Wage Differentials between the Metropolitan Regions of Belo Horizonte and Salvador: a Discussion Based on the Blinder-Oaxaca Decomposition

ABSTRACT

This study quantifies and explains the difference in hourly wages between the metropolitan regions of Belo Horizonte (RMBH) and Salvador (RMS). The choice of these two regions was due to the interest in verifying how the difference of economic dynamism between the regions would impact the labor market is return. The method employed was the Blinder-Oaxaca decomposition. As expected, the results showed that the income per hour in RMBH were on average higher than those found in the RMS, due to greater economic concentration and agglomeration of the former region when compared to the latter.

KEY WORDS

Wage Differentials. Metropolitan Region of Belo Horizonte. Metropolitan Region of Salvador. Blinder-Oaxaca Decomposition.

Jader Fernandes Cirino

- Ph.D. in Applied Economics by the Federal University of Viçosa.
- Associate Professor in the Economics Department of the Federal University of Viçosa.

João Eustáquio de Lima

- Ph.D. In Rural Economics – Michigan State University.
- Postdoctorate in Quantitative Methods – University of Florida;
- Full Professor in the Rural Economics Department of the Federal University of Viçosa.

1 – INTRODUCTION

The occurrence of wage differentials in labor market has been widely discussed in the economic literature, especially when it concerns the gap between men and women and blacks and whites. In Brazil, Soares (2000); Campante; Crespo & Leite (2004); Giuberti & Menezes-Filho (2005) and Matos & Machado (2006) are relevant examples. As to the international literature, Stanley & Jarrel (1998); Horrace & Oaxaca (2001); Weichselbaumer & Winter-Ebmer (2003) and Lin (2007) are good examples.

This study focuses on comparing wages between regions, more specifically the Metropolitan Regions of Belo Horizonte (RMBH) and Salvador (RMS). The relevance of such study lies in the importance of decomposing the portion of wage differential that is due to the workers' productive characteristics from that related to regional characteristics of the labor market. Thus, it is possible to identify public policies suitable for equalize wages between the markets concerned, prioritizing those whose wages are less attractive.

The choice of the RMBH, to be compared with RMS, is due to two main aspects: firstly, the fact that the labor market of the former is more organized and institutionalized in what concerns formality and, secondly, the fact that the RMBH has a greater and more developed economy in comparison with that of the RMS.

As for the first aspect, data from the 2005 PED (Employment and Unemployment Research, DIEESE, 2007) showed that while a 6% hidden unemployment in the RMBH was subdivided in 3.2% precarious jobs and 2.8% discouragement, the 10.2% hidden unemployment in the RMS corresponded to 7.4% precarious jobs. This was higher more than twice the one found in the RMBH.[1] As for employed distribution, according to their occupation post, the 2005 research showed that, taking into consideration less stable and more precarious occupation forms identified by means of legal insertion status, the total sum of the rate of paid workers without proper registration, self-employed workers and domestic servants, indicated a greater vulnerability for posts offered by the economy of the RMS (42.9% of employed people in such situation, against 38.4% in the RMBH). .

As for the second aspect, Braga & Rodarte (2006) highlight the fact that the RMBH is an economically dynamic region, marked by the development of industrial activities and services sector with a greater concentration in properly registered work, which makes it offer better job opportunities if compared with the RMS. According to Dieese (2009), while the total unemployment rate in the RMS was the highest among all the Metropolitan Regions in the country analyzed by PED, peaking at 20.3% in 2008, for the RMBH it was the lowest level for that year, reaching 9.8%. Data from PNAD (National Household Survey, IBGE, 2007) related to occupation levels in 2006 reinforce this situation, since the rate in the RMS was 53.6% while it was 59.4% in the RMBH.

Thus, since the labor market in the RMBH shows greater and better job opportunities when compared with that of the RMS, it is expected that wages in the former region are higher than those in the latter. For that reason, the problem is put forward of investigating the wage differential between both markets, as well as separating the gap portion relating to workers' productive characteristics related to regional particularities in each market.

Therefore, this study aims to quantify the wage differential between the RMBH and the RMS, as well as decomposing it between the portion due to workers' productive attributes and that one caused by the characteristics of each region's economy and market. Specifically, we intend to: a) study, within each of the portions identified, the most important variables for wage gap formation; b) compare the regional wage differential behavior between genders; and c) analyze wage determinants and occupation probability for men and women in the markets.

2 – THEORETICAL BACKGROUNDS

In order to determine the wage differentials between the RMBH and the RMS, it is necessary to estimate

1 In this research, the unemployeds are divided into three categories: open unemployment, hidden unemployment by precarious jobs and hidden unemployment by discouragement. For further details about this classification and research, see Dieese (2008).

the earned income equations for the individuals in each of the markets. To do that, we use the human capital theory, whose pioneering essays are the ones by Schultz (1961) e Becker (1962). According to that theory, an individual's earned income can be expressed in terms of their professional qualities derived from their education and years of training and experience. This is due to the fact that, when the agent invests in any of the mentioned aspects, they improve in their human capital and, as consequence, in their marginal productivity, increasing the expected value of their work in the market.

In order to integrate the theory of investment in human capital in an empirical context compatible with the formal models of economic theory, Mincer (1974) put forward an equation that considers the influence of schooling and experience in people's wages. This equation, known as the salary function or Mincerian equation, was put forward as follows:

$$\ln Y_i = a + b_1 s_i + b_2 j_i + b_3 j_i^2 + v_i \quad (1)$$

where $ln\ Y$ is the natural logarithm for the individual's salary or earnings; s, worker's schooling in years; j, worker's experience accounted for their years in labor market; a and b_l (l = 1 to 3) are the parameters to be estimated; and v is the stochastic error term with usual properties.

The assumptions on equation (1) are that additional years of schooling and experience have a positive impact on wages, i.e., coefficients b_1 and b_2 are greater than zero. However, increases caused by the addition of experience would be subject to decreasing returns, being coefficient b_3 negative. Therefore, as pointed out by Berndt (1996), the relation between earnings and experience is not linear but rather parabolic, with a peak near the individual's average age, due to the depreciation of human capital with age.

It is important to point out that it is not always possible to collect data on individuals' experience, and this is the case of the database of this study (PNAD). Therefore, one possible solution is to consider the actual age as a proxy for individual's work experience, as suggested by Smith (2000). This author also highlighted that, for a better specification of equation (1), its regressors set should consider, besides productive attributes of schooling and age, other individual and labor market insertion characteristics, such as activity sector, region and legal bond with the employer.

3 – METHODOLOGY

The methodology of this study is divided into two parts. In the first, we present the necessary procedures to estimate the earnings equations. In the second, we discuss the methodology used to decompose the wage differentials between the RMHB and the RMS.

3.1 – Mincerian Earnings Equations for the Metropolitan Regions of Belo Horizonte and Salvador

In order to estimate the mincerian earnings equations for the labor markets, we used the Sample Selection Model, also known as Type II Tobit model, developed by Heckman (1979).

This model consists of estimating two equations: one to formally define the agent's decision to participate of the sample, called selection equation; the other one to explain the level of a given variable related to the previous decision.

In this study, the variable of interest is the level of earned income by an individual in labor market. The problem is that such variable is only observed if the person is employed in a paid activity. Otherwise, if the individual is inactive, idle or inserted in an unpaid activity, the earned income cannot be observed.

To solve this problem, the Sample Selection model explicitly considers the individual's decision to participate in the sample through the selection equation. By this approach, it is possible to introduce the λ variable, known as inverse Mills ratio, into the earnings equation, thus obtaining consistent estimates to the earnings equation parameters.[2]

In order to obtain the selection equation, a probit model is used that tries to analyze factors that influence an individual's probability to participate in

2 For a detailed presentation of Heckman's procedure (1979), see the reference work and Verbeek (2000).

the sample, i.e., to be employed with positive earnings in labor market. In this study, the selection equation's explanatory variables, according to works on the subject such as Kassouf (1994, 1997); Scorzafave & Menezes-Filho (2005) & Menezes, Fernandez e Dedecca (2005), are the same of those considered in modeling an individual's probability of being in the Economically Active Population (PEA). Thus, the selection equation of this study was:

$$L_i = \alpha_1 + \alpha_2 \text{PcHI}_i + \alpha_3 E_{1i} + \alpha_4 E_{2i} + \alpha_5 E_{3i} + \alpha_6 E_{4i} + \alpha_7 \text{Age}_i + \alpha_8 \text{Age}^2_i + \alpha_9 CD_{1i} + \alpha_{10} CD_{2i} + \alpha_{11} CD_{3i} + \alpha_{12} \text{Child}_i + \alpha_{13} R_{1i} + \alpha_{14} R_{2i} + \mu_i \quad (2)$$

where L is a binary dependent variable assuming a value 1 if the individual is employed with positive earnings; otherwise, the value is 0; αj (j = 1 to 14) are the parameters to be estimated. PcHI is the per capita household income of all sources, except that arising from the individual's work; E_k (k = 1 to 4) is a discrete variable indicating the individual's years of schooling, with base group formed by those with 0 years of schooling, and E_1, E_2, E_3 e E_4 formed, respectively, by those with 1 to 4, 5 to 8, 9 to 11 and over 11 years of schooling; Age is the economic agent's years of life; CD_l (l = 1 to 3) are dummy variables related to the individual's position in the household, being the base group formed by the head or reference person in the household; CD_1, partners; CD_2, children; and CD_3, other positions. Child is a binary variable assuming value 1 if there are children under 14 years of age in the household, and 0 otherwise. R_m (m = 1 to 2) are dummy variables which indicate the economic agent's race, with base group representing white individuals; R_1, blacks; e R_2, brown; and μ_i i is the random error term with average 0 and variance σ^2_u.

The earnings equation was based on the human capital theory, including some other variables, with the aim of controlling the wage differentials derived from sectors of activity and position and type of work and insertion in the labor market. Thus, it is possible to obtain a more accurate measure for wage differentials between the regions. Then, this is the earnings explanation model:

$$ln\ W_i = \beta_1 + \beta_2 E_{1i} + \beta_3 E_{2i} + \beta_4 E_{3i} + \beta_5 E_{4i} + \beta_6 Age_i + \beta_7 Age^2_i + \beta_8 \lambda_i + \beta_9 R_{1i} + \beta_{10} R_{2i} + \beta_{11} S_{1i} + \beta_{12} S_{2i} + \beta_{13} S_{3i} + \beta_{14} S_{4i} + \beta_{15} Poc_{1i} + \beta_{16} Poc_{2i} + \beta_{17} Poc3_i + \beta_{18} Poc_{4i} + \beta_{19} Poc_{5i} + \beta_{20} Poc_{6i} + \beta_{21} Toc_{1i} + \beta_{22} Toc_{2i} + \beta_{23} Toc_{3i} + v_i \quad (3)$$

where $ln\ W_i$ is the natural logarithm for hourly wages arising from the individual's main job; β_j (j = 1 to 23) are the parameters to be estimated; λ is the inverse Mills ratio; S_k (k = 1 to 4) are dummy variables related to worker's sector of activity, with industry as the base group; S_1, agriculture; S_2, construction; S_3, commerce and S_4, services; Poc_l (l = 1 to 6) are dummy variables that define the position of employment's main occupation, with the base group assumed to be regular, properly the registered worker; Poc_1, unregistered, irregular worker; Poc_2, military and statutory staff; Poc_3, properly registered domestic servants; Poc_4, unregistered domestic servants; Poc_5, self-employed worker; and Poc_6, employer; Toc_m (m = 1 to 3) are dummy variables indicating the kind of occupation, being Toc_1 for managers in general, Toc_2 arts and sciences professional and Toc_3 undergraduate technicians, with the base group formed by the other kinds of workers[3] ; and v_i is the random error term with average 0 and variance σ^2_u. The other variables are the same shown for equation (2).

As to the activity sectors considered in equation (3), the following aggregation were based on PNAD: i) in industry, other industrial activities and transformation industry were considered; and ii) in the services sector, the following activities were included: lodging and food; transportation, storage and communication; public administration; education, health and social work; domestic services; other collective, social and personal services; and other activities.

The estimation of the models presented must take into consideration the characteristics of the sample. In this case, due to the incorporation of PNAD's sampling

3 Administrative service workers, service workers, salespeople and commerce service suppliers, agriculture workers, worker in the production of goods and services, and repairing and maintenance services, members of the military and auxiliary forces and ill-defined or undeclared work.

and weight plan, the Maximum Likelihood Method (MLM) could not be used, with the assumption that the observations are the result of random independent and identically distributed processes. This hypothesis is not suitable for complex sampling data such as PNAD's,[4] especially when it concerns standard error calculation of the estimates and hypothesis tests. Thus, the adjustment of parametric models in this study was made by means of Maximum Pseudo-Likelihood Method (MPLM), which use in complex samples began in the works of Binder (1983), and was consolidated by Skinner; Holt and Smith (1989).

3.2 – Decomposition of the Wage Differential between the Concerned Regions

After discussing the procedures for obtaining the earnings equation, we can present the methodology which allows us to separate the difference in earned incomes between the two regions into two components: the first is related to productive differences and other aspects related to employed workers, while the second one refers to differences in earned incomes of these attributes in both markets. The most suitable methodology for this is the Blinder-Oaxaca Decomposition, developed by Oaxaca (1973) and Blinder (1973).

Although most applications of this methodology are concentrated in the literature of discrimination in labor market, it can be also used with other focuses. As Jann (2008b) highlights, the Blinder-Oaxaca decomposition can be used to study differences between groups for any result variable.

To achieve such procedure, firstly we need to estimate the wage determinants equations, as shown in expression (3), which can be represented in the following matrix form:

$$\ln W = X'\beta + v \qquad (4)$$

where $\ln W$ is the vector of the natural logarithm of the individual's earnings; X is the matrix of explanatory variables, including the intercept; and v is the random error vector.

The difference of ln of the mean earnings between workers in the RMBH and the RMS can be expressed as the difference of the linear prediction taken from the mean midpoint of the regressors in each group, as follows:

$$\begin{aligned} D &= E(\ln W_{BH}) - E(\ln W_{Sal}) = \\ &E(X'_{BH}\beta_{BH} + v) - E(X'_{Sal}\beta_{Sal} + v) \\ &= E(X_{BH})'\beta_{BH} - E(X_{Sal})'\beta_{Sal} \end{aligned} \qquad (5)$$

where $E(\beta) = \beta_I$ and $E(v_I) = 0$, by hypothesis, with the i index representing the group formed by workers in the RMBH or the RMS.

To identify the contribution of the differences of the average of the regressors – which indicate the productive characteristics of the agents and other aspects related to their employment in the market – and the difference of coefficients β_I – which are the returns, in terms of earnings, of such attributes – for D, Jones & Kelly (1984) and Daymont & Andrisani (1984) put forward that the expression (5) should be rearranged as following:

$$\begin{aligned} D = &\left[E(X_{BH}) - E(X_{Sal})\right]'\beta_{Sal} + \\ &E(X_{Sal})'(\beta_{BH} - \beta_{Sal}) + \\ &\left[E(X_{BH}) - E(X_{Sal})\right]'(\beta_{BH} - \beta_{Sal}) \end{aligned} \qquad (6)$$

According to Jann (2008b), the expression (6) is known as "three-fold" decomposition, i.e., the total differential between the RMBH and the RMS workers' earnings is divided into three components: explained or characteristic effect; unexplained or price effect; and interaction term.

The first term on the right of equation (6) represents the explained component or characteristic effect, indicating the wage differentials due to average differentiation of workers' positive and personal attributes and other aspects related to their insertion in each region's labor market. The region with the greatest amount of such attributes and with the greatest quality job offer in its economy will offer higher returns to their respective workers.

The second term is the unexplained component or price effect, which represents a wage differentiating

4 For more details on the complexity of PNAD data and the implications of these characteristics for estimates of econometric models, see Silva; Pessoa & Lila (2002) and Cirino (2008).

measure between the regions, regardless of respective averages of workers' work characteristics and personal (productive and non-productive) attributes. This term could represent a different valuation of the same characteristics and attributes between the markets in both regions. If, as expected, the coefficients of earnings equation for men and women in the RMBH were really higher than those found in the same equations for the RMS, there would be evidence of higher valuation of the workers' characteristics and employment in the former region at the expense of those in the latter, regardless of the working characteristics and personal attributes of the employed in each region. However, it is important to point out that, besides that "pricing" difference between the regions, which is here called a proxy of regional effect, the unexplained term also catches the potential effects of the differences of non-observed variables in the earnings equation, as pointed out by Jann (2008b) and Scorzafave & Pazello (2007).

Finally, the third component on the right of equation (6) measures the interaction between the differences in the averages of working characteristics and personal attributes, and the differences in the coefficients in between both regions.

Still regarding equation (6), we point out that it is formulated from the RMS workers' point of view, i.e., the differences of regressors between the groups to determine the characteristic effect are weighed by the coefficients of the RMS earnings equation. Likewise, for regional effect, the differences of coefficients are weighed by the averages of the RMS workers' explained variables, i.e., such component measures the expected change in the hourly wage of the latter, if they were equally paid as the RMBH workers, i.e., with the coefficients of such group.

Although there are several studies which have implemented the Blinder-Oaxaca decomposition, two aspects, especially in national terms, have not been duly exploited: the estimation of sampling variances of decomposition components and the uncertainty problem that arises from obtaining the portion of unexplained term due to the dummy variable groups.

According to Jann (2008b), most international studies which use such decomposition, give the point estimates of the coefficients only, without mentioning sampling variances and standard errors at all. Exception can be found in studies by Oaxaca & Ransom (1998), Horrace & Oaxaca (2001), Heinrichs & Kennedy (2007) and Lin (2007). In terms of national literature, this aspect is still barely exploited..

Procedures to obtain consistent and non-biased estimates of standard errors from Blinder-Oaxaca (1973) decomposition results can be found in Jann (2008a). Thus, it is possible to analyze the decomposition results not only by point estimates for the differentials found, but also by the dispersion measures of the latter. As Jann (2008b) pointed out, such possibility is important, since it is indispensable to obtain precision measures for the differentials found for a good interpretation of decomposition results and the performance of statistical inference.

As for the uncertainty problem that arises from obtaining the portion of unexplained term due to dummy variable groups, it arises from the fact that the decomposition result for categorical regressors depends on which base group was chosen, i.e., which category will be omitted (OAXACA; RANSOM, 1999; NIELSEN, 2000; HORRACE; OAXACA, 2001; YUN, 2005; SCORZAFAVE; PAZELLO, 2007). The effect of a categorical variance is usually modeled, including dummy variables in the equation to be estimated, assuming a value 0 or 1 for different categories, one of the latter to be called base group or reference group, omitted in order to avoid multicollinearity problems. (WOOLDRIDGE, 2006). Since the coefficients associated with each specific binary variable quantify differences in comparison with the base group, such coefficients change when the reference group is changed. Thus, the results of decomposition of price effect that measures the coefficient differences between the groups for each of the dummy variables are not invariant to the choice of different base groups.

As Oaxaca & Ransom (1999) noted, if the aim is to estimate only the total effects of decomposition expressed in (6), the problem of uncertainty for dummy variables is not relevant, since such total effects are invariant to the choice of reference groups of indicator variables. Similarly, Jann (2008b) argues

that, for the explained component, the unitary and total contribution for the decomposition of a dummy variables set representing an individual's given attribute is not affected by the choice of the base group either. However, for the unexplained component or price effect, the author stresses that the change in reference group not only changes the results of each unitary dummy, but also changes the total contribution to the binary variables group that expresses a given attribute.

In order to solve this problem, allowing an interpretation with economic meaning for detailed decompositions in the presence of binary variables, Yun (2005) put forward a procedure based on normalization of earnings equations for the estimation of decomposition coefficients. Such procedure consists in estimating the models for the groups using the same setting for dummy variables, and then transforming the vectors of the respective estimated coefficients, so that they are expressed in terms of deviations of a mean coefficient for each group of categorical variables. If such transformation is applied to the estimation of earnings equations, the detailed results of Blinder-Oaxaca decomposition, including those referring to dummies in the calculation of price effect, become independent of the choice of reference groups. Yun (2005) also emphasized that this procedure is the same as estimating earnings equations by varying the reference groups in order to obtain the estimation averages from different results to use them in Oaxaca-Blinder decomposition, with the advantage of estimating only one of the equation set for both groups.

To illustrate the procedure put forward by Yun (2005), take the following model:

$$Y = \beta_0 + \beta_1 D_1 + ... + \beta_{k-1} D_{k-1} + \varepsilon \quad (7)$$

where β_0 is the intercept; and D ($j = 1$ to k) are dummy variables representing a qualitative variable with k categories.

Assuming category k as the base group, model (7) can be alternatively formulated as follows:

$$Y = \beta_0 + \beta_1 D_1 + ... + \beta_{k-1} D_{k-1} + \beta_k D_k + \varepsilon \quad (8)$$

where β_k is restricted to zero. Since:

$$c = (\beta_1 + ... + \beta_k)/k \quad (9)$$

and defining:

$$\tilde{\beta}_0 = \beta_0 + c \quad e \quad \tilde{\beta}_j = \beta_j - c, \quad j = 1,...,k \quad (10)$$

the normalized model is then given by:

$$Y = \tilde{\beta}_0 + \tilde{\beta}_1 D_1 + ... + \tilde{\beta}_k D_k + \varepsilon, \quad \sum_{j=1}^{k} \tilde{\beta}_j = 0 \quad (11)$$

It is important to highlight that the normalized model is mathematically equivalent to the original one, i.e., both produce identical predictions.

As for the data base, we use the PNAD for the year 2006, with the sample used being comprised of men and women between 16 and 65 years of age. The choice of this age group was based on the Brazilian law related to minimum age to start working and retiring.

4 – RESULTS AND DISCUSSION

The results and discussion of this study are presented in two parts: In the first part, there is the analysis of the determinants of earnings in both markets. In the second part, the focus is the presentation and discussion of wage differentials verified between the RMBH and the RMS.

4.1 – Determinants of Insertion and Labor Earnings in the Metropolitan Regions of Belo Horizonte (RMBH) and Salvador (RMS)

The difference in the labor return between the RMBH and the RMS will be analyzed for both male and female workers. For this purpose, the initial sample for selection equation estimation comprised people between 16 and 65 years old, being 3,887 and 4,790 men, and 4,382 and 5,587 women for the RMBH and the RMS, respectively. It is important to point out that, since only employed people with positive earnings are considered in the earnings equations, the amount of observations used in these regressions and in the earnings decomposition is smaller than that observed in the selection equations, in which employed people with no payment, unemployed people and idle people have also been considered. Thus, the sample was reduced to, respectively, 3,068 and 3,458 men and 2,475 and 3,032 women in the RMBH and the RMS.

The results for the selection and earnings equations necessary to run the wage differential decomposition between the markets can be found in tables 1, 2 and 3.

Starting with the analysis of the selection equations, we noted that most variables were significant at 1%, stressing the importance of such regressors to explain the probability of participation of workers in the condition of employed people with positive income in the main job. As for the coefficients signal, they were as expected (table 1).[5]

For *per capita* Household Income variable, it was observed that it decreases the probability of both men and women be employed with positive income in their main jobs. This is because the greater the per capita household income is, the smaller will be the motivation and need of the individual to have a paid job. (Table 1).

As for the schooling indicator variables[6], we noticed that, for both genders, the greatest the schooling, the biggest are the chances of the individual be working. The reason for such behavior lies in the fact that education level is positively related to job opportunities and access to higher wages. (Table 1).

For 'age' variable, we noticed that it follows a quadratic behavior, result of the normal depreciation of the human capital with age. As for the indicator variables 'position in the household', it was perceived that, while no significant differences for men between breadwinners and partners with respect to the probability of being employed, the same was not confirmed for women. This can be explained by the fact that women usually take more intensely a secondary role in providing the household income in comparison with men, since household chores are traditionally associated with women. (Table 1).

Regarding the presence of young children, such variable has shown a different behavior between genders with respect to its impact on the probability of being employed. The explanation of the negative signal for women can be associated with the fact that bringing up and caring for children are still typically female activities, as stressed by Sanches & Gebrim (2003). However, the presence of young children eventually pushes men even harder to provide household livelihood, consequently driving them to the labor market.

Regarding the impact of race on the probability of employment in labor market, it was observed that generally its effect is not important for the estimated probability, except for women in the RMBH. For this region, such fact could reflect a greater need for work by black women, due to the fact that they generally occupy lower social classes than white women.

As to the wage determinants, it was noted that most variables was significant at 1%, which stresses the importance of such regressors to explaining the regressed concerned. It is important to highlight that the inverse Mills ratio (λ) was significant at 10% in 3 out of 4 earnings equation estimated, which stresses the importance of including such variable in order to eliminate the problem of sampling selection. Regarding the coefficient signals, they behaved as expected. (Tables 2 and 3)

Regarding the marginal effects to the 'schooling' variable, it was observed that, according to the theory of human capital, the higher the worker's schooling level, the higher was their earned income in the labor markets analyzed. In the four regressions, the marginal effect of schooling was positive and increasing, i.e., going from the base group comprised of worker without any level of schooling towards the most educated ones, we noticed that such effect is increasingly higher. Taking the earnings equation for the RMBH men as an example, while worker with 1 to 4, 5 to 8 and 9 to 11 years of schooling receive, respectively, 21.52%, 46.01% and 68.65% more than the base group, those with over 11 years of schooling receive, on average, wages which are 121.06% higher than those in the reference group. (Tables 2 and 3).[7]

5 The marginal effects shown in the selection equations indicate the impact, in percentage, of changes in each of the regressors on the probability of interest. Thus, taking the equation for men in the the RMBH as an example, the marginal effect of -0,0128 for per capita household income indicates that, if such variable increases R$ 100, the probability of the individual be employed with positive income decreases 1.28%.

6 For a binary variable such as schooling, the marginal effect of 0.0243 indicates that the fact worker has more than 11 years of schooling increases their chances of employment with positive income by 2.43% when compared with the base group (no schooling).

7 Marginal effect calculated as per the methodology put forward by Hoffmann & Kassouf (2005).

Table 1 – Selection Equations by Gender, in the RMBH and the RMS, 2006

Variables	RMBH				RMS			
	Coeff.	S.D.	P-v	M.E	Coeff.	S.D.	P-v	M.E
Man								
Constant	-1,0180	0,7023	0,15	–	-0,3244	0,6924	0,64	–
Net pcHI	-0,0004	0,0001	0,00	-0,0128	-0,0001	0,0002	0,70	-0,0007
1 to 4 years of schooling	0,2617	0,2647	0,32	0,0070	0,1813	0,3337	0,59	0,0014
5 to 8 years of schooling.	0,4103	0,2919	0,16	0,0110	0,4208	0,2965	0,16	0,0032
9 to 11 years of schooling	0,9077	0,3012	0,00	0,0246	0,5250	0,3112	0,09	0,0047
> 11 years of schooling.	1,9350	0,6125	0,00	0,0243	0,8575	0,3557	0,02	0,0039
Age	0,1577	0,0260	0,00	0,0005	0,2006	0,0338	0,00	0,0001
Age squared	-0,0020	0,0003	0,00	–	-0,0027	0,0004	0,00	–
Partner	0,1425	0,2451	0,56	0,0040	-0,2121	0,3355	0,53	-0,0025
Child	-0,2666	0,1874	0,16	-0,0096	-0,7003	0,1891	0,00	-0,0104
Other	0,5710	0,3732	0,13	0,0108	-0,5274	0,2838	0,06	-0,0093
Child <14 years of age	-0,0805	0,1527	0,60	-0,0026	0,0091	0,1914	0,96	0,0001
Black	-0,1550	0,2633	0,56	-0,0056	-0,5795	0,3199	0,07	-0,0078
Brown (mixed ethn.)	-0,1203	0,1503	0,42	-0,0039	-0,3244	0,6924	0,64	-0,0030
Woman								
Constant	-0,6085	0,5236	0,25	–	1,1443	0,6845	0,10	–
Net pcHI	-0,0002	0,0001	0,06	-0,0209	-0,0001	0,0001	0,08	-0,0043
1 to 4 years of schooling	0,4010	0,1748	0,02	0,0375	0,2572	0,2154	0,23	0,0070
5 to 8 years of schooling	0,5398	0,1884	0,01	0,0510	0,4318	0,2355	0,07	0,0115
9 to 11 years of schooling	1,1377	0,2009	0,00	0,1131	0,8176	0,2419	0,00	0,0284
> 11 years of schooling.	1,2961	0,2592	0,00	0,0825	1,0852	0,2800	0,00	0,0177
Age	0,1106	0,0222	0,00	0,0002	0,0304	0,0322	0,35	0,0003
Age squared	-0,0015	0,0003	0,00	–	-0,0003	0,0004	0,48	–
Partner	-0,4139	0,1295	0,00	-0,0505	-0,4825	0,1428	0,00	-0,0194
Child	-0,1281	0,1738	0,46	-0,0155	-0,3642	0,1863	0,05	-0,0151
Other	0,3227	0,2463	0,19	0,0299	-0,0583	0,3011	0,85	-0,0021
Child <14 years of age	-0,2229	0,0898	0,01	-0,0265	0,2074	0,1382	0,14	0,0067
Black	0,4371	0,1916	0,02	0,0386	-0,0279	0,1837	0,88	-0,0009
Brown (mixed ethn.)	-0,1355	0,0966	0,16	-0,0158	0,0218	0,1706	0,90	0,0007

Source: Research results.

Caption: Coeff.: Coefficients;
S.D.: Standard deviation
P-v: P-value;;
M.E.: Marginal Effect

As for age - used as a proxy of work experience - according to the theory of human capital, all equations indicated a parabolic relationship between it and the workers' wage. Thus, the impact of age on wage was positive, however decreasing, until the maximum point of the parabola formed by the relation of these two variables, from which the mentioned impact becomes negative.

For the 'race' variable, it was observed that, in both markets, the wages of blacks and brown (people of mixed ethnicity), all else being constant, are lower

Table 2 – Earnings Equations by Gender, in the RMBH, 2006

Variables	Men				Women			
	Coeff.	S.D.	P-v.	M.E.	Coeff.	S.D.	P-v.	M.E.
Constant	-0,7928	0,1440	0,00	-	-0,6723	0,1894	0,00	-
1 to 4 years of schooling	0,2200	0,0597	0,00	0,2152	0,1603	0,0776	0,04	0,1461
5 to 8 years of schooling	0,4677	0,0608	0,00	0,4601	0,3077	0,0860	0,00	0.2885
9 to 11 years of schooling	0,7034	0,0622	0,00	0,6865	0,5772	0,0882	0,00	0,5334
> 11 years of schooling	1,2280	0,0766	0,00	1,2106	1,2619	0,1073	0,00	1,2267
Age	0,0729	0,0069	0,00	0,0227	0,0544	0,0093	0,00	0,0144
Age squared	-0,0007	0,0001	0,00	–	-0,0006	0,0001	0,00	–
λ	0,3426	0,0812	0,00	0,3426	0,2411	0,1379	0,08	0,2411
Black	-0,1390	0,0326	0,00	-0,1352	-0,0872	0,0479	0,07	-0,1022
Brpwn (mixed ethn.)	-0,1185	0,0268	0,00	-0,1158	-0,1076	0,0279	0,00	-0,1017
Agriculture	-0,4742	0,1054	0,00	-0,4742	-0,0582	0,1196	0,63	-0,0582
Construction work	-0,1874	0,0363	0,00	-0,1874	0,2661	0,1308	0,04	0,2661
Trade	-0,1723	0,0357	0,00	-0,1723	0,0626	0,0543	0,25	0,0626
Services	-0,0552	0,0280	0,05	-0,0552	0,1311	0,0481	0,01	0,1311
Without formal contract	-0,2024	0,0334	0,00	-0,2024	-0,0872	0,0373	0,02	-0,0872
Mil. and statutory.	0,4315	0,0538	0,00	0,4315	0,2909	0,0508	0,00	0,2909
Domestic servant with formal contract	-0,3219	0,0874	0,00	-0,3219	-0,0872	0,0500	0,08	-0,0872
Domestic servant w/o formal contract	-0,3128	0,1778	0,08	-0,3128	-0,1480	0,0524	0,01	-0,1480
Self-employed	-0,0151	0,0394	0,70	-0,0151	-0,0767	0,0527	0,15	-0,0767
Employers	0,1202	0,0696	0,09	0,1202	0,3379	0,1109	0,00	0,3379
Managers in general	0,6451	0,0642	0,00	0,6451	0,3456	0,0790	0,00	0,3456
Science & Arts professional	0,6627	0,0670	0,00	0,6627	0,2922	0,0625	0,00	0,2922
Undergraduate technician	0,4088	0,0435	0,00	0,4088	0,2369	0,0439	0,00	0,2369

Source: Research results.
Caption: Coeff.: Coefficients;
S.D.: Standard deviation;
t: 't' statistics;
P-v P-value.

than those received by white people. These results are similar to the studies which suggest the occurrence of racial discrimination in labor market, such as Cavalieri & Fernandes (1998) for the Brazilian Metropolitan Regions, and Soares (2000) and Matos & Machado (2006) for Brazil.

As for the sector of activity, employment in industry and services were the most profitable for men in the markets. As for women, while higher wages are concentrated in the services in the RMBH, no significant difference was noticed in the RMS as to the earnings of female workers in different economic sectors. Also noteworthy is the fact that agriculture in metropolitan regions, as well as the number of women in construction work are insignificant, although such variables have been kept on the respective equations with the sole aim of keeping analytical homogeneity, using the same variables of the economic sector in all regressions.

As for employment positions, informality (work without a formal contract) has decreased the workers' wage, as can be seen through the negative and

Table 3 – Earnings Equations by Gender, in the RMS, 2006

Variables	Men				Women			
	Coeff.	S.D.	P-v.	M.E.	Coeff.	S.D.	P-v.	M.E.
Constant	-0,2495	0,1442	0,09	–	-0,4501	0,1672	0,01	–
1 to 4 years of schooling	0,1467	0,0571	0,01	0,1469	0,1919	0,0706	0,01	0,1894
5 to 8 years of schooling	0,2848	0,0582	0,00	0,2853	0,2512	0,0714	0,00	0,2471
9 to 11 years of schooling	0,5922	0,0571	0,00	0,5929	0,4933	0,0723	0,00	0,4831
> 11 years of schooling	1,2437	0,0725	0,00	1,2443	1,1339	0,0859	0,00	1,1263
Age	0,0523	0,0068	0,00	0,0193	0,0502	0,0074	0,00	0,0136
Age squared	-0,0005	0,0001	0,00	–	-0,0005	0,0001	0,00	–
λ	-0,0750	0,1061	0,48	-0,0750	0,1822	0,0520	0,00	0,1822
Black	-0,2033	0,0372	0,00	-0,2038	-0,1746	0,0373	0,00	-0,1743
Brpwn (mixed ethn.)	-0,1391	0,0356	0,00	-0,1400	-0,0969	0,0362	0,01	-0,0971
Agriculture	-0,4734	0,1217	0,00	-0,4734	-0,4537	0,1399	0,00	-0,4537
Construction work	-0,1492	0,0407	0,00	-0,1492	0,2387	0,1568	0,13	0,2387
Trade	-0,2583	0,0377	0,00	-0,2583	-0,1125	0,0629	0,08	-0,1125
Services	-0,1439	0,0342	0,00	-0,1439	0,0745	0,0536	0,17	0,0745
Without formal contract	-0,2891	0,0295	0,00	-0,2891	-0,1967	0,0345	0,00	-0,1967
Mil. and statutory.	0,3490	0,0463	0,00	0,3490	0,2914	0,0505	0,00	0,2914
Domestic servant with formal contract	-0,2182	0,0570	0,00	-0,2182	-0,2721	0,0336	0,00	-0,2721
Domestic servant w/o formal contract	-0,3756	0,1112	0,00	-0,3756	-0,4149	0,0390	0,00	-0,4149
Self-employed	-0,1514	0,0356	0,00	-0,1514	-0,2157	0,0450	0,00	-0,2157
Employers	0,4633	0,0839	0,00	0,4633	0,5235	0,1241	0,00	0,5235
Managers in general	0,5421	0,0658	0,00	0,5421	0,5371	0,0755	0,00	0,5371
Science & Arts professional	0,7272	0,0717	0,00	0,7272	0,4416	0,0585	0,00	0,4416
Undergraduate technician	0,3274	0,0445	0,00	0,3274	0,1525	0,0405	0,00	0,1525

Source: Research results.
Caption: Coeff.: Coefficients;
S.D.: Standard deviation;
t: 't' statistics;
P-v P-value..

significant coefficient of the dummy 'without formal contract' in the four equations. The same trend was observed in the comparison between domestic servant with and without formal contract, with the former having higher wages — coefficient of the variable 'domestic servant with formal contract' less negative that the one for the same kind of employment without formal contract. Still regarding domestic servants, we noticed that, even when it is formal, it showed lower wages when compared to other workers with formal contracts, being, respectively, 9% and 27% lower than those in the reference group in the women's equations in the RMH and the RMS[8] . Thus, with respect to wages, this kind of employment can be less unfavorable in the RMBH than in the RMS. As for the occupations with the highest wages, these were the military, statutory and employers. The explanation for the first case is due to the stability and occurrence of usually higher wage possibilities for this kind of occupation,

8 Domestic work is typically a female occupation, and was kept in men's equations with the sole aim of maintaining analytical homogeneity

being, in the second case, individuals constituted of company owners. (Tables 2 and 3).

Finally, the results related to the kind of occupation indicate, in the three markets, that other kind of workers receive lower wages than the managers in general, Science and Arts professionals and undergraduate technicians. Such result is associated with the fact that individuals grouped in one of these three last categories usually perform activities that demand higher skills, technical knowledge or professional qualification, in comparison with the other kind of occupation, and therefore tend to receive better wages.

4.2 – Wage Differential between the Metropolitan Region of Belo Horizonte (RMBH) and Salvador (RMS): Magnitude and Analysis of their Main Components

Once the selection and earnings equations for men and women in labor markets under investigation had been estimated, it was possible to perform the Blinder-Oaxaca (1973) decomposition, as described in the methodology section.[9] The aim is to estimate the effect of regional differences on labor market of the RMBH and the RMS. Besides the estimations of decomposition coefficients, the Stata 9.2 package provides the respective standard deviation and normalizes the earnings equations in order to solve the uncertainty problem of methodology associated with qualitative variables. The wage differentials between the regions were analyzed for both men and women, so that we could also compare such aspect between workers of different genders. The results for men can be found in Table 4, while the ones related to women are in Table 5.

Confirming the results found by Cavalieri & Fernandes (1998), it was noticed that the hourly wage is greater in the RMBH than in the RMS by 19.53% for men and 19.36% for women.

Decomposing the mean wage differentials of the two regions, it was noticed that both the characteristic effect and the regional effect contributed to increase the said differential between men and women.

The first effect, statistically significant at 5% for both genders, contributed, respectively, 51.34% and 50.62% to the total differential between the markets for men and women. Alternatively, the portion of the wage differential[10] between the RMBH and the RMS attributable to differences in the distribution of employed people's attributes and in the work characteristics in each of the regions, caused workers (both male and female) from the former region receive 9.59% (male)/9.37% (female) more than those in the latter. The regional effect, statistically significant at 1% for men and women, was responsible for 64.81% in the former group and 73.78% in the latter, of the logarithm gap of mean hourly wage between the RMBH and the RMS. This means that different dimensions, complexities and aspects of the labor markets in these regions make the mean work return of the RMBH superior, in comparison with the RMS, by 12.25% and 13.95% for men and women, respectively (Tables 4 and 5).

Proceeding to a more detailed analysis by variable groups for the performed decomposition, with respect to the characteristic effect, only the dummy variables related to race and position in the occupation were significant at least at 10%, for both genders.

The explanation for the fact that race-related variables have provided the main contribution of the characteristics effect for the wage differential between the RMBH and the RMS is due to the existence of a much higher rate of white workers employed in the former market, while the rate of blacks is much higher in the latter.[11] Since wages are comparatively higher in the former group (Tables 2 and 3), it was verified that the differences in the composition of race groups in the two regions, all else constant, caused the RMBH workers to receive an hourly wage 5.15%/4.26% higher than that received by a male/female employed worker in the RMS (Tables 4 and 5).

9 The Oaxaca routine in Stata 9.2 software was used to work out the decomposition.

10 The figures shown are calculated from the antilog of the coefficients.

11 For the sample used, in the RMBH, among men, 42.5% declared themselves as white and 12.1% stated they were black. In the RMS, though, black (32.1%) are predominant over white (15.4%). Among women, we notice the same trend: While the rate of white and black in the RMBH is, respectively, 45.2% and 11.7%, in the RMS black (30%) are predominant over white (17%).

Table 4 – Decomposition of the Difference of the Hourly Wage Logarithm between the RMBH and the RMS for Male Workers, in 2006

Differential of the expected value of the hourly wage logarithm	Coefficients	S.D.	"t" stat.	P>|t|	Hourly wage
RMBH	1,3948	0,0283	49,2400	0,0000	4,0343
RMS	1,2164	0,0289	42,1200	0,0000	3,3751
Difference	0,1784	0,0405	4,4100	0,0000	1,1953
Characteristic effect					**% Difference**
Schooling	-0,0022	0,0189	-0,1200	0,9070	-1,24
Age	0,0011	0,0053	0,2000	0,8400	0,60
Race	0,0502	0,0098	5,1100	0,0000	28,15
Position at work	0,0191	0,0062	3,0800	0,0020	10,71
Sector of activity	0,0051	0,0038	1,3600	0,1730	2,87
Kind of occupation	0,0183	0,0112	1,6300	0,1040	10,24
Total	0,0916	0,0383	2,3900	0,0170	51,34
Regional Effect					
Schooling	0,0323	0,0147	2,2000	0,0280	18,08
Age	0,4245	0,1729	2,4500	0,0140	237,93
Race	0,0032	0,0083	0,3800	0,7010	1,78
Position at work	0,0485	0,0372	1,3000	0,1930	27,18
Sector of activity	0,0268	0,0300	0,9000	0,3710	15,03
Kind of occupation	-0,0208	0,0323	-0,6400	0,5210	-11,64
Intercept	-0,3988	0,1885	-2,1200	0,0350	-223,55
Total	0,1156	0,0198	5,8400	0,0000	64,81
Interaction between the two effects					
Schooling	-0,0005	0,0042	-0,1300	0,8970	-0,30
Age	-0,0013	0,0025	-0,5100	0,6080	-0,71
Race	-0,0143	0,0120	-1,1900	0,2350	-7,99
Position at work	-0,0113	0,0039	-2,9200	0,0040	-6,34
Sector of activity	-0,0035	0,0035	-1,0100	0,3110	-1,99
Kind of occupation	0,0021	0,0026	0,8200	0,4130	1,19
Total	-0,0288	0,0137	-2,1000	0,0360	-16,14

Source: Research results.

As for the categorical variables related to position in the occupation, generally, the RMBH workers are concentrated in higher wage positions (worker with formal contract, military and statutory staff and employers) when compared to those in the RMS, and there are fewer of them in those occupations with lower wages (workers without a formal contract, domestic servant and self-employed). Thus, when the percentages of male/female workers in the former and in the latter wage groups in the RMBH are, respectively, 62.7%/50% and 37.3%/50%, the percentages in the RMS are 57.7%/43.2% for those in the first wage group and 42.3%/56.8% for those in the second wage group. Therefore, men earn, on average, 1.93% more in the RMS than in the RMS, whilst women earn 2.20% more.

As to the details of the regional effect for each variable group for men, we verified that the dummies related to schooling and age was statistically

Table 5 – Decomposition of the Difference of the Hourly Wage Logarithm between the RMBH and the RMS for Female Workers, in 2006

Differential of the expected value of the hourly wage logarithm	Coefficients	S.D.	"t" stat.	P>|t|	Hourly wage
RMBH	1,1933	0,0320	37,2900	0,0000	3,2980
RMS	1,0164	0,0295	34,4500	0,0000	2,7631
Difference	0,1770	0,0435	4,0700	0,0000	1,1936
Characteristic effect					**Difference %**
Schooling	0,0156	0,0184	0,8500	0,3970	8,81
Age	-0,0056	0,0038	-1,4600	0,1460	-3,16
Race	0,0417	0,0101	4,1300	0,0000	23,57
Position at work	0,0218	0,0093	2,3500	0,0190	12,31
Sector of activity	0,0027	0,0040	0,6800	0,4940	1,55
Kind of occupation	0,0134	0,0087	1,5300	0,1260	7,55
Total	0,0896	0,0378	2,3700	0,0180	50,62
Regional Effect					
Schooling	0,0230	0,0220	1,0400	0,2970	13,00
Age	0,1068	0,2082	0,5100	0,6080	60,36
Race	-0,0050	0,0082	-0,6100	0,5440	-2,81
Position at work	0,0163	0,0216	0,7500	0,4520	9,20
Sector of activity	-0,0527	0,0513	-1,0300	0,3050	-29,76
Kind of occupation	0,0511	0,0323	1,5900	0,1130	28,90
Intercept	-0,0091	0,2359	-0,0400	0,9690	-5,12
Total	0,1306	0,0279	4,6700	0,0000	73,78
Regional Effect					
Schooling	-0,0001	0,0041	-0,0300	0,9730	-0,08
Age	-0,0006	0,0011	-0,5500	0,5850	-0,35
Race	-0,0150	0,0142	-1,0600	0,2890	-8,50
Position at work	-0,0125	0,0056	-2,2100	0,0270	-7,05
Sector of activity	-0,0066	0,0040	-1,6500	0,0990	-3,76
Kind of occupation	-0,0083	0,0040	-2,0600	0,0400	-4,67
Total	-0,0432	0,0161	-2,6800	0,0080	-24,41

Source: Research results.

significant at 5%. In the first case, we noticed that, on average, male workers in the RMBH, all else constant, receive 3.28% more than those in the RMS, due to their years of schooling. As for the 'age' variable, which is a proxy of experience in labor market and whose contribution for the total differential (237.93%) was the most important of all, the wage difference of such attribute between the regions was responsible, all else constant, for an hourly wage 52.88% higher in the RMBH than in the RMS. (Table 4) This higher wage of productive attributes 'age' and 'schooling' in the RMBH could reflect two aspects of the regional labor markets: i) a higher economic concentration and agglomeration of productive activities in the RMBH tends to make its labor market pay their workers better wages than those in the RMS. Furthermore, the higher dynamism of the RMBH economy can somehow influence the higher wage of the 'age' attribute, since the years of experience in the labor market of this region tend to add more productivity to the worker than

those acquired in the RMS; ii) the existence, in general, of a greater number of formal jobs and more skilled occupations in the RMBH, compared with the RMS, may be causing the attributes in question to be more necessary to carry out activities in the first market, being, therefore, better paid.

Still on the workers' regional effect, we observed that the difference between the intercept of both regions was significant at 5%, accounting for an important contribution to the wage differential verified (-223.55%). This means that a white, male worker, without schooling or experience, working with a formal contract in the industry, in other kind of occupation, all else kept constant, would be better paid in the RMS than in the RMBH. However, it is important to point out that, according to Wooldridge (2006), most of the times, the intercept does not show a precise economic meaning, and the analysis generally focus on the coefficient of the explained variables. Thus, since the intercept in this study would also take into consideration the age equals zero, which does not exist in the data base used, as well as a set of barely significant characteristics in the sample, such parameter is set up here just as a mathematical component to the calculation of total wage differential between regions than as an important distinguishing factor between both markets.

Regarding the interaction term, although it has also been significant at 5%, its contribution to regional wage differential (-16.14%) was much lower than the other two effects, causing a small reduction in the gap (-2.84%) in favor of the RMS.

As to the details of regional effect for women, it was noticed that, although such effect has been highly significant in the aggregate, the same did not happen to the coefficients of decomposition by variables groups, due to the occurrence of high standard deviations associated with their estimations. Such exception made, it was noticed that, as happened to men, the most important variable to the regional effect was age once again, contributing to increase the mean hourly wage differential by about 11% in favor of the RMBH. With respect to the interaction term, also for women, it was significant, this time at 1%, being, however, less important to explain the wage gap between the metropolitan regions, in comparison with the contribution made by the characteristic and regional effects. The impact of the interaction term among women is to decrease the gap in favor of the RMBH by -4.23%. (Table 5).

5 – CONCLUSIONS

The difference in the level of development among Brazilian regions, especially with respect to the Southeast and the Northeast, raises the possibility of the existence of wage differentials between these two. Thus, this study quantified in average terms the value of the difference in wages received in the RMBH and that observed in the RMS, highlighting the main aspects which cause the occurrence of such gap between those two regions.

The findings indicate that, as expected, the hourly wage was higher in the RMBH when compared to that in the RMS for both genders. The explanation for such setting between the two regions reflected a combination of two aspects. The first was related to the fact that workers in the RMBH receive more because they had more of the most valued personal attributes in the labor market, especially with respect to race. As well as the workers in the RMBH find more favorable working characteristics than those workers in the RMS. As for the 'race' attribute, it was found out that, although racial discrimination was present in both of the markets, its impact was greater in the RMS, due to the fact that this region has a higher rate of black people among the employed. As for the second aspect, the gap between wages is due to the highest economic concentration and agglomeration of the RMBH that pays better for the same productive attributes, especially schooling and age, in comparison with RMS.

In terms of public policies and regarding the wage differential in the RMS, actions can be suggested that are aimed at increasing the economic dynamism of this region, such as public investment and provision of subsidized credit for productive activities, among others, in order to raise their income levels. However, since the productive restructuring process of the 1990s affected all the labor markets in the country and their precariousness is, in general, a common structural characteristic of all of them, policies must

be implemented which are aimed at improving working conditions and wages all over the country.

REFERENCES

BECKER, G. S. Investiment in human capital: a theorical analysis: part 2. **Journal of Political Economy**, v. 70, n. 5, p. 9-49, 1962.

BERNDT, E. R. **The practice of econometrics classic and contemporary**. 8. ed. Boston: Addison-Wesley, 1996. 702 p.

BINDER, D. A. On the variances of asymptotically normal estimators from complex survey. **International Statistical Review**, v. 51, n. 3, p. 279-292, Dec. 1983.

BLINDER, A. S. Wage discrimination: reduced form and structural estimates. **Journal of Human Resources**, v. 8, n. 4, p. 436-455, 1973.

BRAGA, T. S.; RODARTE, M. M. S. A inserção ocupacional e o desemprego dos jovens: o caso das regiões metropolitanas de Salvador e Belo Horizonte. **Pesquisa & Debate**, São Paulo, v. 17, n. 1, p. 103-123, 2006.

CAMPANTE, F. R.; CRESPO, A. R. V.; LEITE, P. G. Desigualdade salarial entre raças no mercado de trabalho urbano brasileiro: aspectos regionais. **Revista Brasileira de Economia**, v. 58, n. 2, p. 185-210, abr./jun. 2004.

CAVALIERI, C. H.; FERNANDES, R. Diferenciais de salários por gênero e cor: uma comparação entre as regiões metropolitanas brasileiras. **Revista de Economia Política**, São Paulo, v. 18, n. 1, p. 158-175, 1998.

CIRINO, J. F. **Participação feminina e rendimento no mercado de trabalho:** análises de decomposição para o Brasil e as Regiões Metropolitanas de Belo Horizonte e Salvador. 2008. 188 f. Tese (Doutorado em Economia Aplicada) – Universidade Federal de Viçosa, Viçosa, MG, 2008.

DAYMONT, T. N.; ANDRISANI, P. J. Job preferences, college major, and the gender gap in earnings. **Journal of Human Resources**, v. 19, n. 3, p. 408-428, Summer 1984.

DIEESE. **Mercado de trabalho metropolitano em 2008**. Avaible in: <http://www.dieese.org.br/ped/metropolitana/pedmetropolitanaAnual2008.pdf>. Access in: sept. 23 2009.

_____. **O mercado de trabalho metropolitano 1998-2005:** estrutura e dinâmica nas seis regiões brasileiras do Sistema PED. Avaible in: <http://www.seade.gov.br/produtos/ped/metropolitana/pdfs/analise_fev2007.pdf>. Access in: dez. 8 2007.

_____. **Pesquisa de emprego e desemprego**. Avaible in: <http://www.dieese.org.br/ped/ped.xml>. Acess in: oct. 20 2008.

GIUBERTI, A. C.; MENEZES-FILHO, N. Discriminação de rendimentos por gênero: uma comparação entre o Brasil e os Estados Unidos. **Economia Aplicada**, Ribeirão Preto, v. 9, n. 3, p. 369-383, jul./set. 2005.

HECKMAN, J. J. Sample selection bias as a specification error. **Econometrica**, v. 47, n. 1, p. 153-161, Jan. 1979.

HEINRICHS, J.; KENNEDY, P. A computational trick for calculating the Blinder-Oaxaca decomposition and its standard error. **Economics Bulletin**, v. 3, n. 66, p. 1-7, 2007.

HOFFMANN, R.; KASSOUF, A. Deriving conditional and unconditional marginal effects in log earnings equations estimated by Heckman´s procedure. **Applied Economics**, v. 37, n. 11, p. 1303-1311, jun. 2005.

HORRACE, W. C.; OAXACA, R. L. Inter-industry wage differentials and the gender wage gap: an identification problem. **Industrial and Labor Relations Review**, v. 54, n. 3, p. 611-618, Apr. 2001.

IBGE. **Banco de dados agregados:** SIDRA: Sistema IBGE de Recuperação Automática. Avaible in: <http://www.sidra.ibge.gov.br/pnad/default.asp>. Access in: nov. 10 2007.

JANN, B. Standard errors for the Blinder-Oaxaca decomposition. In: GERMAN STATA USERS GROUP

MEETING, 3., 2005, Berlin. **Proceedings**... [S.l.], 2005. Avaible in: <http://repec.org/dsug2005/oaxaca_se_handout.pdf>. Access in: sept. 15 2008a.

_____. **A stata implementation of the Blinder-Oaxaca decomposition**. Zurich: ETH Zurich Sociology, 2008b. (Working Paper, n. 5).

JONES, F. L.; KELLEY, J. Decomposing differences between groups. A cautionary note on measuring discrimination. **Sociological Methods and Research**, v. 12, n. 3, p. 323-343, Feb. 1984.

KASSOUF, A. Retornos à escolaridade e ao treinamento nos setores urbano e rural do Brasil. **Revista de Economia e Sociologia Rural**, Rio de Janeiro, v. 35, n. 2, p. 59-76, 1997.

_____. The wage rate estimation using the Heckman Procedure. **Revista de Econometria**, Rio de Janeiro, v. 14, n. 1, p. 89-107, Apr./Oct.1994.

LIN, E. S. On the standard errors of Oaxaca-type decompositions for inter-industry gender wage differentials. **Economics Bulletin**, v. 10, n. 6, p. 1-11, 2007.

MATOS, R. S.; MACHADO, A. F. Diferencial de rendimento por cor e sexo no Brasil (1987-2001). **Econômica**, v. 8, n. 1, p. 5-27, jun. 2006.

MENEZES, W. F.; FERNANDEZ, J. C.; DEDECCA, C. Diferenciações regionais de rendimentos do trabalho: uma análise das regiões metropolitanas de São Paulo e Salvador. **Revista Estudos Econômicos**, São Paulo, v. 35, n. 2, p. 271-296, abr./jun. 2005.

MINCER, J. **Schooling, experience, and earnings**. New York: National Bureau of Economic Research: Columbia University, 1974. 152 p.

NIELSEN, H. S. Wage discrimination in Zambia: an extension of the Oaxaca-Blinder decomposition. **Applied Economics Letters**, v. 7, n. 6, p. 405-408, Jun. 2000.

OAXACA, R. L. Male-female differentials in urban labor market. **International Economic Review**, v. 14, n. 3, p. 693-709, Oct. 1973.

OAXACA, R. L.; RANSOM, M. R. Calculation of approximate variances for wage decomposition differentials. **Journal of Economic and Social Measurement**, v. 24, n. 1, p. 55-61, 1998.

_____. Identification in detailed wage decomposition. **The Review of Economics and Statistics**, v. 81, n. 1, p. 154-157, Feb. 1999.

SANCHES, S.; GEBRIM, V. L. M. O trabalho da mulher e as negociações coletivas. **Estudos Avançados**, São Paulo, v. 17, n. 49, p. 99-116, set./dez. 2003.

SCHULTZ, T. W. Investiment in human capital. **American Economic Review**, v. 51, n. 1, p. 1-17, Mar. 1961.

SCORZAFAVE, L. G.; MENEZES-FILHO, N. A. Impacto da participação das mulheres na evolução da distribuição de renda no trabalho no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 35, n. 2, p. 245-266, ago. 2005.

SCORZAFAVE, L. G.; PAZELLO, E. T. Using normalized equations to solve the indetermination problem in the Oaxaca-Blinder decomposition: an application to the gender wage gap in Brazil. **Revista Brasileira de Econometria**, Rio de Janeiro, v. 61, n. 4, p. 535-548, Oct./Dec. 2007.

SILVA, P. L. N.; PESSOA, D. G. C.; LILA, M. F. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. **Cadernos de Saúde Coletiva**, Rio de Janeiro, v. 7, n. 4, p. 659-670, 2002.

SKINNER, C. J.; HOLT, D.; SMITH, T. M. F. **Analysis of complex surveys**. Chichester: John Wiley & Sons, 1989. 309 p.

SOARES, S. S. D. **O perfil da discriminação no mercado de trabalho:** homens negros, mulheres brancas e mulheres negras. Rio de Janeiro: IPEA, 2000. (Text for Discussion, n. 769).

STANLEY, T. D.; JARREL, S. B. Gender wage discrimination bias?: a meta-regression analysis. **Journal of Human Resources**, v. 33, n. 4, p. 947-973, Sept. 1998.

VERBEEK, M. **A guide to modern econometrics. Chichester:** John Wiley & Sons, 2000. 384 p.

WEICHSELBAUMER, D.; WINTER-EBMER, R. **A meta-analysis of the international gender gap**. Vienna: Institute for Advanced Studies, 2003. (Economic Series, n. 143).

WOOLDRIDGE, J. M. **Introdução à econometria:** uma abordagem moderna. São Paulo: Pioneira Thomson Learning, 2006. 684 p.

YUN, M. S. A simple solution to the identification problem in detailed wage decompositions. **Economic Inquiry**, v. 43, n. 4, p. 766-772, Oct. 2005