

Earning Differentials by Occupational Categories and Discrimination: Gender, Race and Regions

Francisca Livia S. Menezes (CAEN/UFC)

Ronaldo A. Arraes (CAEN/UFC)

Andrei G. Simonassi (CAEN/UFC)

Resumo

Analisar os diferenciais de rendimentos nas ocupações dos trabalhadores brasileiros é o foco central deste trabalho. Em virtude da vasta diversidade de ocupações catalogadas pelo IBGE (cerca de 800), aplica-se um procedimento teórico para reduzi-las em apenas sete a fim de permitir tratamento estatístico dos dados. São estimadas equações mincerianas quantílicas, sobre as quais são feitas decomposições para verificar o hiato entre os rendimentos dos indivíduos com atributos inerentes distintos, bem como entre aqueles residentes em regiões mais e menos desenvolvidas do país. Os resultados das estimações ratificam a importância da análise desagregada em estratos ocupacionais e por quantis, uma vez que os efeitos das variáveis explicativas são distintos ao longo das distribuições de rendimentos e entre as ocupações. Em relação às lacunas existentes entre as distribuições de rendimentos, verifica-se um *glass ceiling effect* em algumas ocupações, ou seja, a defasagem é maior no topo da distribuição. Entretanto, na maioria dos casos, a ocorrência do hiato é maior na base da distribuição – *sticky floor effect*. Além disso, diferente das lacunas entre gêneros, que são decorrentes apenas da remuneração às características dos indivíduos (discriminação), os diferenciais de rendimentos entre raças e regiões também decorrem das diferentes características dos trabalhadores.

Palavras-Chave: Distribuição de Renda em Ocupações, Decomposição Contrafactual, Equações Mincerianas Quantílicas

Abstract

Analyzing the income differentials amongst the Brazilian workers' occupations is the focus of this paper. Due to the wide diversity of occupations cataloged by the IBGE (around 800), a theoretical procedure is applied to reduce them to only seven in order to allow statistical treatment of the data. The methodological approach is based on Mincer quantile equations to be estimated in various strata of the workers' income distribution, on which a breakdown is made to check the gap among the individuals' income with distinct inherent attributes and between those living in both more and less developed regions. The estimation results ensure the importance of breakdown analysis for occupational strata as well as for quantiles, since the effects of explanatory variables are distinct along the income distribution and vary among occupations. Regarding the gaps in this distribution, there is a glass ceiling effect in some occupations, that is, the gap is greater at the top of distribution. However, in most cases the gap is greater at the bottom of the distribution – sticky floor effect. Moreover, contrary to the gap in gender, which is due to the compensation characteristics of individuals only (discrimination), income differentials between regions and races also occur from the heterogeneous characteristics of workers.

Keywords: Earning Distributions in Occupations, Counterfactual Decomposition, Mincer Quantile Equations

JEL: J31, J44, J71, C21

1. INTRODUCTION

Economic growth and the increase in demand for jobs has generated significant changes in the Brazilian job market dynamics. Besides that, the recent phenomenon of income inequality reduction has a direct implication for the transformations that have occurred in the market occupational distribution, be it transformations in demand, be it supply ones. (BARROS, 2007).

However, centuries old problems such salary inequalities between individuals who work at similar activities and also have the same qualifications, are still observable. For instance, the salary gap

between men and women, between whites and non-whites and between residents of more and less developed regions. In some cases, acquired attributes such as education level and work experience seem to have less relevance than a worker's gender or race. As an illustration to grasp this idea, data from the Brazil's National Household Sampling Survey (PNAD) of 2011 reveals that for individuals 10 years and older, the averages of education years of women and men are respectively 7,5 and 7,1, which means a 5,6% superiority in favor of women. On the other hand, the men's average monthly earning are 73,7% higher than that for women, that is, R\$1.129 and R\$650, respectively.

So, it becomes relevant to investigate how education and other personal attributes have affected the Brazilian worker. It is also believed to be indispensable such analysis to take into account not only distinctions between inherent attributes and regionalization, but also differences among the diverse occupational categories due to requirements on each of them regarding human capital level.

On an empirical application of such nature, it is usual to get theoretical support based on Mincer equations that spot the effects of explanatory factors on the average variation of the individual worker's income. Worth reminding is that the impact of every factor can significantly vary among the income distribution strata. It is expected, for instance, that effects of inherent characteristics such as ethnicity and gender, or acquired characteristics such as human capital to differentiate according to occupation and income stratification. Therefore, the results obtained by applying this methodological technique on non-stratified samples, which is often applied in literature, produce inconsistent and distorted effects.

This paper relies on quantile Mincer equations to verify the impact of explanatory variables on every earning distribution quantile, allowing for the occurrence of differentiated effects when stratifying workers' income distribution by occupation. Starting from strata, an investigation about the lagging between workers' earning caused by gender, ethnicity and region differences, according to the methodology by Machado and Mata (2005), which expands the formulation of the Blinder-Oaxaca Decomposition (BLINDER, 1973; OAXACA, 1973). For this purpose, the conditional wage distributions of these categories are estimated via quantile equations, marginal distributions are obtained, when comparisons are made between the abovementioned categories, they turn possible the decomposition of differences in income distributions into two components: one that comes from observable characteristics of workers and another from the returns to these characteristics, which may be interpreted as discrimination.

The main methodological hurdle for the actualization of the estimations is the considerable number of more than eight hundred occupational categories designated by The Brazilian Institute Of Geography and Statistics Foundation (FIBGE) . To overcome such hurdle and make estimations possible, it is vital to implement a special handling on the data so as to create occupation groups and then have a statistically-practicable numbers that lead to inferences of greater degree of freedom and accuracy, despite the possibility of information loss.

The seminal contribution to the investigation of occupational choices is believed to have been made by Roy (1951), who identified in such choices, the skills and these skills' return distribution for every occupation. In order for a proper continuity of this analysis through the application of adequate inference methods with discrete choices, some researchers suggested the grouping of those numerous occupational categories in smaller strata.

One of the earliest proposals was made by Erickson, Goldthorpe and Portocarrero (1979) which was later honed by Goldthorpe, Llewellyn and Payne (1987), who designed a scheme known as EGP (named after the initials of the first authors). It is comprised of seven occupational classes and it is internationally used, including by the CASMIN Project (an acronym for *Comparative Analysis of Social Mobility in Industrial Nations*), whose purpose is centered on the analysis of labor mobility in the industrial realities of various European Union countries. Later on, many other grouping formulations came up, and all of show similarity to the EGP proposal. Among them, the Vale Silva (1988) formulation stands out, for it incorporates significant axes for the identification of similarities and differences between categories, such as the dichotomy between rural and urban sectors as well as the dichotomy between manual and non-manual labor, so it grouped the numerous categories into eighteen sets. Nevertheless, such scheme has not been widely used due to its analysis difficulty, so researchers have been opting for the implementation of a grouping proposal which contains a lower number of categories.

It is expected that this study will contribute to the debate about this subject when it conducts an analysis of the earning distribution quantiles of every occupation, and by individual characteristics, it identifies the factors that affect them. When using the EGP scheme, firstly a recoding of the PNAD (Brazil's National Household Sampling Survey) data is required to effectively group those various occupational categories into just seven sets, and then an analysis of the Brazilian job market by occupational categories is conducted. Such analysis is approached by the estimation of Mincer quantile equations and by identifying and decomposing the existing hiatus between men's income and women's incomes, whites and non-whites, individuals who reside in the Northeast region or in another more-developed region (namely the South, the Southeast or the Midwest). Besides verifying the strata in which the applied explanatory variables affect the income distribution, this procedure makes possible the detection of distortions that may occur between personal attribute effects and in addition it also detects the existence of likely effects such as *glass ceiling* or *sticky floor*, as previously seen in economic literature. Once the distribution strata that are affected by each factor are identified, it becomes less difficult to implement policies that aim to reduce income inequalities, which leads to the improvement of workers' well-being.

This article comprises four additional sections. A selective review of the literature is highlighted in the next section, followed by a discussion of both the methodology and the results. The last section is reserved to the final remarks.

2. A SELECTIVE REVIEW OF LITERATURE: MINCER EQUATIONS, DECOMPOSITION OF INCOME DISTRIBUTION AND OCCUPATIONS

In the 20th century, Schultz (1961) and Ben-Porath (1967) formulated, despite the absence of theoretical strictness, the relation between education and income. However, the debates and the empirical applications were intensified after the publishing of the seminal article by Mincer (1974), which formalized the theory of the derivation of such relation whose equation specifies the worker's income as a function of his or her formal education and experience acquired on the job. It is worth mentioning, that most applications use a simplified version of the Mincer model in which the rigidity hypothesis is assumed, that is, returns from education and experience are invariable among individuals. However, the most general of his model, and the most realistic one, acknowledges the possibility of such returns to vary among individuals. And that leads to a random-coefficient model (HARMON et al, 2003). So, when one stratifies earning distribution as it is done herein, this is the model to be applied.

In this sense, this paper uses equations which have that methodological structure, despite having a focus that differs from the focus usually seen in literature. For each occupational stratus, analyses of the Brazilian workers' income distribution and of the factors that cause changes in such distribution are conducted. Initially, quantile mincerian equations are estimated to investigate the effects of variables like education and experience on the workers' income distribution. From these results, conditional distributions are obtained from which in turn, the marginal ones will be extracted. Thus, it becomes possible to decompose the workers' income distributions into various components.

The first works that contained this kind of analysis were done by Oaxaca (1973) and by Blinder (1973), which decomposed the income distributions of males and females into a component that is attributed to distinct individual characteristics, and into another component derived from the differences between the returns to these characteristics. Such decomposition, however, was conducted via the Least Squares method (LS), only for the average of the distribution. Later studies applied quantile regressions in order to perform such analysis, thus giving information pertaining not only to the income distribution center, but also pertaining to the substrata.

A similar procedure was applied by Blau and Kahn (1996) to analyze the income inequalities in the United States in comparison to nine OECD countries, with data from the 1980's. Through the estimation of the income densities of male workers, the authors concluded that salary inequality in the U.S. is greater than in all other OECD countries. However, after disaggregating the income distributions into several parts, Blau and Kahn verified that inequality in the United States is considerably larger/smaller than in other countries in the lower/higher quantiles, respectively.

DiNardo *et al.* (1996) applied nonparametric methods to measure the effects of institutions and the effects of job market factors on the changes of income distribution in the United States from 1979 to 1988. They concluded that labor unions and supply and demand shocks are vital to the understanding of the income inequality rising in that period. Besides, they also found evidence that the decline of the actual minimum salary value explains most of such inequality increase, especially due to the gender factor.

Based on the Blinder–Oaxaca decomposition by applying quantile regressions, Albrecht *et al.* (2003) found in the Swedish workers' income distribution what is conventionally named in literature as the *glass ceiling effect*, that is, the difference in income between men and women is positive, it grows along the distribution and intensifies at the top.

The technique proposed by Machado and Mata (2005) has been applied in many recent articles. It utilizes quantile regressions to identify and decompose gaps between the incomes of different categories at different points in time. They applied the method with the purpose of verifying the changes in income distribution in Portugal regarding the period from 1986 to 1995, and concluded that the changes of the individuals' attributes as well as the returns of such attributes do contribute to the increase of income inequality. They also concluded that education contributed in a decisive way to this increase.

Albrecht *et al.* (2009) provided with greater theoretical support to the methodology of Machado and Mata (2005) by evidencing the asymptotic consistency and normality regarding such technique, besides taking into consideration sample selection adjustments. The empirical application of such technique aimed to test the hiatus between genders among Dutch workers' income, and it was then that these researchers noticed the presence of a *glass ceiling effect* on that distribution. Using these authors' methodology, Paredes (2012) also analyzes the gender discrimination in the Chilean economy by grouping up occupations in sets of nine¹, and identifies that the income hiatus between men and women has been narrowed in categories such as managers, upscale self-employed professionals, and especially in occupations dealing with finances and insurance. Similarly for Brazil and using data from PNAD/2007, Coelho *et al.* (2010) have also identified such an effect through the earning distribution gap of black and white Brazilian women by performing quantile regression with semiparametric correction for sample selection bias. Their main finding is that the returns to education are higher at the top of the distribution.

This work also follows the Methodology of Machado and Mata (2005), but it also verifies the changes in the Brazilian workers' income distributions by occupational groups which has more accurate results. So it will benefit policies that aim to reduce salary inequalities as those policies will be designed in a more efficient way.

The studies abovementioned analyze the variations of salary distributions either in an aggregate way or considering gender differences. This research proposes an analysis that takes into account individuals' different occupations. The most difficult task in researches that focus on this approach is to narrow down occupational categories due to the high number of disaggregated job specifications. So it becomes vital to aggregate them in order to accomplish a practicable analysis, even if that may lead to some information loss. Thus, various researchers have proposed different ways to aggregate such categories into smaller groups while simultaneously trying to avoid information loss.

The first versions of occupational distribution were proposed by Erickson, Goldthorpe and Portocarrero (1979) and improved by Goldthorpe, Llewellyn and Payne (1987), scheme internationally known as EGP (first authors' initials), which is comprised of seven classes. Based on this aggregation, occupations were then grouped in several countries studied by the CASMIN Project (*Comparative Analysis of Social Mobility in Industrial Nations*), which was coordinated by Goldthorpe in the early 1980's.

This methodological scheme takes into account the work relations between employers and employees in both rural and urban activity sectors. According Ribeiro and Scalón (2001) the resulting classes at the end of the aggregation are classes of comparable individuals in terms of their income level,

¹ The author makes no reference to the applied methodology for grouping the occupations.

financial safety degrees, career-advancement probabilities and these individuals' autonomy to perform work.

There are other class-aggregation proposals, but all of them share some degree of similarity with the proposal by Goldthorpe. Pastore (1979), by proposing that the aggregation of numerous categories to be done by forming six groups and that accentuates the heterogeneity of each of them due to the reduced number of strata. Aiming to reduce such in-between groups difference, Silva (1988) proposed the use of eighteen categories. Silva's proposal was not used by other researchers as they considered this too high a number of categories. The opposite happened to Scalon (1999) who rearranged Silva's (1988) eighteen categories into nine but managed to minimize the loss of information. To achieve this, Scalon applied cluster analysis techniques and also a criterion for combining categories based on log-linear models.

With the goal of verifying which are the determinants of the distinctions in Brazilian workers' earning distributions, this research uses the EGP scheme's theoretical principle and comes with a recoding of the PNAD data about occupations, for it grouped more than eight hundred occupational category classifications into just seven as follows: professional managers, routine non-manual laborers, small business owners, rural owners/employers, skilled manual laborers such as technicians and supervisors, unskilled manual laborers, and rural-sector unskilled manual laborers. Opposite to the analyses usually found in literature, where job market approaches are done in an aggregate fashion, this research stands out for its partitioning of the market by occupational category strata, and that leads to inferences that produce more realistic results for more efficient policies.

3. METHODOLOGY

The applied methodology comprises two procedures. In the first procedure it is measured the returns to workers' education through quantile Mincer equations which were estimated for each occupational stratum. The second procedure is a counterfactual decomposition of workers' earning, taking into account differences of gender, ethnicity and region. The data used and the variables that make up the models are to be described in the next sections.

3.1 Earning Equations

The first empirical exercise is the estimation of earning equations in the traditional way, however a little differently as this estimation applies a quantile regression model created by Koenker and Basset (1978). This way, a detailed analysis of workers' income distribution can be conducted and it differs from the usual application of the LS method in this kind of investigation which provides only with the characteristics of a part of the distribution of the response variable. However, for developing countries such as Brazil, where there is a marked heterogeneity among workers in terms of income and also in terms of inherent and acquired attributes, where the impact of individual characteristics on earnings is utterly significant, the analysis by quantile regressions is the most suitable one. Equations like these are also used in order to obtain the conditioned income distributions that are so necessary for the decomposition proposed by Machado and Mata (2005), which is described as follows.

3.2 Decomposition of Earning Distribution

Counterfactual decomposition can be understood as a generalization of the Blinder–Oaxaca (1973) procedure, that applies the LS method to obtain the estimations of the median of the distribution, but with the advantage of using quantile regressions that are modeled on each distribution quantile. This technique allows for the decomposition of the differences of income distributions of various categories. For instance, men and women become two components which derived from differences between observable characteristic distributions (education, experience,etc) for genders, and also derived from the returns of these characteristics. Therefore, decomposition makes possible to identify the existing gap between the incomes of distinct classes and also to measure how much of this gap is distributed as a consequence of characteristics and coefficients. The effect of this may be considered as discrimination between categories. This decomposition has some stages to be described as follows.

3.2.1 Conditioned Income Distributions

The first phase of decomposition is to derive the conditioned income distributions which are estimated through the quantile Mincer equations. This kind of equation models the relation between regressors type (z) – education level, experience, square experience, gender, ethnicity, regionalization – in the corresponding quantiles of the dependent variable [$w = \log(\text{earning})$]:

$$Q_q(w|z) = z' \beta_q \quad (1)$$

Where $Q_q(w|z)$, for $q \in [0,1]$, is the conditional q^{th} quantile of the distribution of the dependent variable and β_q is the q^{th} coefficient vector of the mincerian equation.

As the sample average is the solution to minimizing a residual sum of squares and the median is the solution to a sum of absolute values of the residuals, the quantile regression can be defined as a solution to an asymmetric minimization of a weighted sum of the absolute values of the residuals. This way, β_q can be estimated by solving this following problem (KOENKER AND BASSET, 1978):

$$\min_{\beta \in R^k} \sum_{i:w_i \geq z_i' \beta} q |w_i - z_i' \beta_q| + \sum_{i:w_i < z_i' \beta} (1 - q) |w_i - z_i' \beta_q| \quad (2)$$

Despite not having an explicit form, this problem can be efficiently solved by the application of linear programming methods and the standard errors can be known by *bootstrapping* (MARTINS and PEREIRA, 2004). Besides, Cameron and Trivedi (2009) emphasize that quantile regressions become robust in the presence of *outliers* and the absence of normality of the dependent variable, and they do not need the existence of the conditional mean to be consistent. These characteristics favor and justify its application in numerous economic investigations just like this one.

3.2.2 Marginal Density of Wages and Decomposition of Changes In Wage Densities.

After specifying the conditional quantile function, the estimated parameters are used to simulate the conditional distribution of w given z , by applying the Probability Integral Transform Theorem : If Y is a continuous random variable with *fc*p $F_Y(y)$, then, the random variable $F_Y^{-1}(X)$, where X is $U(0, 1)$, has a distribution $F_Y(x)$, if U is a uniformly-distributed variable in $[0,1]$, then $F^{-1}(U)$ has density function FF . If q_1, q_1, \dots, q_M derives from a uniform distributio $(0,1)$, the corresponding M estimations of the quantiles of the w conditionals to z_i , $\hat{w}_i \equiv \{z_i' \hat{\beta}_q\}_{i=1}^M$, constitute a random sample of the (estimated) conditional distribution of the income given z_i .

This simulation procedure characterizes the conditional quantiles for all z , but it does not gives the w marginal density, once this marginal density depends not only on the conditional function $\hat{\beta}_q$, but also on the distribution of covariates $g(z)$. Aiming to generate a random sample of the marginal density w , one can simulate $g(z)$ data lines and, for every z_i line, starting from a distribution $U(1)$, simulate a random q_i and then form $\hat{w}_i \equiv z_i' \hat{\beta}_{q_i}$, which is a simulation of the w density implied by the model. By replicating this technique several times, it is possible to simulate a random sample of the desired distribution. This procedure is equivalent to numerically integrating the estimated quantile conditional function $\hat{Q}_q(w|z)$ with the distribution of s z and q to obtain $f(\hat{w}) = \iint \hat{Q}_q(w|z)g(z)dzdq$.

Inferring about the gap between the incomes of two categories (e.g. between men and women) requires to estimate two marginal densities of earnings: the density function estimated according to the women's earning distribution – based on a sub-sample – can be defined as the marginal distribution derived from the model; and the counterfactual density function, which in this case corresponds to the men's density function related to the women's characteristic distribution (MACHADO and MATA, 2005).

The procedure to calculate the marginal distribution derived from the model is summarized by these steps:

1. A size- M random sample is generated corresponding to the M estimates of the conditional quantiles, from a distribution $U(1)$: q_1, q_1, \dots, q_M .

2. Using data about female individuals, it is then estimated that: $Q_{q_i}(w|z) = z'\beta_q$. As a result, M estimates for the quantile regression coefficients, $\hat{\beta}_{fq_i}$; where the f symbol refers to females.
3. A size- M random sample is generated, with replacement of the vector of covariates, z_f ;
4. Using this sample, an estimate is done of the density function through the multiplication of the estimated vector of coefficients times the generated size- M vector of covariates: $\{\hat{w}_i \equiv z_{if}\hat{\beta}'_{fq_i}\}_{i=1}^M$.

The counterfactual distribution is calculated according to the procedure above, but, on step 2, it is applied data of male workers to estimate the vector of coefficients and then obtain: $\{\hat{w}_i \equiv z_{if}\hat{\beta}'_{mq_i}\}_{i=1}^M$, where the m symbol represents males. The difference of these two density functions defines gender discrimination in a given period.

The described procedure conducts the investigation on the gaps between earning distributions caused by differences of gender, ethnicity and region. Then, for every occupational stratum, threefold replications of the procedure are required, for every set of the binary characteristics, namely: gender, race and region.

3.3 Database

The database for the estimation of the proposed models was obtained from Brazil's National Household Sampling Research (PNAD), which is conducted yearly by The Brazilian Institute Of Geography and Statistics (IBGE). Information regarding the year 2011 was gathered (the most recent period of available data).

3.3.1 Occupation Definitions

Starting from the theoretical principle of the EGP scheme, a recoding of the PNAD data was done so as to aggregate more than 800 occupational categories into just seven which are described as follows:

1. Professional general directors and managers (P.G.D.M.)
Legislators, general managers of government departments, operation and production managers, bioengineering professionals, biotechnology and genetic engineering professionals, mechatronics engineers, mathematicians, physicists, chemists, architects, doctors, surgeon dentists, veterinarians, pharmacists, teachers, lawyers, judges, managers and accountants.
2. Routine non-manual laborers (R.N.M.L.)
Registrars in general, agents, assistants and managerial assistants, secretaries, office machinery operators, office juniors, storekeepers and warehouse keepers, postmen, checkout operators and box office staff, receptionists, telephone operators, telemarketers, clerks, pollsters/interviewers.
3. Small business owners (S.B.O.)
Employers of more than five employees, providing they have declared themselves as employers or self-employed individuals: wholesale agents and sales technicians, dry-cleaners, laundry workers and similar kinds of professionals, replenishers and retail associates responsible for markups and markdowns, sales supervisors.
4. Skilled manual laborers such as technicians and supervisors (S.M.L.T.S.)
Technical designers and fashion designers, civil engineering/construction technicians, petrochemical technicians, topographic/land survey/hydrographic technicians, production technicians and food technologists, industrial lab staff, medical and dental instrument technicians.
5. Unskilled Manual Laborers (U.M.L.)
Workers in extraction of solid minerals, earthworks and foundations, masonry structures, weaving preparation, textile craft, feeders in production lines.
6. Rural owners/employers (R.O.E.)
7. Rural-sector unskilled manual laborers (R.S.U.M.L.)

Due to the heterogeneity among rural and urban workers, as well as the different work environments in which these two groups operate, the categories (6) and (7) were excluded from sampling. These categories are the ones including rural workers. The purpose of such exclusion was to give the analysis the necessary robustness. With the purpose of homogenizing the sample, the northern

workers were left aside, due to certain socio-geographic peculiarities of the north region that may disguise the urban characteristics of native workers as compared to the other regions of the country.

3.3.2 Variable Definition

For each occupational group mentioned in the previous subsection, quantile Mincer equations were estimated, and then from these equations, a counterfactual decomposition of wage distribution was conceived. The equations were estimated according to the following specification:

$$w_i = \alpha + \beta_1 ed_i + \beta_2 xp_i + \beta_2 (xp_i)^2 + \beta_3 ra\acute{c}a_i + \beta_4 gen_i + \beta_5 co_i + \beta_6 ne_i + \beta_7 se_i + \beta_8 s_i + u_i \quad (2)$$

- (w) is the main work's wage logarithm
- (ed) are the education years – corresponds to a grade or a level. For the purpose of equivalence, each accomplished passed grade is considered as one education year.
- (xp) is the calculated experience, it conceived as age minus six, minus the education years ². When adopting this procedure, it is assumed that all individuals started out their school years at the age of six and they do not spend any time outside school in the market (BERNDT, 1991 *apud* KASSOUF, 1998), that is, school-attending individuals do not work. Besides, it is hypothesized that the sample accounts for only individuals in the range of 15 – 55 years of age.
- (ETH) is for ethnicity - 1 for whites and 0 for non-whites;
- (GEN) is for gender - 1 for female and 0 for male;
- (MW) represents Midwest- 1 for the midwest region residents and 0, otherwise;
- (NE) Northeast- 1 for the northeast region residents and 0, otherwise;
- (SE) Southeast- 1 for the southeast region residents and 0, otherwise;
- (S) South- 1 for the south region residents and 0, otherwise;

The descriptive statistics of the variables applied in the model of each occupational category is shown in Table 1. The wide amplitude in the variations of the continuous variables points for heterogeneity among the occupational categories, which in turn justifies the suitability of the disaggregation performed. It is worth noting that the low relative representativeness of the small business owners category may lead to results that diverge from those of other categories.

Table 1- Descriptive Statistics by Occupational Category

Variable	1. P.G.D.M.		2. R.N.M.L.		3. S.B.O.		4. S.M.L.T.S.		5. U.M.L.	
	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
w	3421.27	4263.71	1177.91	1442.35	5917.30	7409.11	1094.28	1077.51	844.72	816.86
ed	14.14	2.82	11.78	3.22	12.71	3.33	9.37	3.75	8.28	3.84
xp	16.93	10.02	15.24	11.12	22.72	9.61	19.85	11.83	21.83	11.89
GEN	0.47	0.50	0.59	0.49	0.31	0.46	0.33	0.47	0.43	0.49
ETH	0.69	0.46	0.54	0.50	0.78	0.42	0.47	0.50	0.41	0.49
MW	0.16	0.36	0.15	0.35	0.15	0.36	0.13	0.34	0.14	0.34
NE	0.23	0.42	0.29	0.45	0.20	0.40	0.27	0.45	0.33	0.47
SE	0.39	0.49	0.36	0.48	0.35	0.48	0.38	0.48	0.34	0.47
S	0.23	0.42	0.21	0.41	0.30	0.46	0.22	0.41	0.19	0.39
Number Of Observations	11211		36150		743		27904		28050	

Source: Own calculations using PNAD/2011 data.

4 RESULTS

In this section, the results of both proposed empirical enterprises are presented. In the first subsection brings the quantile mincerian equations, followed by the earning decomposition.

4.1 Quantile Mincerian Equations

The quantile mincerian equations estimates by occupational stratum, as well as by LS are shown in Table 2. They reveal that all the coefficients are significant up to a maximum of 5%. In the majority of

² This procedure was adopted because the PNAD did not have information about individuals' experience.

the cases, besides differing from the estimation by LS, the effect of the explanatory variables is distinct along the distribution of the dependent variable distribution. Such characteristic is statistically proved by the Wald Test, whose results are shown in Table 4 of Appendix A. This effect also varies between the occupational strata and can be easily seen on the graphs 1–5 of Appendix B. These graphs, besides showing the results by LS estimation, also evidence the influence of each explanatory variable on the dependent quantiles by occupational category.

The coefficients of education (*ed*) and experience (*xp*) variables are positive in all cases, with stronger effect on the general directors' and managers' category (P.G.D.M.) This implies, following the literature, that human capital contributes significantly to the increase of workers' earnings. It is worth emphasizing that, except for the unskilled manual laborers' category (U.M.L.), the effect of these variables is greater at the top of the earnings distribution. For the total number of workers, a similar effect was noticed in the U.S.A. (Chamberlain, 1994; Buchinsky, 1994), Germany (Fitzenberger and Kurz, 1997), Uruguay (González and Miles, 2001), Zambia (Nielsen and Rosholm, 2001), Portugal (Machado and Mata, 2005) and Brazil (Coelho *et al*, 2010) for a sample of women in the range of 20–60 years of age. The xp^2 variable coefficients measure to what extent additional experience years translate into salary increases. This, in most cases, shows the expected negative sign, evidencing that at some point the effect of experience on salaries goes null or even negative. In certain points of the distribution of some occupational strata, these coefficients differ from what was expected. Although positive, they are in fact practically null.

Gender and ethnicity show a negative-and positive sign coefficients, respectively, in all quantiles and occupations, evidencing that if a worker is a female or a non-white person, she or he will get smaller earnings. Regarding gender, this result is similar to the result found by Machado and Mata (2005) in their study about Portuguese workers. The effects of these variables between the quantiles are distinct in the magnitudes of the coefficients between occupations. For the effect is stronger on some of higher quantiles while in some others, the opposite is true.

Inserted in the model to detect differences in earnings variation between residents of more-and less-developed regions, the dummy variables proved to be relevant in most of the studied cases. As expected, the negative effects are predominant on the earnings of individuals who reside in the least-developed region (the northeast) in almost all quantiles and occupations. Except for Small Business Owners, on which the null effect prevails in all regions, but principally in the southeast, and also income strata whose results might have been derived from the homogeneity degree and low sampling representativeness relative to the other occupations.

4.2 Decomposition of Earning Distribution

In order to decompose the workers' earning distribution into two components – a component due to individual distinct characteristics and another one that comes from the coefficients (that is, remuneration to these characteristics) – the procedure described in section 3 was conducted with a replication number *M* equals to 1000. The characteristics included were: years of education, experience, ethnicity (when conducting decomposition by gender) and gender (when conducting decomposition by ethnicity and by region).

The results of the decomposition of the gap that exists in-between men's and women's income, as well as results of the whites and non-whites and workers residing in the northeast region residents in comparison to other regions, by occupational stratum, are shown in table 3.

Regarding the gap between the men's income and women's income, it is clear that in all occupational strata, the gap is negative while it is significantly wider in the unskilled manual laborer category, where in smaller quantiles, it surpasses 100% mark, and that the proportion originated by the characteristics is positive. Therefore, the wage differential has to do with the coefficients, that is, it has to do with the remuneration given to the individual due to his or her characteristics, and this means the presence of discrimination by gender in the job market. A similar result was verified by Albrecht (2008) for the Netherlands and by Arulampalam, Booth and Bryan (2005) for several European countries.

Tabela 2 – Quantile Mincerian Equations by Occupational Category

Variable	1. P.G.D.M.					2. R.N.M.L.					3. S.B.O.					4. S.M.L.T.S.					5. U.M.L.				
	5%	25%	50%	75%	LS	5%	25%	50%	75%	LS	5%	25%	50%	75%	LS	5%	25%	50%	75%	LS	5%	25%	50%	75%	LS
<i>ed</i>	0.095 (0.005)	0.121 (0.003)	0.136 (0.004)	0.145 (0.004)	0.128 (0.003)	0.088 (0.003)	0.067 (0.002)	0.084 (0.001)	0.104 (0.002)	0.097 (0.001)	0.102 (0.016)	0.093 (0.017)	0.108 (0.016)	0.109 (0.014)	0.099 (0.010)	0.041 (0.004)	0.039 (0.001)	0.051 (0.001)	0.073 (0.001)	0.065 (0.001)	0.061 (0.004)	0.055 (0.001)	0.050 (0.001)	0.061 (0.001)	0.063 (0.001)
<i>xp</i>	0.054 (0.004)	0.054 (0.003)	0.059 (0.003)	0.058 (0.004)	0.056 (0.003)	0.050 (0.002)	0.031 (0.001)	0.033 (0.001)	0.043 (0.001)	0.040 (0.001)	0.045 (0.020)	0.056 (0.017)	0.013* (0.020)	0.026* (0.020)	0.032 (0.012)	0.034 (0.004)	0.024 (0.001)	0.030 (0.001)	0.037 (0.001)	0.034 (0.001)	0.042 (0.003)	0.037 (0.001)	0.032 (0.001)	0.034 (0.001)	0.039 (0.001)
<i>xp²</i>	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.001 (0.000)	0.000* (0.000)	-0.001* (0.000)	0.000* (0.000)	0.000* (0.000)	0.000* (0.000)	-0.001 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.001 (0.000)	-0.001 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.001 (0.000)
<i>GEN</i>	-0.323 (0.025)	-0.367 (0.016)	-0.388 (0.016)	-0.388 (0.020)	-0.370 (0.014)	-0.262 (0.013)	-0.216 (0.006)	-0.250 (0.006)	-0.316 (0.009)	-0.308 (0.006)	-0.510 (0.120)	-0.450 (0.080)	-0.465 (0.081)	-0.410 (0.084)	-0.490 (0.063)	-0.621 (0.030)	-0.339 (0.009)	-0.390 (0.007)	-0.458 (0.009)	-0.452 (0.007)	-1.182 (0.025)	-0.746 (0.013)	-0.549 (0.009)	-0.558 (0.008)	-0.712 (0.007)
<i>ETH</i>	0.088 (0.027)	0.129 (0.018)	0.172 (0.021)	0.211 (0.022)	0.170 (0.016)	0.076 (0.014)	0.064 (0.006)	0.084 (0.007)	0.128 (0.009)	0.107 (0.006)	0.297 (0.150)	0.168* (0.089)	0.163* (0.123)	0.082* (0.083)	0.157 (0.072)	0.067 (0.016)	0.047 (0.007)	0.073 (0.008)	0.106 (0.009)	0.097 (0.007)	0.036* (0.021)	0.058 (0.010)	0.063 (0.008)	0.091 (0.008)	0.079 (0.008)
<i>NE</i>	-0.313 (0.036)	-0.221 (0.025)	-0.148 (0.026)	-0.07* (0.029)	-0.140 (0.022)	-0.501 (0.024)	-0.327 (0.010)	-0.314 (0.009)	-0.280 (0.013)	-0.320 (0.009)	-0.15* (0.175)	0.002* (0.125)	0.267 (0.133)	0.465 (0.117)	0.222 (0.086)	-0.673 (0.033)	-0.370 (0.010)	-0.353 (0.010)	-0.324 (0.012)	-0.373 (0.011)	-0.651 (0.034)	-0.471 (0.014)	-0.395 (0.011)	-0.343 (0.011)	-0.429 (0.011)
<i>W</i>	0.006* (0.035)	0.093 (0.031)	0.182 (0.028)	0.310 (0.035)	0.190 (0.023)	0.016* (0.020)	-0.044 (0.009)	-0.022 (0.011)	0.055 (0.017)	0.038 (0.010)	-0.10* (0.213)	-0.04* (0.108)	0.263 (0.107)	0.391 (0.107)	0.211 (0.093)	-0.070 (0.021)	-0.108 (0.011)	-0.077 (0.013)	-0.003* (0.017)	-0.02* (0.012)	-0.02* (0.033)	-0.038 (0.015)	-0.032 (0.012)	-0.006* (0.012)	0.002* (0.013)
<i>E</i>	-0.067 (0.028)	-0.03* (0.020)	0.001* (0.020)	0.060 (0.025)	0.012* (0.018)	0.001* (0.016)	-0.046 (0.008)	-0.043 (0.008)	-0.02* (0.011)	-0.027 (0.008)	-0.14* (0.144)	-0.07* (0.078)	0.016* (0.090)	0.109* (0.088)	0.032* (0.072)	-0.090 (0.019)	-0.080 (0.009)	-0.050 (0.008)	-0.024 (0.012)	-0.044 (0.009)	-0.03* (0.026)	-0.047 (0.013)	-0.039 (0.010)	-0.029 (0.011)	-0.027 (0.011)
<i>c</i>	4.952 (0.073)	5.085 (0.052)	5.197 (0.052)	5.444 (0.064)	5.356 (0.045)	4.845 (0.051)	5.617 (0.019)	5.640 (0.016)	5.595 (0.021)	5.463 (0.016)	5.045 (0.333)	5.774 (0.294)	6.302 (0.300)	6.627 (0.272)	6.304 (0.193)	5.583 (0.056)	6.052 (0.020)	6.111 (0.018)	6.104 (0.021)	5.950 (0.017)	5.125 (0.052)	5.698 (0.023)	6.022 (0.022)	6.180 (0.021)	5.854 (0.019)

Source: Own estimations using PNAD/2011 data.

Notes: *not significant to 5% ; ** standard errors- in parentheses - calculated by Bootstrapping with 1000 repetitions; The South variable was excluded in order to prevent from perfect collinearity

Table 3 – Decomposition of Earning Differentials by Gender, Ethnicity and Regions by Selected Quantiles.

		Gender				Ethnicity				Region			
		5%	25%	50%	75%	5%	25%	50%	75%	5%	25%	50%	75%
<i>P.G.D.M.</i>	Total	-0.1913 (0.0171)	-0.2715 (0.0126)	-0.3228 (0.0129)	-0.3499 (0.0149)	-0.1997 (0.0185)	-0.2777 (0.0149)	-0.3017 (0.0158)	-0.3067 (0.0180)	-0.3220 (0.0287)	-0.2893 (0.0192)	-0.2694 (0.0200)	-0.2426 (0.0232)
	Coefficients	-0.2909 (0.0160)	-0.3582 (0.0124)	-0.3949 (0.0127)	-0.4051 (0.0140)	-0.0886 (0.0144)	-0.1712 (0.0108)	-0.2076 (0.0113)	-0.2218 (0.0131)	-0.2640 (0.0129)	-0.2482 (0.0100)	-0.2407 (0.0105)	-0.2233 (0.0121)
	Characteristics	0.0996 (0.0172)	0.0867 (0.0132)	0.0721 (0.0133)	0.0552 (0.0142)	-0.1111 (0.0195)	-0.1065 (0.0134)	-0.0941 (0.0128)	-0.0850 (0.0140)	-0.0579 (0.0185)	-0.0411 (0.0132)	-0.0288 (0.0129)	-0.0193 (0.0144)
<i>R.N.M.L.</i>	Total	-0.2143 (0.0131)	-0.1383 (0.0035)	-0.2037 (0.0043)	-0.2754 (0.0060)	-0.2595 (0.0151)	-0.1588 (0.0039)	-0.2169 (0.0048)	-0.2827 (0.0067)	-0.5945 (0.0206)	-0.2970 (0.0066)	-0.3160 (0.0059)	-0.3599 (0.0094)
	Coefficients	-0.3016 (0.0127)	-0.2088 (0.0049)	-0.2731 (0.0060)	-0.3535 (0.0081)	-0.1904 (0.0118)	-0.1049 (0.0040)	-0.1551 (0.0051)	-0.2073 (0.0070)	-0.5691 (0.0093)	-0.2820 (0.0032)	-0.3008 (0.0042)	-0.3400 (0.0059)
	Characteristics	0.0873 (0.0113)	0.0705 (0.0051)	0.0694 (0.0064)	0.0781 (0.0084)	-0.0690 (0.0141)	-0.0539 (0.0047)	-0.0618 (0.0049)	-0.0754 (0.0065)	-0.0254 (0.0103)	-0.0149 (0.0040)	-0.0153 (0.0046)	-0.0199 (0.0064)
<i>S.B.O.</i>	Total	-0.3655 (0.0942)	-0.3668 (0.0657)	-0.3795 (0.0602)	-0.3656 (0.0720)	-0.1980 (0.1274)	-0.1887 (0.0795)	-0.0726 (0.0924)	0.0365 (0.0871)	-0.1750 (0.1693)	-0.0462 (0.0975)	0.1048 (0.0984)	0.2817 (0.1061)
	Coefficients	-0.4684 (0.0676)	-0.4556 (0.0442)	-0.4540 (0.0425)	-0.4277 (0.0495)	-0.2316 (0.0596)	-0.2017 (0.0403)	-0.0786 (0.0378)	0.0383 (0.0442)	-0.1446 (0.0543)	-0.0220 (0.0411)	0.1243 (0.0366)	0.3069 (0.0426)
	Characteristics	0.1029 (0.0638)	0.0888 (0.0460)	0.0745 (0.0476)	0.0621 (0.0506)	0.0336 (0.0666)	0.0130 (0.0468)	0.0060 (0.0429)	-0.0018 (0.0480)	-0.0304 (0.0651)	-0.0242 (0.0510)	-0.0194 (0.0445)	-0.0252 (0.0474)
<i>S.M.L.T.S.</i>	Total	-0.6427 (0.0283)	-0.2201 (0.0044)	-0.3944 (0.0058)	-0.4620 (0.0089)	-0.2735 (0.0200)	-0.1362 (0.0034)	-0.1953 (0.0052)	-0.2395 (0.0064)	-0.7232 (0.0271)	-0.2796 (0.0085)	-0.3770 (0.0066)	-0.4280 (0.0094)
	Coefficients	-0.6944 (0.0121)	-0.2679 (0.0044)	-0.4479 (0.0047)	0.0557 (0.0059)	-0.2527 (0.0160)	-0.1111 (0.0052)	-0.1595 (0.0058)	-0.1834 (0.0077)	-0.7048 (0.0118)	-0.2575 (0.0040)	-0.3418 (0.0042)	-0.3782 (0.0057)
	Characteristics	0.0517 (0.0109)	0.0479 (0.0051)	0.0536 (0.0053)	-0.5177 (0.0064)	-0.0207 (0.0171)	-0.0251 (0.0053)	-0.0358 (0.0057)	-0.0561 (0.0070)	-0.0184 (0.0129)	-0.0222 (0.0044)	-0.0352 (0.0047)	-0.0498 (0.0062)
<i>U.M.L.</i>	Total	-1.1780 (0.0203)	-0.7111 (0.0124)	-0.4656 (0.0059)	-0.5994 (0.0061)	-0.3330 (0.0175)	-0.2371 (0.0089)	-0.2089 (0.0047)	-0.2255 (0.0062)	-0.7512 (0.0268)	-0.5371 (0.0144)	-0.3873 (0.0065)	-0.3914 (0.0084)
	Coefficients	-1.1978 (0.0142)	-0.7279 (0.0043)	-0.4804 (0.0049)	-0.6134 (0.0065)	-0.2666 (0.0218)	-0.1793 (0.0075)	-0.1490 (0.0060)	-0.1622 (0.0079)	-0.7257 (0.0168)	-0.4985 (0.0052)	-0.3521 (0.0044)	-0.3509 (0.0059)
	Characteristics	0.0197 (0.0137)	0.0168 (0.0046)	0.0148 (0.0052)	0.0139 (0.0067)	-0.0663 (0.0231)	-0.0578 (0.0086)	-0.0600 (0.0058)	-0.0632 (0.0070)	-0.0255 (0.0173)	-0.0386 (0.0065)	-0.0352 (0.0053)	-0.0405 (0.0061)

Source: Own estimations using PNAD/2011 data.

Notes: Standard errors in parentheses - sampling by bootstrapping with 1000 repetitions.

The gaps between whites' and non-whites' incomes who reside in the northeast or in another region are also negative gaps. Except for the small business owners category, is that there are positive gaps in some points of the distribution. Just like in the decomposition by gender, the largest part of these lags come from the coefficients, which evidences the existence of discrimination. However, the proportion that comes from the workers' characteristics is also negative. Such fact means that non-white individuals who reside in the northeast region have characteristics (years of education and experience) that make them get smaller incomes than those of white individuals or individuals who reside in other regions. The implementation of policies that promote social integration of these categories becomes highly advisable, in the light of the evidences shown herein.

Graphs 6, 7 and 8 in appendix C generated by the Stata software after the decomposition procedure, illustrate, for every occupational stratum, the gaps between the men's and women's income

distributions, as well as between whites and non-whites and northeast residents and other regions, coming from the coefficients (discrimination) with a 95% trust interval. Some occupations show a glass ceiling effect, that is, the lag is higher at the top of the distribution. That can be observed, for instance, in the Professional General Directors and Managers category (P.G.D.M.), likewise in decompositions by gender and by race, where the income hiatus is higher than 19% in the bottom part of the distribution and higher than 30% at the top part. However, the incidence of a sticky floor effect is more likely. It is an effect whose name was used by Arulampalam, Booth and Bryan (2005) for situations when the hiatus is higher in the bottom part of the distribution. In the decomposition by regions, such effect is observed in all occupational strata, indicating that discrimination decreases as distribution evolves.

The results of the decomposition over all of the distribution can be better seen in graphs 9, 10 and 11, depicted in appendix F. It is clear the hiatus variation between the incomes along the quantiles and between occupations. In the decomposition by ethnicity, for example, in the category of Professional General Directors and Managers, the income differential significantly varies along the quantiles and does a strikingly distinct trajectory, different from the trajectory seen in the Small Business Owners (S.B.O.) category. Facts like these ratify the importance of disaggregated analysis of occupational strata and by quantiles.

FINAL REMARKS

By grouping hundreds of occupational categories (as they were characterized by the PNAD/2011 survey) into just five categories, and also by following the schematic EGP theoretical principle, an analysis has been conducted herein, of the Brazilian urban job market, by occupational stratum.

Quantile Mincer equations were estimated to test the effect of inherent and acquired attributes on various quantiles along workers' earning distribution. Abiding by the Machado and Mata (2005) methodology, a decomposition of such distribution has been conducted in order to verify the earning differentials between men and women, between whites and non-whites, between northeast region residents and other regions' residents. This procedure made it possible to evidence how much of the difference between these categories' incomes is caused by distinctions in characteristics (experience and education) and how much of the difference is caused by dissimilarities in remuneration to such characteristics, which conclusively means discrimination.

The results of the quantile Mincer equations ratify the importance of disaggregated analysis of occupational strata by quantiles, once the effects of the explanatory variables are distinct along income distributions and do vary between occupations. This suggests an alternative and realistic view of Mincer equations with variable coefficients, especially regarding human capital return. Variables such as years of education and experience, bring on a growth effect on workers' wages, mainly in the Professional General Directors and Managers category. On the other hand, regional variables show that the wages earned by workers who reside in the northeast is smaller, except for the Small Business Owners category.

Regarding the gaps between men's and women's income distributions, between whites and non-whites, and between northeast-region residents and other-region residents, some occupations display the presence of glass ceiling effect, that is, the lag is bigger at the top of the distribution. However, in most cases, there is a sticky floor effect, meaning that the hiatus is bigger in the bottom part of the distribution. In the decomposition by region, such effect is observable in all occupational strata, indicating that discrimination decreases as distribution progresses.

Regarding the income differential between men and women, it is negative in all occupational strata, expressively higher in the Unskilled Manual Laborers category and the proportion derived from the characteristics is positive. It shows that wage differential is originated mainly from the coefficients, that is, from remuneration to individual characteristics, which evidences discrimination by gender in the Brazilian job market.

Similarly, the gaps between the incomes of whites and non-whites as well as between the incomes of northeast residents and other-region residents are negative. Solely in the Small Business Owners category, and on few distribution points, these gaps are positive. These lags are mostly consequential to coefficients that indicate discrimination. However, unlike to what happens to gender, the differentials are also resultant from these workers' distinct characteristics. This indicates that non-white individuals and

northeast region residents have characteristics, such as years of education and experience, that cause them to get smaller income than the income earned by white individuals and residents of other regions. This fact confirms the imperativeness of policies that foster aid and the social integration of these categories.

BIBLIOGRAPHY

- ALBRECHT, J.; BJORKLAND, A.; VROMAN, S. Is there a glass ceiling over Sweden? **Journal of Labor Economics**. n.21, p.145-177, 2003.
- ALBRECHT, J.; van VUUREN, A.; VROMAN, S. Counterfactual distributions with sample selection adjustments: Econometric theory and an application to the Netherlands Labour Economics. **Labor Economics**. v.16, n.4; p.383-396, 2009.
- ARULAMPALAM, W., A. L. BOOTH, and M. L. BRYAN: **Is there a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution**. Mimeo, 2005.
- Blinder, A. Wage discrimination: Reduced form and structural estimates. **Journal of Human Resources** n.8, p.436-55, 1973.
- BLAU, Francine D.; KAHN, Lawrence. International Differences in Male Inequality: Institutions versus Market Forces. **The Journal of Political Economy**. v.104, n.4, p.791-837, 1996.
- Blinder, Alan S. Wage Discrimination: Reduced Form and Structural Estimates. **Journal of Human Resources**, v.8, n.4, p.436-455, 1973.
- BRASIL. Instituto Brasileiro de Geografia e Estatística. **IBGE**, 2012. Disponível em: <www.ibge.gov.br>. Acesso em: julho de 2012.
- BUCHINSKY M. Changes in the U.S. wage structure 1963–1987: an application of quantile regression. **Econometrica** v.62, p.405–458, 1994.
- CAMERON, A.C.; TRIVEDI, P.K. **Microeconometrics – Methods and Applications**. Cambridge University Press, Cambridge, 2005.
- _____. **Microeconometrics Using Stata**. Texas: Statacorp LP, 2009.
- CHAMBERLAIN G. Quantile regression, censoring and the structure of wages. In *Advances in Econometrics*, Sins C (ed.). **Cambridge University Press**: Cambridge, p. 171–209, 1994.
- DINARDO, John; FORTIN, Nicole; LEMIEUX, Thomas. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. **Econometrica**. v.64, n.5, p. 1001-1044.
- ERICKSON, Robert e GOLDTHORPE, John H. The CASMIN Project and the American Dream. **European Sociological Review**, v. 8, pp. 283-306, 1992.
- _____. **The Constant Flux: A Study of Class Mobility in Industrial Societies**. Oxford, Oxford University Press, 1993.
- _____ e PORTOCARRERO, Luciene. Intergenerational Class Mobility in Three Western European Societies. **British Journal of Sociology**, v. 30, 1979.
- FITZENBERGER B, KURZ C. **New insights on earnings trends across skill groups and industries in West Germany**. Universität Konstanz, mimeo, 1997.
- GOLDTHORPE J H: **Social Mobility and Class Structure in Modern Britain**, 2nd ed. Oxford: Clarendon Press, 1987.
- GONZÁLEZ X, MILES D. Wage inequality in a developing country: decrease in minimum wage or increase in education returns. **Empirical Economics** v. 26, p. 135–148, 2001.
- HARMON, C., HOGAN, V.; WALKER, I. Dispersion in the Economic Return to Schooling. **Labor Economics**, v.10, n.2, p.205-214, 2003.
- KASSOUF, A. L. Wage gender discrimination and segmentation in the Brazilian labor market. **Economia Aplicada**, v. 2, n. 2, p. 243-269, abr./jun. 1998.
- KOENKER, R.; BASSETT, G. W. Regression Quantiles. **Econométrica**, v. 46, n.1, p. 33–50. Jan. 1978.
- KOENKER, R.; HALLOCK. Quantile Regression. **Journal of Economic Perspectives**, v.15, n.4, p. 143-156, 2001.
- MACHADO, José A. F.; MATA, José. Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. **Journal of Applied Econometrics**. n.20, p. 445-465, 2005.

- MARTINS, Pedro S.; PEREIRA, Pedro T. Does Education reduce wage inequality? Quantile Regression evidence from 16 countries. **Labour Economics**. n.11, p. 355-371, 2004.
- NIELSEN H, ROSHOLM M. The public–private sector wage gap in Zambia? A quantile regression approach. **Empirical Economics** v.26, p.169–182, 2001.
- OAXACA, R.L. Male female wage differentials in urban labor markets. **International Economic Review**. v.14, n.3, p.693-709, 1973.
- PAREDES, Dusan. The impact of the occupations and economic activities on the gender wage gap using a counterfactual framework. **Serie de Documentos de Trabajo en Economía**, WP2012-08, 34 p., UCN, Chile, 2012.
- SCALON, Maria Celi. **Dados**, v. 41, n. 2, 1998.
- _____; RIBEIRO, Carlos Antônio Costa. Mobilidade de Classe no Brasil em Perspectiva Comparada. **Dados**, v.44, n.01, 2001.
- _____. **Mobilidade Social no Brasil: Padrões e Tendências**. Rio de Janeiro, Revan/IUPERJ-UCAM, 1999.
- SILVA, Nelson do Valle; RODITI, Déborah. "Et Plus Ça Change. Tendências Históricas da Fluidez Social no Brasil", in C. Hasenbalg e N. V. Silva (eds.), **Estrutura Social, Mobilidade e Raça**. Rio de Janeiro, IUPERJ/Vértice, 1988.

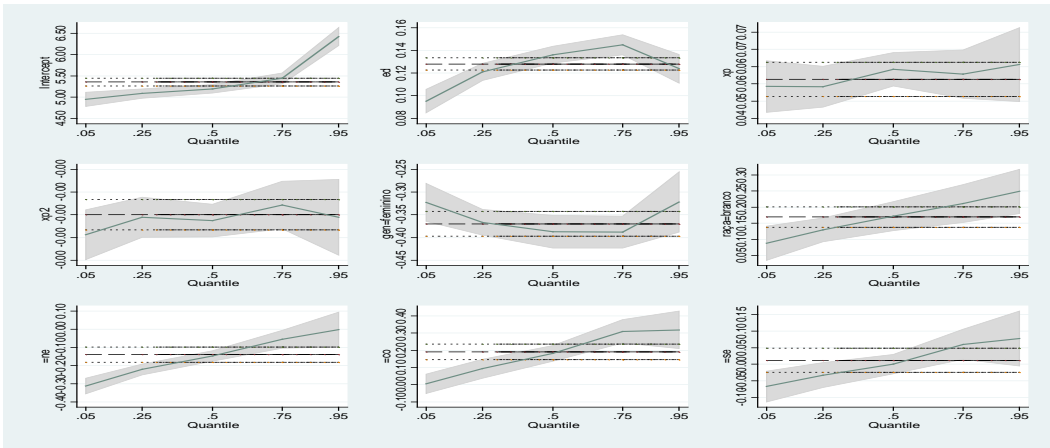
APPENDIX A – COEFFICIENT EQUALITY BETWEEN QUANTILES

Table 4 – Wald Test

Occupations	Coefficient	p-Value						
		q05=q25	q05=q50	q05=q75	q25=q50	q25=q75	q50=q75	q05=q25=q50=q75
1. P.G.D.M.	<i>ed</i>	0.0000	0.0000	0.0000	0.0000	0.0000	0.0101	0.0000
	<i>xp</i>	0.9742	0.2505	0.4888	0.0845	0.3761	0.6821	0.3861
	<i>xp²</i>	0.1437	0.2976	0.0590	0.7151	0.3357	0.1394	0.2076
	<i>GEN</i>	0.0679	0.0163	0.0293	0.1685	0.3160	0.9852	0.1090
	<i>ETH</i>	0.1235	0.0079	0.0003	0.0245	0.0008	0.0488	0.0017
	<i>NE</i>	0.0104	0.0000	0.0000	0.0022	0.0000	0.0004	0.0000
	<i>SE</i>	0.2242	0.0302	0.0002	0.0712	0.0003	0.0044	0.0007
	<i>co</i>	0.0210	0.0000	0.0000	0.0010	0.0000	0.0000	0.0000
2. R.N.M.L	<i>ed</i>	0.0000	0.2286	0.0000	0.0000	0.0000	0.0000	0.0000
	<i>xp</i>	0.0000	0.0000	0.0066	0.0060	0.0000	0.0000	0.0000
	<i>xp²</i>	0.0000	0.0000	0.0000	0.0003	0.3393	0.0378	0.0000
	<i>GEN</i>	0.0001	0.3710	0.0004	0.0000	0.0000	0.0000	0.0000
	<i>ETH.</i>	0.3658	0.5852	0.0007	0.0005	0.0000	0.0000	0.0000
	<i>NE</i>	0.0000	0.0000	0.0000	0.1161	0.0002	0.0010	0.0000
	<i>SE</i>	0.0018	0.0089	0.2953	0.6428	0.0147	0.0089	0.0007
	<i>co</i>	0.0011	0.0665	0.1189	0.0238	0.0000	0.0000	0.0000
3. S.B.O.	<i>ed</i>	0.6458	0.7464	0.7421	0.2900	0.3946	0.9879	0.7533
	<i>xp</i>	0.6296	0.2076	0.4830	0.0119	0.1677	0.4727	0.0961
	<i>xp²</i>	0.6319	0.1934	0.4512	0.0099	0.1462	0.4968	0.0826
	<i>GEN</i>	0.6030	0.7257	0.4597	0.8492	0.6721	0.4916	0.8484
	<i>ETH</i>	0.3753	0.4307	0.1890	0.9584	0.4035	0.4475	0.5874
	<i>NE</i>	0.3755	0.0302	0.0021	0.0260	0.0013	0.1062	0.0037
	<i>SE</i>	0.6385	0.3132	0.1189	0.2694	0.0623	0.2722	0.2609
	<i>co</i>	0.7580	0.0970	0.0301	0.0038	0.0007	0.2145	0.0048
4. S.M.L.T.S.	<i>ed</i>	0.5397	0.0034	0.0000	0.0000	0.0000	0.0000	0.0000
	<i>xp</i>	0.0045	0.2095	0.3819	0.0000	0.0000	0.0000	0.0000
	<i>xp²</i>	0.0025	0.0168	0.0690	0.0913	0.0099	0.0740	0.0014
	<i>GEN</i>	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	<i>ETH</i>	0.1760	0.7155	0.0306	0.0001	0.0000	0.0001	0.0000
	<i>NE</i>	0.0000	0.0000	0.0000	0.0827	0.0005	0.0041	0.0000
	<i>SE</i>	0.5876	0.0319	0.0016	0.0001	0.0000	0.0067	0.0000
	<i>co</i>	0.0560	0.7392	0.0069	0.0047	0.0000	0.0000	0.0000
5. U.M.L.	<i>ed</i>	0.0722	0.0014	0.9870	0.0000	0.0002	0.0000	0.0000
	<i>xp</i>	0.0756	0.0019	0.0141	0.0005	0.0690	0.1030	0.0008
	<i>xp²</i>	0.1582	0.0062	0.0081	0.0004	0.0030	0.9967	0.0016
	<i>GEN</i>	0.0000	0.0000	0.0000	0.0000	0.0000	0.2687	0.0000
	<i>ETH</i>	0.2263	0.1737	0.0095	0.5710	0.0014	0.0002	0.0009
	<i>NE</i>	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	<i>SE</i>	0.4525	0.6682	0.9723	0.5113	0.1909	0.2977	0.5098
	<i>co</i>	0.4375	0.5996	0.7918	0.6124	0.0416	0.0301	0.1061

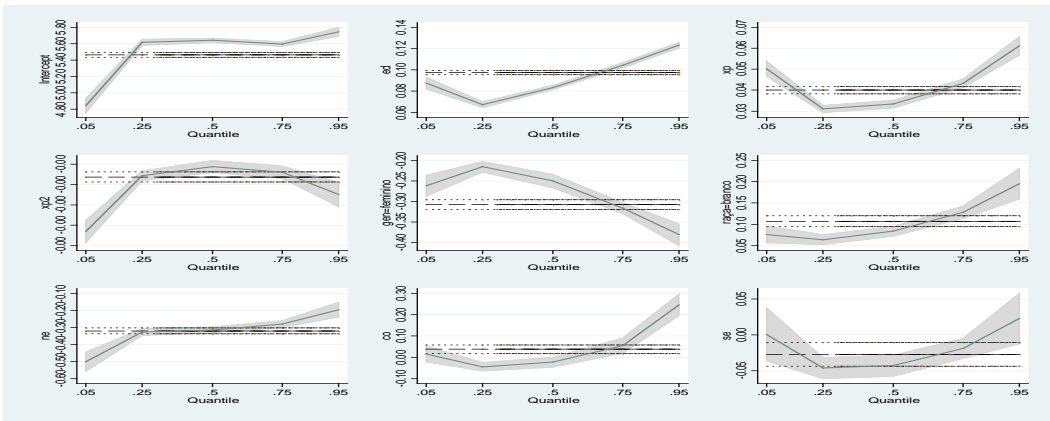
Source: Own estimations using PNAD/2011 data

APPENDIX B – QUANTILE MINCER EQUATIONS – THE EFFECTS OF EXPLANATORY VARIABLES



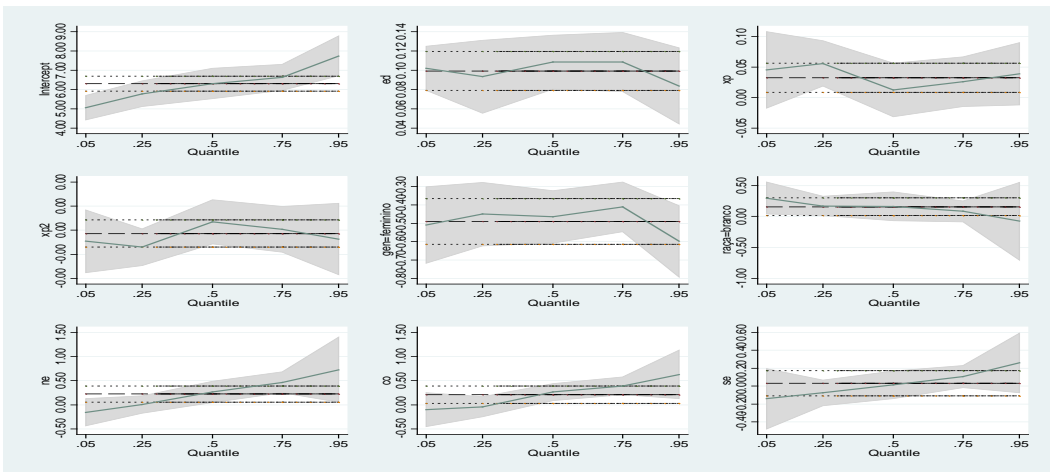
Graph 1 – Occupation 1

Source: Own elaboration using PNAD/2011 data.



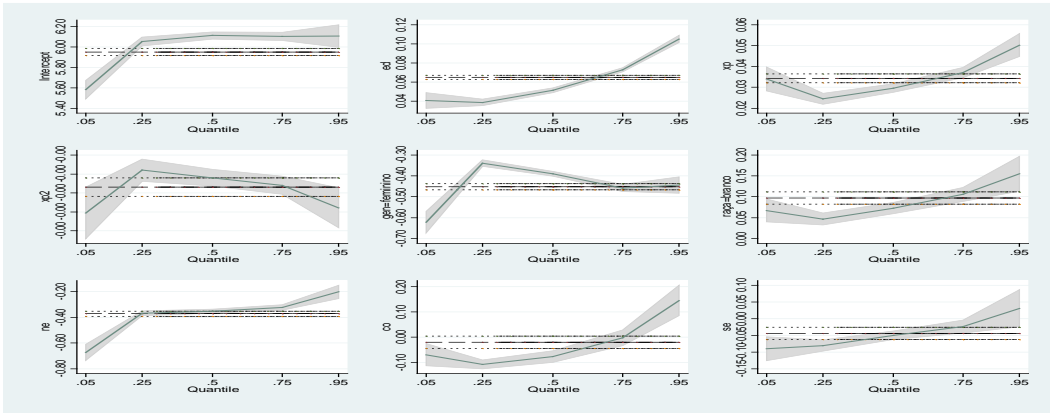
Graph 2 – Occupation 2

Source: Own elaboration using PNAD/2011 data.

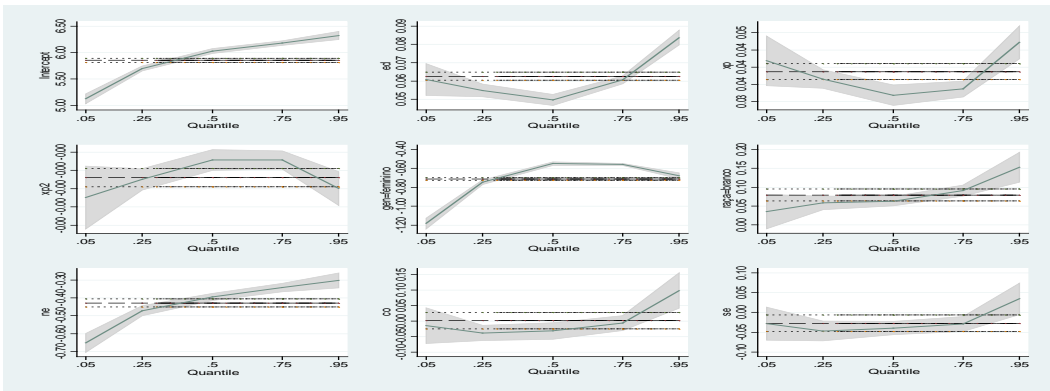


Graph 3 – Occupation 3

Source: Own elaboration using PNAD/2011 data.

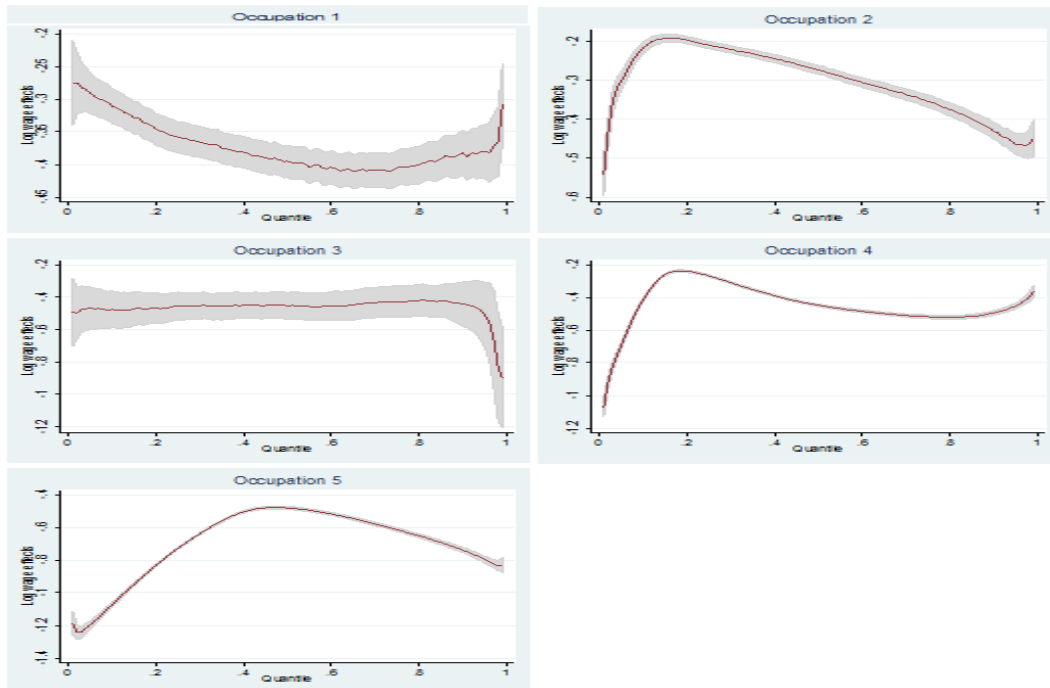


Graph 4 – Occupation 4
 Source: Own elaboration using PNAD/2011 data.

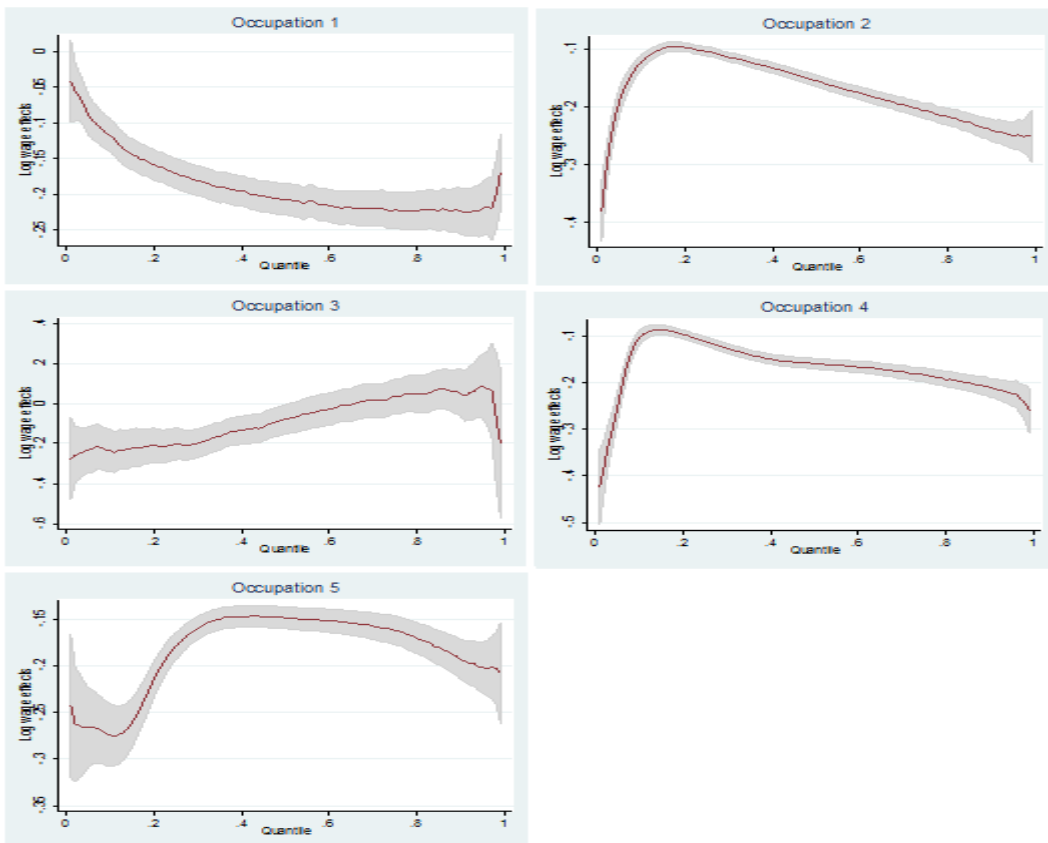


Graph 5 – Occupation 5
 Source: Own elaboration using PNAD/2011 data.

APPENDIX C – COEFFICIENT EFFECTS (DISCRIMINATION)

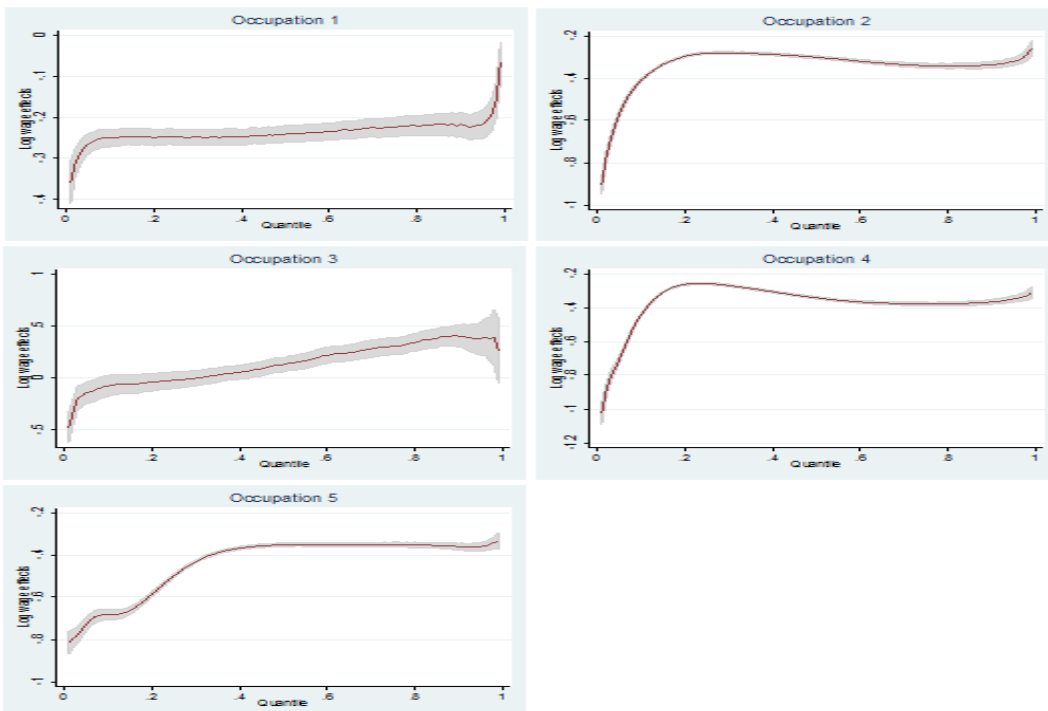


Graph 6 – Gender
 Source: Own elaboration using PNAD/2011 data.



Graph 7 – Ethnicity

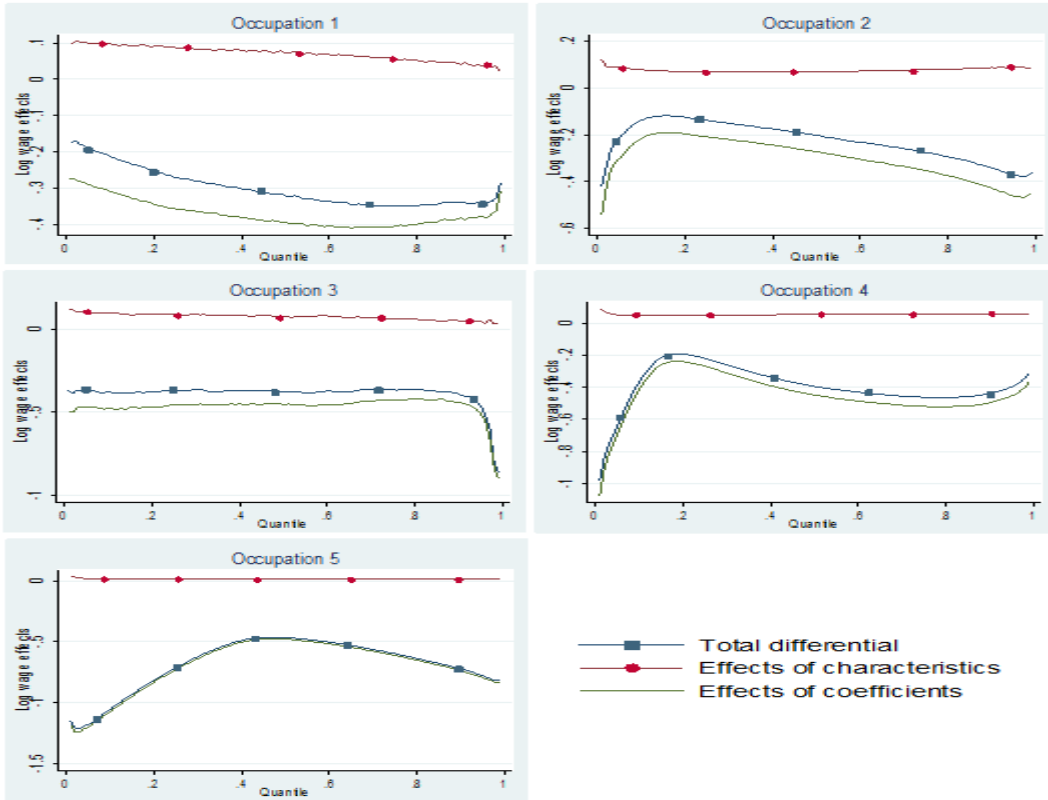
Source: Own elaboration using PNAD/2011 data.



Graph 8 – Northeast Region.

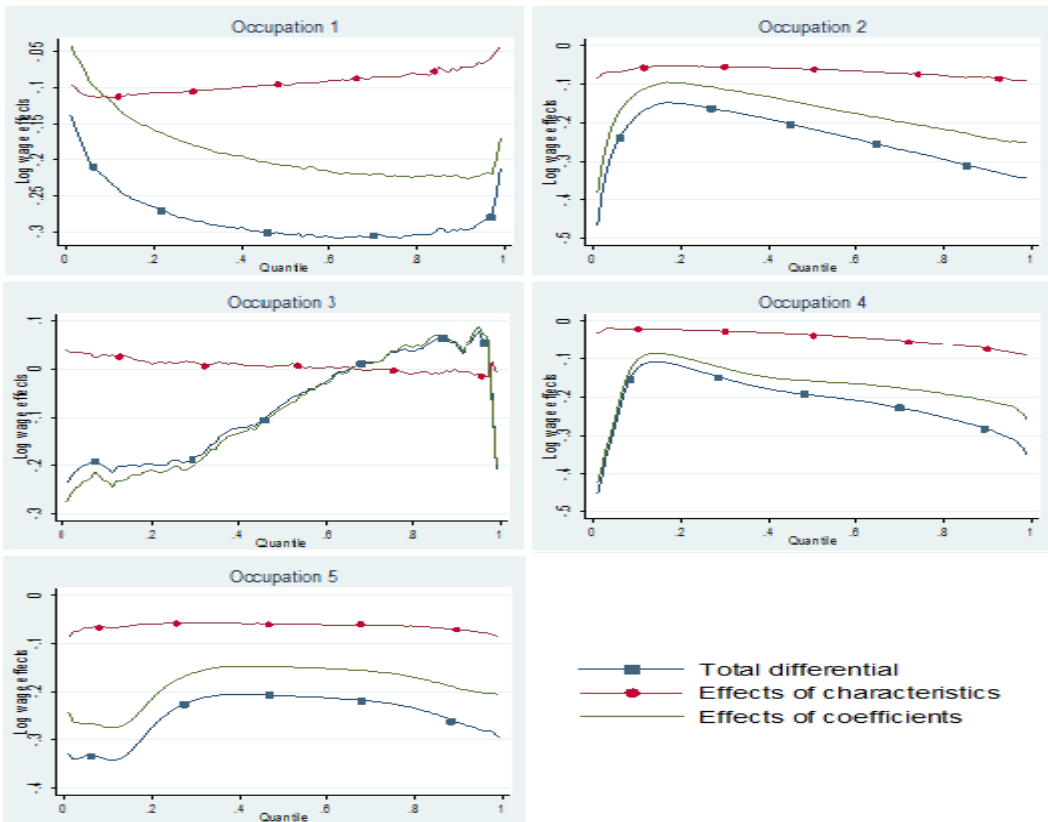
Source: Own elaboration using PNAD/2011 data.

APPENDIX F – DECOMPOSITION OF DIFFERENCES IN DISTRIBUTION



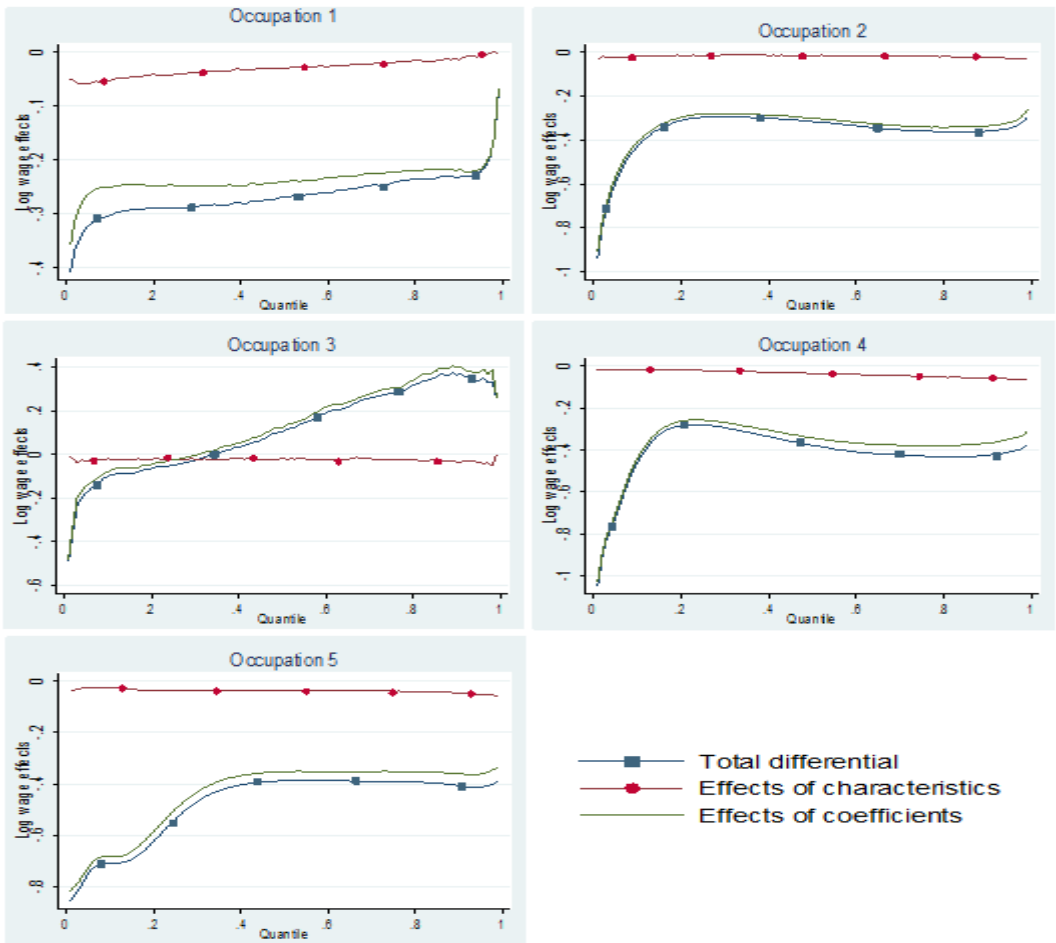
Graph 9 – Gender

Source: Own elaboration using PNAD/2011 data.



Graph 10 – Ethnicity

Source: Own elaboration using PNAD/2011 data.



Graph 11 – Northeast Region.
 Source: Own elaboration using PNAD/2011 data.