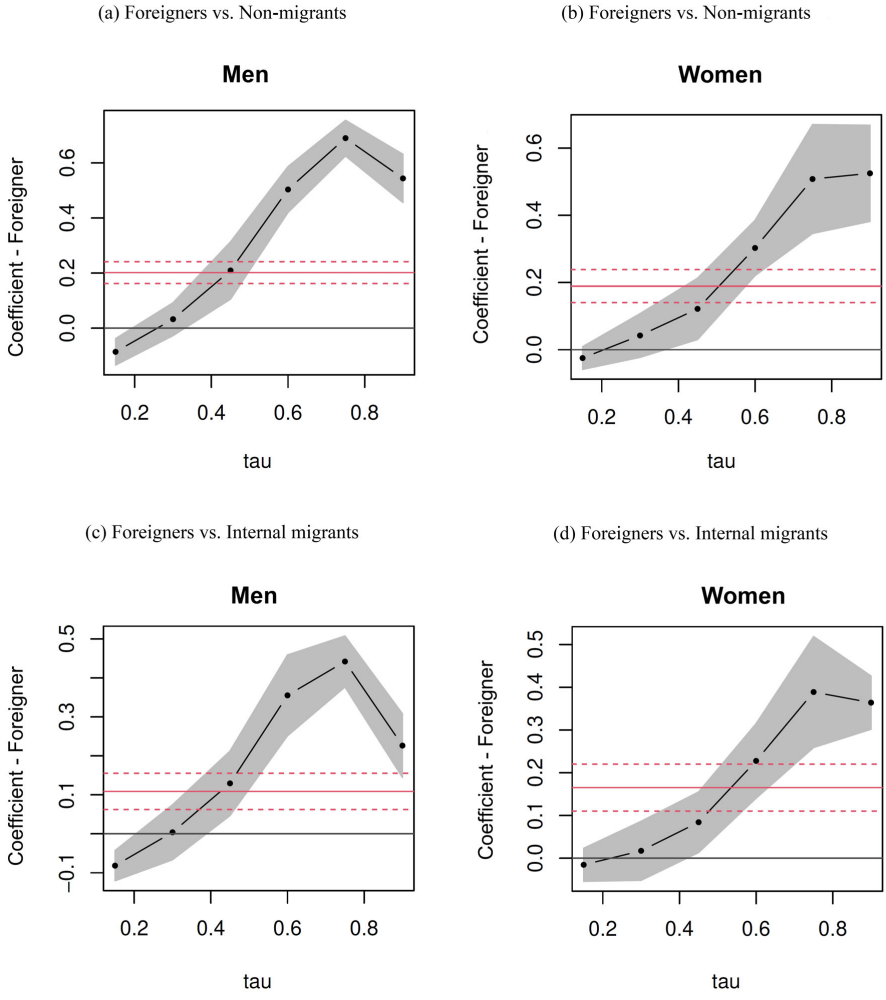

Figure 4. Quantile Regression results by gender - 2010 Census

tives due to presence of endogeneity. When treated, the wage gap reduces by approximately 50%, indicating a quite deep bias. Additionally, for all years and specifications the estimates by instrumental variable suggest that immigrants constitute a positively selected group, the difference being reduced over the decades.

However, the evidences for the mean are not very informative, as they ignore the existence of heterogeneity in the sample, and are still subject to the identification bias of multiplicative models. In order to overcome this issue, new inferences were made through the Quantile Regression method. Our results indicate that for the years 1991 and 2000, the wage gap between foreigners and natives has an inverted U-shape, in a way that the greater differential is seen in intermediate occupations. However, throughout the decade there is a reduction in wage differences in low-skilled occupations, especially for women. Such findings are supported by the predictions of the Heckscher-Ohlin-Samuelson model (HOS), in which trade opening causes changes in relative prices fa-

voring factor-intensive industries. When analyzing the year 2010, the inverted U-shape is disfigured and wage differences decrease or become negative for both genders in low-skilled occupations. These results suggest an increase in immigration of foreigners from underdeveloped countries (Haitians, Venezuelans, etc.) whose salaries are relatively cheaper, which can lead to the replacement of the native workforce by foreigners.

Finally, it is noteworthy that the findings of this work should be interpreted with reservations, as there is a significant portion of immigrants who reside in the country illegally and who have access to the Brazilian labor market through informality, which makes it difficult to achieve an undistorted reality. However, this is a limitation of the available data. In relation to the empirical strategy, a considerable effort was made to ensure that the evidence presented were robust to the problems of model identification.

Funding We are grateful for CNPQ(Conselho Nacional de Desenvolvimento Científico e Tecnológico) financial support for research project number $433485/2018-3$.

Availability of data and materials The datasets and codes used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests None.

Bibliography

Becker, Gary S. and H. Gregg Lewis (1973): "On the interaction between the quantity and quality of children," *Journal of Political Economy*, 81 (2, Part 2), S279–S288. [4, 6]

Bonacich, Edna (1973): "A theory of middleman minorities," *American Sociological Review*, 38, 583–594. [16]

Borjas, George J. (1987): "Self-selection and the earnings of immigrants," Tech. Rep., National Bureau of Economic Research. [2, 3, 9, 13]

——— (1990): "Self-selection and the earnings of immigrants: Reply," *The American Economic Review*, 80 (1), 305–308. [3]

——— (1995): "Assimilation and changes in cohort quality revisited: What happened to immigrant earnings in the 1980s?" *Journal of Labor Economics*, 13 (2), 201–245. [3]

Brito, Fausto (1995): "Ensaio sobre as migrações internacionais no desenvolvimento do capitalismo," *Revista Brasileira de Estudos de População*, 12 (1/2), 21–34. [1]

Carliner, Geoffrey (1980): "Wages, earnings and hours of first, second, and third generation American males," *Economic Inquiry*, 18 (1), 87–102. [2]

Castles, Stephen (1998): *The age of migration: International population movements in the modern world*, Macmillan International Higher Education. [2]

Chiquiar, Daniel and Gordon H. Hanson (2005): "International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 113 (2), 239–281. [3]

Chiswick, Barry R. (1978): "The effect of Americanization on the earnings of foreign-born men," *Journal of Political Economy*, 86 (5), 897–921. [2, 3, 9, 10, 16]

Croix, David and Matthias Doepke (2003): "Inequality and growth: Why differential fertility matters," *American Economic Review*, 93 (4), 1091–1113. [4, 6]

de Cerqueira Rondon, Igor Raoni, Roney Fraga Souza, and Sheila Cristina Ferreira Leite (2013): "Migração e seleção: Um estudo sobre os migrantes interestaduais brasileiros para os anos de 2000 e 2010," *Revista Debate Econômico*, 1 (2), 5–27. [3]

Fernandes, Lendina de Castro (2019): "Diferença salarial entre imigrantes e nativos: Evidência empírica em Portugal," Master's thesis, Universidade do Minho, Braga, Portugal. [5]

FGV (2020): "A economia de Roraima e o fluxo venezuelano," Tech. Rep., Diretoria de Análise de Políticas Públicas da Fundação Getulio Vargas. [5]

Figueiredo, Erik, Luiz Renato Lima, and Georg Schaur (2014): "Robust estimation of gravity equations and the WTO impact on trade inequality," in *CESifo Conference on Estimation of Gravity Model of Bilateral Trade*. [3, 4, 6, 7, 18]

Freitas, Gregory (1981): "The earnings of immigrants in the American labor market," in *Fourth Annual Meeting*, 58. [2]

Gama, Luiz Carlos Day and Ana Maria Hermeto (2017): "Diferencial de ganhos entre migrantes e não migrantes em Minas Gerais," *Revista Brasileira de Estudos de População*, 34 (2), 341–366. [9]

Gheasi, Masood, Peter Nijkamp, and Piet Rietveld (2017): "Wage gaps between native and migrant graduates of higher education institutions in the Netherlands," *Letters in Spatial and Resource Sciences*, 10 (3), 277–296. [5]

Harmon, Colm and Ian Walker (1995): "Estimates of the economic return to schooling for the United States," *The American Economic Review*, 85 (5), 1278–1286. [5, 17]

Hausman, Jerry A. (1978): "Specification tests in econometrics," *Econometrica*, 46 (6), 1251–1271. [8, 16]

Heckman, James J., Lance J. Lochner, and Petra E. Todd (2006): "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond," *Handbook of the Economics of Education*, 1, 307–458. [3, 5]

Koenker, Roger and Gilbert Bassett Jr (1978): "Regression quantiles," *Econometrica*, 46 (1), 33–50. [4, 7]

Mariano, Francisca Zilania and Ronaldo de Albuquerque (2011): "Endogeneidade da educação na previsão da taxa de retorno: Avaliação metodológica e aplicação para regiões brasileiras e estados selecionados," *Relatório*, (7). [5]

Mincer, Jacob A. (1974): "Schooling and earnings," in *Schooling, Experience, and Earnings*, NBER, 41–63. [3, 4, 5]

Petit, Juan Miguel (2002): “Migraciones, vulnerabilidad y políticas públicas,” Tech. Rep., Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL). [2]

Piore, Michael J. and Sean Safford (2007): “Preliminary thoughts on identity and segmentation in primary sector labor markets,” *Economies et sociétés*, 41 (6), 925. [16]

Rahman, Mustafizur and Md Al-Hasan (2018): “Returns to schooling in Bangladesh revisited: An instrumental variable quantile regression approach,” *Development Studies*, 41 (2), 27–42. [17]

Sales, Teresa (1991): “Novos fluxos migratórios da população brasileira,” *Revista Brasileira de Estudos de População*, 8 (1/2), 21–32. [2]

Santos Junior, Enestor da Rosa dos (2002): “Migração e seleção: O caso do Brasil,” Ph.D. thesis. [3, 5]

Sassen, Saskia (1990): *The mobility of labor and capital: A study in international investment and labor flow*, Cambridge University Press. [7]

Silva, Sidney Antonio da (2006): “Bolivianos em São Paulo: Entre o sonho e a realidade,” *Estudos Avançados*, 20 (57), 157–170. [9]

Sjaastad, Larry A. (1962): “The costs and returns of human migration,” *Journal of Political Economy*, 70 (5, Part 2), 80–93. [2]

Soares, Weber, Carlos Lobo, and Ralfo Matos (2015): “Mobilidade espacial dos imigrantes estrangeiros no Brasil-1991/2010,” *REMHU: Revista Interdisciplinar da Mobilidade Humana*, 23 (44), 191–205. [2]

Stock, James H. and Mark W. Watson (2015): *Introduction to Econometrics*, Pearson Education, 3 ed. [5]

Suliano, Daniel Cirilo and Marcelo Lettieri Siqueira (2012): “Retornos da educação no Brasil em âmbito regional considerando um ambiente de menor desigualdade,” *Economia Aplicada*, 16 (1), 137–165. [6]

Vilela, Elaine Meire (2011): “Desigualdade e discriminação de imigrantes internacionais no mercado de trabalho brasileiro,” *Dados*, 54 (1), 89–128. [3, 8, 9, 17]

Wu, De-Min (1974): “Alternative tests of independence between stochastic regressors and disturbances: Finite sample results,” *Econometrica*, 42 (3), 529–546. [8]