BRAZILIAN FEMALE LABOR MARKET: RACIAL-SKIN COLOR DISCRIMINATION AND INEFFICIENCY

Gustavo Andrey de Almeida Lopes Fernandes [*]

Resumo

O mito da democracia racial é ainda muito presente na sociedade brasileira, embora a literatura econômica aponte para a existência de um alto grau de discriminação racial. Este artigo visa inferir sua existência nos salários eficientes do mercado de trabalho feminino, levando-se em conta o problema do viés de seleção. As causas da ineficiência são também investigadas. Os resultados mostram que o diferencial de salário eficiente, não explicadas por diferenças no capital humano, é significativo, sendo maior entre as trabalhadoras mais eficientes. Por conseguinte, políticas tradicionais possuem apenas um efeito limitado para reduzir as diferenciais salariais entre raças.

Palavras-chave: Discriminação Racial; Discriminação de Gênero; Fronteira de Produção Estocastica

Abstract

The myth of racial democracy is still widespread in the Brazilian society, although economic literature has continuously documented pervasive racial discrimination. This study analyzes racial discrimination in the Brazilian female labor market using a Mincer stochastic wage frontier, corrected for sample selection. The results, using Oaxaca-Blinder decomposition, showed that an efficient wage differential, not explained by differences in human capital does exist and is large among the most efficient workers. Wage inefficiencies are also explained and it is a claim of the study that traditional policies might be a tool in reducing underpayment dispersion, but cannot cope with pure discrimination.

Keywords: Racial discrimination; Female Labor Market; Affirmative Action

---

[*] Professor Assistente do Departamento de Gestão Pública da EAESP-FGV. São Paulo, São Paulo, Brasil

1   Introdução

Racial conflicts are almost absent in Brazil and that might be the reason why Brazilians, in general tend to believe that discrimination is not a relevant problem. Freyre (1933), one of the most important Brazilian anthropologists, wrote a book containing the results of his lifelong research, in which he epitomized racial democracy in Brazil. That belief is still subscribed to by the Brazilian elite, which believes that racial democracy continues to prevail in the country [1].

This view has been seriously questioned in the economic literature. Henriques (2001), using the Pesquisa Nacional por Amostra de Domicilios (PNAD) — the national household sample survey carried out annually since 1976 by Instituto Brasileiro de Geografia e Estatística, IBGE, the Brazilian Bureau of the Census, presents strong evidence against Freyre's perception of an ideal racial democracy, as far as economic outcomes are concerned. His study has shown that in Brazil, Afro-Brazilian men earn only 40 to 50 percent of what white men earn. The average monthly income in all regions of Brazil was the highest for white men and the lowest for non-white women, with white women earning more than non-white men in all regions of the country.[2]

The Brazilian Census has a peculiar way of classifying people, when it comes to race. Actually, people are not classified by race; they are, instead, classified by color. The miscegenation of races resulted in a special color for some people who were classified as pardo by IBGE, meaning an individual of mixed race. Pardos are treated as persons of African descent. If we add the black and pardo people, Brazil has the largest African descent population in the world, outside Africa. For 2005, PNAD data indicated that 47.37 percent of the Brazilian population was white while black and pardo accounted for around 52.37 percent of the 187 million inhabitants of Brazil. This means the size of the black population, including persons classified as pardo[3], is second only to that of Nigeria.

However, if one looks at the distribution of income in Brazil, it is easy to see that the black constitute only about 26.7 percent of the population in the highest income decile, while they represent 59.76 percent of the population in the poorest one. Figure 1 shows the income distribution as per 2005 PNAD data. It clearly shows that the black population in Brazil has a strong presence among the poor and a very weak presence among the high-income brackets.

Barros et al. (2000) and Ferreira (2000) argue that differences in educational attainment are a major factor contributing to unequal distribution of income in Brazil. The presence of a large proportion of blacks among the poorest might be because education levels are very different between racial groups. The discrepancy in years of schooling between blacks and whites is 2.13 years, according to 2005 PNAD data. It might look relatively small but it is rather large when compared to the seven years of average schooling for

---

[1]Broadly speaking, racial democracy is the conception of an ideal society, of equals, in which the color of the skin presents no barriers for social mobility. The opposite conception is discrimination based on social classes. For a detailed description about the Brazilian racial debate, see Telles (2004).

[2]Average monthly income in 2005 was Real the Brazilian currency) 1077.27 for white men, 679.83 for white women, 549.29 for nonwhite men and 356.89 for non-white women (25 years or older).

[3]This approach will be followed in the rest of the article.

Figure 1
Population distribution by decile and by color - Brazil: 2005
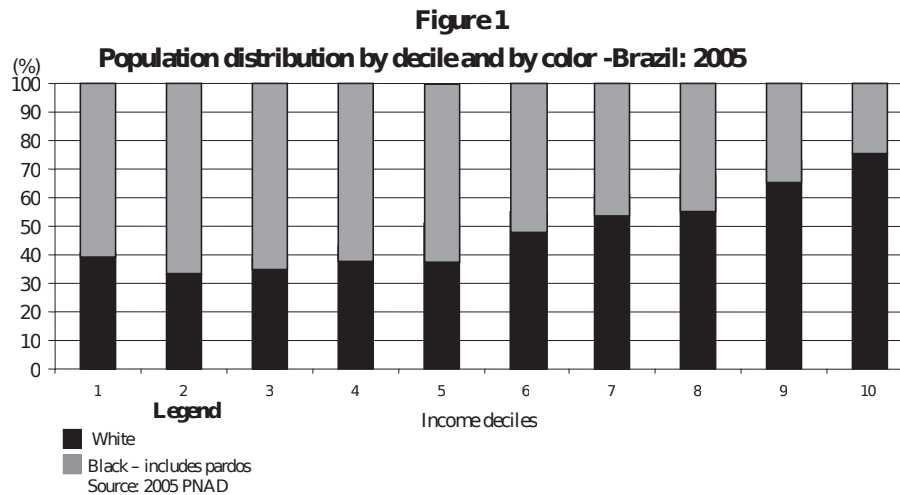
Figura 1: Population distribution by decile and by color

the entire population above 25 years of age. The picture becomes clearer when one looks at the evolution of education among groups. Defining a cohort as someone born between 1929 and 1974, average years of schooling for the whole population increased steadily, but the difference in years remained constant between blacks and whites throughout this period.

The differences between skin color groups become even more pronounced when a gender distinction is made. Kassouf (1998) shows that even after controlling for human capital, measured by educational attainment, a woman could make only 75 percent of a man's salary in Brazil in 1995. More recently, Carvalho et al. (2006) found that 97% of the difference between white men and black women wages is due to discrimination. Both papers, showing gender wage differentials favoring men, support the view presented in this study - i.e. if both race and gender differences are taken into account, then black women are likely to be at the bottom of the Brazilian socio-economic ladder. In fact, the average wages of black women in 2005, according to PNAD, were 64 percent of those of a black man and only 33 percent of the average salary of a white man.

Racial discrimination, from the perspective of economics, is measured by wage differentials unexplained by differences in contribution to economic outcomes, which are usually explained according to a scheme presented by Mincer (1970, 1974) that became known in the literature as the Mincer Equation. Oaxaca (1973) and Blinder (1973), working with Mincer's formulation, developed a methodology to decompose the wage gap among groups of persons into a first component attributed to individual characteristics (or human capital) that affect productivity and another component associated with discrimination. This framework has also been used to analyze variants of decomposition techniques by simulating counterfactual distributions, combining data on individual characteristics from one distribution with estimated parameters from another.

Zuchi & Hoffman (2004) and Zuchi (2006) show that wage differentials between white and blacks[4], unexplained by human capital assets as measured

---

[4]They also include *pardos* as part of the Black population.

by years of education and experience between whites and blacks, could be as high as 30.9 percent for 2003. Arcand & d'Hombres (2004) worked with data of the 1998 survey covering men in the 25-65 age group. Applying basically the Mincer Oaxaca framework but using a quantile Mincer-Oaxaca regression analysis, they encountered wage differential favoring male workers that goes from 38.95 percent for the first decile up to 48.1 percent for the last decile of the wage distribution. Their results show rising wage differentials as one goes up in the wage distribution scale and the wage gap. According to Arcand and d'Hombres, discrimination is more prevalent as the color of the skin gets darker. Matos & Machado (2006), also using PNAD data, applied the Oaxaca decomposition to Mincer equations and found that from 1987 to 2001, race and gender had a significant contribution to explain wage inequalities in Brazil.

The brief literature review has documented pervasive racial discrimination in Brazil and there is little doubt that dark skin is a handicap to social advancement. Our aim is to go even further, analyzing the Brazilian female labor market, so that discrimination will be the result of the Oaxaca decomposition in an estimated efficient Mincer wage function corrected for selection bias. Simultaneously, the factors that might contribute to pulling black female workers away from their respective Mincer efficient frontiers will be presented.

After this introduction, estimation procedures are in the second part detailed. Data is described in the following section. PNAD 2005, a nationwide household survey is used. In the fourth part, before the conclusion, results are presented. Racial discrimination seems to be very important in Brazil even when one looks to efficient wages, which are larger the more efficient the worker is. This analysis adds to the evaluation of policy instruments that could be used to improve efficiency as well as reduce wage disparities within wage groups, leading to improvement in the wage distribution. Moreover, if a wage gap attributed to racial discrimination is encountered, it can be argued that the effect of traditional remedies, very often proposed to eliminate or diminish discrimination, reach a bound given by the estimated value of pure form of discrimination. In this sense, even the most efficient black worked would still be discriminated.

2   Conceptual Framework

The methodology used to infer discrimination in the labor market relies heavily on the stochastic frontier model which has been traditionally used in the framework of studies concerning productive efficiency ever since Aigner et al. (1977) and Meeusen & Van Den Broeck (1977) defined and used the concept in a simultaneous and yet independent way. One pioneering application of this technique in labor economics was that of Robinson & Wunnava (1989). Other prior studies include Hofler & Murphy (1992, 1994). The former estimated the extent to which effective wages of workers are below the potential maximum they could earn, given their marginal productivity, while the latter estimated the worker's reservation wage. Both papers adopted the framework of search theory. McClure et al. (1998) also take up the question of wage inefficiency but compare Stochastic Frontier estimations for the United States and Canada. Dawson et al. (2001) analyze the question of relative underpayments

to male and female workers. Prieto et al. (2003) use the Stochastic Frontier wage model to estimate returns to education for potential and effective wages. They estimated the impact of formal schooling on workers potential maximum wages. Lang (2005), analyzing discrimination against migrant labor in Germany, claims that human-capital endowments determine earning potentials rather than actual earnings, and uses stochastic earnings frontier to estimate differences between natives and migrants for the year 2000.

This article adds to this body of literature by using the Oaxaca decomposition in a Mincer frontier wage function corrected for sample selection bias. It is our correction for sample selection that makes it possible to obtain consistent estimates of the Stochastic Frontier function, thus allowing an Oaxaca decomposition based on a consistent estimation procedure. Furthermore, by separating efficient from inefficient wages, the analysis allows a better understanding of the unexplained wage differential between groups of laborers.

2.1 The Stochastic Frontier

The starting point of our estimation procedure will be the definition of a Stochastic Frontier Mincer (log) earnings function, obtained from a set of variables, all of them reflecting their income generating inputs according to the following specification:

$$\ln(w_i^*) = x_i'\beta + v_i \tag{1}$$

Where w_i^* is the potential wage earned by the worker, x_i is a column vector of worker characteristics, and occupational and regional differences in labor markets, β is a vector of parameters associated with each variable to logarithm of potential or efficient wage earned by the worker and v_i a vector of parameters associated with each variable to logarithm of potential or efficient wage earned by the worker and:

$$\ln(w_i) = \ln(w_i^*) - u_i \tag{2}$$

where u_i is a random, non-negative variable.

The potential wage of each worker is the frontier wage or the efficient wage. Effective or observed wage is given by the Stochastic Frontier minus the non-negative random variable that captures the difference between potential and effective wage:

$$\ln(w_i) = x_i'\beta + v_i - u_i \tag{3}$$

The earning function is perceived as a frontier. It represents the potential or maximum wage that a worker would receive given her capital endowment as well as occupational and regional characteristics of the labor market. The non-negative random term u_i represents the extent to which observed earnings fall short of potential earnings.

The difference between efficient and observed wage a worker earns is explained by a set of variables reflecting individual and labor market characteristics, $S_i = (S_{i1}, \ldots, S_{im})$ and a new rannor term such that:

$$u_i = S_i'\delta + \eta_i \tag{4}$$

where $\delta' = (\delta_1,...,\delta_m)'$ is a vector of parameters to be estimated, $u_i \geq 0$ and $\eta_i(0, \sigma_u^2)$ are truncated at the point $-S_i'\delta$, in such a way that $\eta_i \geq -S_i'\delta$. Under these assumptions, u_i is a normally distributed variable with its mean depending on the specific explanatory variables reflecting individual and labor market characteristics and truncated at zero, such that $N^+(S_i'\delta, \sigma_u^2)$. Following Aigner et al. (1977), v_i is an idiosyncratic effect specific to an individual worker that could enter the model with either sign, such that we can suppose $v_i(0, \sigma_v^2)$, independent of and of the regressors.

The end result is the Mincer Stochastic Frontier model represented by equation 3 above. The frontier or potential wage is given by equation 1. The inefficiency term is given by u_i, a random variable of particular interest on this setting. Since the data is in log terms, u_i is a measure of the percentage by which the particular individual fails to achieve the particular potential ideal wage rate. The final Mincer Stochastic Frontier model (Equation 3) has a composed error term $\varepsilon_i = v_i - u_i$[5]

2.2   Sample Selection

Once a distribution for u_i is assumed, the next step is the definition of the likelihood function for ϵ_i and to obtain a set of parameters β and δ that maximizes the likelihood. However, there is an additional problem that has to be dealt with in order to obtain consistent estimates of parameters β: sample selection.

In this study, a two-step Heckman procedure will be used to take into account the selection process that might be present in our estimation of the Mincer Stochastic Frontier equation. In the first step of the Heckman procedure, the selection process, responsible for selection bias problems is studied with the so-called selection model. The bias is caused by existence of differences between employed and non-employed (hence earn no wages) women, which are related to their income. So it is necessary to compare these groups (employed and non-employed women) to find out what the differences are. For this purpose, generally, a probit model is estimated[6].

Following the economic literature, in our study, the dependent variable in the probit analysis is a dummy variable indicating whether or not the female is employed. Independent variables in the model are the relevant characteristics of the respondents, available in the data set, such as education, age, formal or informal jobs, number of children 10 years old or younger, and the total number of children in the household. In the probit analysis, we estimate the effects of these variables on the employment decision. Besides that, the effect of unmeasured characteristics which are related to the employment decision

---

[5]To complete the specification of the model and to allow estimation is required a specification of the distribution of the error term. Kumbhakar & Lovell (2003) suggest three possible distributions for the inefficiency term: exponential, half normal and truncated normal. We are going to use the truncated normal, differently from Lang (2005) who used the exponential distribution and afterwards explained the estimated inefficiency as function of explanatory variable in a two-step estimation procedure. We follow Prieto et al. (2003) argument favoring a joint estimation of the frontier parameter vector and of the parameter vector associated to the variables explaining inefficiency for. However, as shown by Greene (2002), the estimated inefficiency differs very little when different distributions are assumed for . The distribution will matter for the specification of the likelihood function to be used in the estimation procedure.

[6]We assume that the error term of this model is normally distributed, one of the assumptions underlying the Heckman model

process can be estimated through the residuals of the probit analysis. After all, the variation which remains in the dependent variable after removing the effect of the known factors can only be caused by the influence of unknown factors. That is the reason why in the Heckman (n.d.) procedure, the residuals of the selection equation are used to construct a selection bias control factor, which is equivalent to the Inverse Mills Ratio[7].

In the second step of the Heckman procedure, Stochastic Frontier analysis of the effects of human capital on the logarithm of wage is carried out, which we call the Mincer Stochastic Frontier equation. In this analysis, we use the Inverse Mills Ratio as an additional independent variable. Because we now have a control factor in the analysis for the effect of the logarithm of wage related unmeasured characteristics, which are also related to the employment decision, the human capital predictors in the Mincer Stochastic Frontier equation are free from this effect: the Mincer Stochastic Frontier analysis, therefore, produces unbiased coefficients.

2.3 Limitations

Before going to the description of the practical estimation of the Heckman model, a word of caution should be brought forward. Although theoretically the procedure sounds appropriate, applying it in practice is not so simple. An important condition for its use is that the selection equation contains at least one variable which is not related to the dependent variable in the substantial equation, otherwise severe problems of multicolinearity may arise and the addition of the correction factor to the substantial equation may lead to difficulties in estimation and to unreliable coefficients.

Formalization of the Mincer Stochastic Frontier model, corrected for sample selection, follows Greene (2003). Assuming that the sample selection follows a general framework, the equation that determines the sample selection process is[8]:

$$\ln(w_i^l) = z'_i + \xi_i \tag{5}$$

while the equation of primary interest is equation (2). The sample selection rule is that wage is observed only if $w_i^l > 0$, which in our analysis means that the observed wage is greater than the reservation wage, and the individual is participating in the labor force. Let us assume that v_i and ξ_i have a bivariate normal distribution with zero mean and correlation ρ, and that

$$E(u_i \xi_i) = E(v_i u_i) = 0 \tag{6}$$

Under those assumptions, the estimated Inverse Mills Ratio is introduced in the Mincer Stochastic Frontier as an additional regressor to correct for sample selection. The equation in the second step is estimated by maximum likelihood, with the Inverse Mills Ratio as additional regressor.

$$\ln(w_i | z_i^* > 0) = x'_i \beta + \lambda \frac{\phi(z'\gamma/\sigma_\xi)}{\Phi(z'\gamma/\sigma_\xi)} + v_i - u_i \tag{7}$$

---

[7]For a definition of the Inverse Mills Ratio, see Greene (2002).

[8]$\mathbf{Z}_i$ is a column vector of worker characteristics, and occupational and regional differences in labor markets. It shares the same variables as $\mathbf{X}_i$ and has at least one additional variable which must be not related to the dependent variable in the substantial equation.

This new model will henceforth be called the Mincer-Heckman stochastic earning frontier model. It can be estimated for each of the two groups of female workers: black and white. Furthermore, for each group of laborers, the parameter vector δ, associated with the mean vector of the truncated normal distribution, is jointly estimated. This estimation is carried in such a way that we can have an estimate of the percentage return on human capital for individuals on the efficient frontier, corrected for sample selection, as well as estimation of the contributing factor to inefficiency in each of the two female labor markets. Once the Mincer-Heckman stochastic frontier is estimated, the next step is to observe if wage differential between black and white still persists.

2.4 Discrimination

Economic discrimination is a broad and inaccurate concept, despite the existence of large differences between population groups which are elaborately empirically documented, showing great stability in due course. A widely accepted definition comes from the fact that identical objects have the same price. Thus, considering the workforce as a commodity, two individuals with identical productivity and in this sense equal, should have the same wage.

Discrimination occurs when workers with the same productive abilities but belonging to different population groups, earn different wages, invalidating the hypothesis of neutrality of trade. Therefore, economic theories that try to explain the phenomenon of discrimination aim at finding the precise conditions in the labor market in which identical goods are priced distinctly.

The job offer for two individuals is considered essentially identical if they display the same degree of productivity in the material process of production of goods. The idea of productivity in this context should be seen in a broad manner, including a variety of factors, such as work attendance, cooperation skills, and expertise — in addition to the actual expected future productivity, among other things.

Evidently, equality of economic agents makes sense only in relation to variables that are exogenous to the process of discrimination. Be a vector of productive characteristics that determine the productivity and observable by the business as well as being exogenous, a dummy that adopts the value 1 if the individual belongs to the segment of the measured population.

$$\alpha = (Y|X, Z = 1) - (Y|X, Z = 0) \tag{8}$$

In a fully egalitarian society α is equal to zero, where Y is the expected value of Y, conditioned by X, i.e.,all-or-equal, with the exception of the social group.

2.5 Oaxaca-Blinder Decomposition

Oaxaca-Blinder decomposition is used to estimate the wage earning differential that can not be attributed to differences in human capital endowment. As it was just shown, unexplained differences in wages between the two groups are usually attributed to racial discrimination. If separate wage regression equations are fitted for black and white female workers, the underlying regression models are:

$$\ln(w_{w,i}) = x_{w,i}{}'\beta_w + \varepsilon_{w,i}, i-1,...,n_w \tag{9}$$

$$\ln(w_{b,i}) = x_{b,i}{}'\beta_b + \varepsilon_{b,i}, i-1,...,n_b \tag{10}$$

where n_w and n_b represent black and white female samples, respectively.

Regressor vectors include socio-demographic variables, such as occupation, region, and human capital variables. A comparison of these two regressions allows us to perform what has been known in the literature as Oaxaca (1973) decomposition, which is based on a comparison of regression functions. For any two vectors of characteristics,

$$\ln(w_{w,i}) - \ln(w_{b,i}) = x'_{w,i}(\beta_{w,i} - \beta_{b,i}) + (x_{w,i} - x_{b,i})'\beta_b \tag{11}$$

The second term in this decomposition is identified with differences in human capital that would explain wage differences due to differences in human capital endowment between white and black female laborers, assuming that labor markets respond to these differences naturally. The first term shows the differential in log wages that is attributable to differences unexplainable by human capital; holding these factors constant at x_w makes the first term attributable to other factors, such as discrimination. Therefore, it is not easy to see that the first term can be used to estimate α:

$$\ln(w_{w,i}) - \ln(w_{b,i}) = x'_{w,i}\alpha + (x_{w,i} - x_{b,i})'\beta_b \tag{12}$$

In this study, the Oaxaca decomposition is performed using estimates generated by the Mincer-Heckman frontier functions for white and black female workers as estimators for β_w and . The difference between this procedure and Oaxaca's is that here, Mincer-Heckman Stochastic Frontiers are used to perform an Oaxaca style decomposition for estimating potential wage for both groups of women, while the conventional procedure uses OLS estimation or Heckman two stage estimation. If the hypothesis that is rejected, then the data shows evidence of wage differentials not attributable to differences in human capital between the two groups of laborers which can be considered an estimate of pure racial discrimination, based on estimated potential wage for black and white female workers, corrected for selection bias.

Dispersion of efficiency within each female group still remains to be explained. Inefficiency dispersion is analyzed with the estimated underpayment for black and white female workers, derived from the estimation of the Mincer-Heckman Stochastic Frontiers. Estimation of the frontiers follows the Kumbhakar and Lovell (2003) procedure, assuming that the random inefficiency term u, in Equation 6 follows a truncated normal distribution[9].

3   The Data

Our empirical analysis uses 2005 PNAD data. After data cleaning, when observations with any missing information were eliminated, the sample was reduced to 140,078 individuals, 25 years or older . Out of this sample, 61,183

---

[9]The normal-truncated normal model used in the estimation of the Mincer Stochastic Frontier and the resulting likelihood function is derived by in Kumbhakar & Lovell (2003). Once the log likelihood is maximized with respect to the unknown parameters, all the information necessary to estimate underpayment for each individual in the sample is at hand. The approximated formula to estimate the underpayment was developed by Jondrow et al. (1982).

were female workers who constituted the sample used in the study, with 30,489 classified as white and 30,551 black, based on the reported color of their skin. Table 1 and 2 present a summary description of variables for all female workers used for model estimation. Their average monthly wage was 519.95, average years of schooling was 9.0, meaning that the mean female worker was almost finishing high school (in Brazil it takes 10 years to finish high school). 12.2 percent of female workers had at least 15 years of schooling, which is required to complete some college study.

Tabela 1: Variables Description

Wage[1]	Wage in Reales
Educ	Years of Formal Education
Educ17[2]	Individual has at least under graduation degree
Exper	Years of experience in the labor market
SE	Individual lives in the Southeast Region
N	Individual lives in the North Region
NE	Individual lives in the Northeast Region
S	Individual lives in the South Region
CW	Individual lives in the Center West Region
Urban	Individual lives in an urban place
Maid formal	Individual works formally as a servant in someone's home
Maid informal	Individual works informally as a servant in someone's home
Formal Labor	Individual has a formal job
Age	In years
Unionized	Individual is a members of a union
Children	Number of Children
Indust. Sector	Individual works in the industrial sector
Child<10[3]	Individual has a less-than-10-year child
Dependent - Partner	Individual is the partner of the head of the family
Dependent - Son	Individual is the son of the head of the family
Dependent - Others	Individual is neither the partner nor the son of the head of the family

The average number of children was 1.8; 17.3 percent were working in domestic servant type of jobs that can be broadly described as "domestic assistance"; 4.3 percent had formal employment; and 16.7 percent were unionized. The northeast and southeast regions have high concentration of female workers and accounted for close to 60 percent of the female labor force. Northeast is the least developed region of Brazil and the southeast the most developed, as measured by per capita income. Thirty-nine point seven percent of black women had at least one child 10 years old or younger, while the figure for white female workers was 35 percent.

Table 2 also presents the same results, for white and black female workers, respectively. Average wage for white female workers was 759.14 reals while for black female workers it was 396.49 reals. A white female worker, on average, earns 1.9 times more money per month than an average black female. If one looks at education and experience, years of education of whites is 26 percent higher than for blacks. Higher education, defined as more than 15 years of schooling, is observed in 20.7 percent of whites in the sample while only 7.5 percent of blacks had higher education. Thus, white females have some edge in terms of human capital. If this is a possible explanation for the huge wage gap between white and black female workers, much still remains to be seen and we hope to shed some new light on this important social issue.

One important limitation for female labor is motherhood; raising small

Tabela 2: Data Summary

Variable	Female Labor Market		White Female Labor Market		Black Female Labor Market	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Wage[1]	514.946	958.755	759.143	1334.816	396.499	637.159
Educ	8.986	4.447	9.783	4.618	7.715	4.540
Educ17[2]	0.122	0.327	0.207	0.405	0.075	0.263
Exper	21.433	14.631	25.784	13.123	26.248	13.221
SE	0.290	0.454	0.347	0.476	0.245	0.430
N	0.123	0.329	0.066	0.249	0.172	0.378
NE	0.304	0.460	0.188	0.390	0.410	0.492
S	0.173	0.378	0.295	0.456	0.057	0.233
CW	0.110	0.312	0.103	0.304	0.116	0.320
Urban	0.819	0.385	0.124	0.329	0.183	0.387
Maid formal	0.042	0.200	0.038	0.192	0.056	0.231
Maid informal	0.131	0.337	0.091	0.288	0.153	0.360
Formal Labor	0.252	0.434	0.285	0.451	0.193	0.395
Age	36.276	13.333	41.157	11.336	40.584	11.262
Unionized	0.168	0.374	0.209	0.407	0.177	0.382
Children	1.778	1.288	1.622	1.124	1.924	1.410
Indust. Sector	0.125	0.331	0.139	0.345	0.111	0.314
Child<10[3]	0.339	0.473	0.350	0.477	0.397	0.489
Dependent - Partner	0.480	0.500	0.576	0.494	0.541	0.498
Dependent - Son	0.226	0.418	0.046	0.209	0.053	0.224
Dependent - Others	0.068	0.252	0.118	0.322	0.117	0.322
	(N=61,183)		(N=30,489)		(N=30,694)	

[1]Wages are measured in Brazilian Currency Reals; the average black market rate for the dollar was 2.7 Reals per dollar

[2]Indicates individuals with 15 or more years of schooling

[1]Women having at least one child 10 years old or younger

children demands intense mother care, making it difficult for women to participate in the labor market. As it has been shown by several studies (e.g., Heckman (n.d.), Kassouf & Silvia (2000) and Killingsworth & Heckman (1986)), the presence of small children imposes an upward drift in female reservation wage. If one omits this aspect of selection of sample of female labor supply, one can obtain only biased estimates of the effect of wages on the female labor supply. As it is possible to see in Table 2, 39.7 percent of black female workers have children 10 year old or younger; in case of white female workers, only 35 percent have small children. Another important fact evident from the table is that a high percentage of black women are working as maids. Maid is a typically unskilled labor job and it is a service used widely by Brazilian households. Maids represent 20.9 percent of black female jobs but only 12.9 percent of white female jobs.

This table reinforces what has already been said in the introduction. Black females are in the lower strata of the Brazilian labor market. Their wages, on average, are lower; their education is lower, and they are over represented in low skill, low wage jobs.

4   Results

4.1   Mincer Equation with Heckman Sample Selection Correction

10

The first part of the analysis deals with sample selection. The results of the estimation of the Mincer equation, corrected for sample selection, are presented in Table 3. It is possible to immediately see that the estimation of Mincer equation without correction for sample selection leads to inconsistent OLS estimates. Likelihood-ratio statistics strongly reject independence of the error term of the selection model from the error term of the wage equation for black, as well as for white female workers. The identifying restriction is the presence of young children in the household, preventing women from being employed if the woman is married. Under these circumstances there might be a larger probability for reservation wage to be above the market wage. This result follows previous findings in Brazil, analyzed by Kassouf (1998), and for other economies, as presented by Killingsworth & Heckman (1986), among others.

Presence of young children has a negative and significant impact on participation of black and white female workers in the labor market. Marriage also contributes to lower labor force participation for black as well as for white female laborers.

4.2   Mincer Stochastic Frontier

Estimation of the Mincer equation in a Stochastic Frontier framework has been used as an instrument to analyze underpayment, defined as:

$$\ln(\text{wage}_{ik}^{\text{obs}}) - \ln(\hat{\text{wage}}_{ik}^{l}) = \hat{u}_{ik}, i_k = 1,...,n_k; k = black, white \qquad (13)$$

where $\ln(\text{wage}_{ik}^{\text{obs}})$ is observed wage and $\ln(\hat{\text{wage}}_{ik}^{l})$ isestimated frontier wage, i. e. the maximum wage a female worker can earn given her human capital endowment; $\hat{u}_{ik}$ is a measurement of underpayment, i. e., how much more she would be able to make if she were getting efficient or frontier wage.

Table 4 shows the estimation of Mincer Stochastic Frontier without the correction for sample selection bias for black and white female laborers. Both estimated Mincer frontier present a somewhat larger return to education than the Mincer equation corrected for sample selection bias reported in Table 5[11]. One year of education, in the Mincer equation of Table 5, leads to a return of 5.9 percent for blacks while without correction for sample selection, the return would be 6.3 percent; for white workers, with the Mincer equation corrected for sample selection bias, one year of education had a return of 9 percent while in the Mincer Stochastic Frontier, the return was 9.2 percent.

---

[10]In a first analysis, Wald testes were carried with the null hypotheses that coefficients are the same for black and white women. For this, pooled regressions were estimated; a dummy for black and interaction dummies with all explanatory variables were created. In all cases, racial equality was rejected.

[11]The significance levels presented in this Table are corrected standard errors due to the use of estimated regressors as explanatory variable in this case the estimated Inverse Mills Ratio for white and black female workers.. To obtain a standard error asymptotically sound two procedures were carried out: The first was to bootstrap the standard error of all coefficients, and the second was to derive analytically the formula for the standard error, afterward deriving their sample values. The bootstrapped values were derived using 500 replications. For the analytical derivation of the standard errors see Murphy & Topel (2002), Greene (2002).

One additional year of superior education would result in return of 58.9 percent for blacks and 47.2 for whites. For black workers, one year of college education would yield a return that is ten times higher than one additional year of education below college level. For whites, the ratio between return of one year of college education and one year of education before college is around five.

The estimated coefficient for the Inverse Mills Ratio is highly significant for both groups of female laborers. This suggests that sample selection is an important factor that makes up for omission of a latent structure of a Heckman type selection.

Table 6 reports the regression results for underpayment. Underpayment means that workers with equal capital endowment earn wages smaller than frontier or potential wages, due to being separated by race. Inefficiency or underpayment decline with more education for both black and white workers. Formal labor relations improve efficiency, though the impact is stronger for black workers. Unionization is an instrument for white workers to reduce inefficiency and yet it is not an instrument for reduction of inefficiency for black workers. Children of any age are a drawback in terms of efficiency. It is very likely that children are a factor that makes women accept a smaller than potential wage. Maids, when informal, contribute positively to the underpayment of women, either white or black.

4.3 Oaxaca Decomposition Results

The main objective of this paper is to present a measure of pure racial discrimination, defined by the first term of the right hand side of equation 10, as discussed above. To achieve this, we used the estimated coefficients of the second stage Mincer-Heckman equations for white and black female workers. The value of pure discrimination is obtained using average sample values of the explanatory variables for the white workers group multiplied by the difference between the estimated frontiers' coefficients for white and for black female laborers. In other words, after taking into account differences in human capital endowment and frontier's estimated coefficients, an unexplained difference ranging from 19.6 to 14.8 percent still remains, when one goes from the highest efficient or potential wage percentile to the lowest.

In order to understand these differences, as we move along the frontier, we compare only efficient wages while separating the sample by percentiles. Table 7 presents the results. As we can observe in the table, Oaxaca difference unexplained by human capital tends to enlarge as efficient wages increase along the frontier. For the highest 1 percent wages, pure discrimination amounts to a 19.6 percent wage differential, while for the 90 percent percentile, the Oaxaca difference falls to 14.8 percent wage differential. Therefore, results show that even the most efficient black worker is discriminated in the Brazilian female labor market. Worse than that, the more efficient she is, the more discriminated she will be.

One could ask why this study separates black and white workers in labor markets. The answer is very simple: if estimated wage differences were small enough, and only individual human capital endowments were able to explain wage differentials, then one could treat all females, black and white, as a distinct labor market. However, if after taking care of differences in human capital endowments, unexplained difference is still large, then to treat

Tabela 3: Estimated Mincer Equations with Sample Selection

| | Black Female Worker | | White Female Worker | |
	Log Wage	Labor Force	Log Wage	Labor Force
Educ	0.087 (0.002)**	0.097 (0.003)**	0.103 (0.002)**	0.078 (0.003)**
Educ17	0.546 (0.025)**		0.475 (0.019)**	
Age	0.053 (0.003)**	0.070 (0.007)**	0.056 (0.004)**	0.071 (0.007)**
Age sq	−0.001 (<0.001)**	−0.001 (<0.001)**	−0.001 (<0.001)**	−0.001 (<0.001)**
N	−0.032 (0.017)	−0.313 (0.037)**	−0.044 (0.023)	−0.265 (0.051)**
NE	−0.371 (0.014)**	−0.195 (0.031)**	−0.313 (0.016)**	−0.256 (0.036)**
S	0.054 (0.024)*	−0.159 (0.053)**	0.049 (0.014)**	−0.311 (0.032)**
CW	0.107 (0.018)**	0.032 −0.045	0.087 (0.019)**	−0.096 (0.047)*
Married		−0.324 (0.056)**		−0.23 (0.065)**
Number Children at Home		−0.037 (0.011)**		−0.043 (0.014)**
Child < 10		−0.028 (0.031)		−0.063 (0.034)
Urban		1.248 (0.024)**		1.299 (0.028)**
Ind. Sector		0.772 (0.047)**		0.922 (0.049)**
Dependent — Partner		−0.704 (0.053)**		−0.729 (0.062)**
Dependent — Son		−0.426 (0.067)**		−0.414 (0.082)**
Dependent — Others		−0.165 (0.078)*		−0.394 (0.083)**
Constant	3.889 (0.078)**	−0.722 (0.169)**	3.798 (0.081)**	−0.591 (0.184)**
Observations	25545	25545	24432	24432
	LR test of indep. eqns. (rho = 0): chi2(1) = 123.85 Prob > chi2 = 0.0000		LR test of indep. eqns. (rho = 0): chi2(1) = 53.94 Prob > chi2 = 0.0000	

Standard errors in parentheses: * significant at 5%; ** significant at 1%

black and white females as different components of the markets is a correct stand and it is a sign that racial discrimination exists and has to be dealt with in empirical analysis.

In fact, results show that a form of pure discrimination still persists in the Brazilian female labor market. On average, there still remains an unexplained wage gap of 13.3 percent. We also encounter rising pure discrimination, when one moves along the efficient wage frontier. Consistent with these findings, the estimated Mincer-Heckman Stochastic frontiers have shown that education is a factor that improves potential wage for blacks as well as for whites; however, the impact of education is greater for whites than for blacks, such that education improves potential wages at a faster rate for whites than for blacks, which leads to increasing differences in potential wages as human capital increases.

Furthermore, when one looks at inefficiency in women labor markets, one can observe that several factors contribute to underpayment to both white and black female laborers in Brazil; children make women, black or white, accept less than potential wages; however, the impact of underpayment is more severe in case of black women. Unionization contributes positively to reduce underpayment of white women but it does not affect underpayment of

Tabela 4: Mincer-Heckman Stochastic Frontier Estimation Without Sample Selection Correction

	Black Female Worker Log Wage		White Female Worker Log Wage	
	SE	Bootstrapping	SE	Bootstrapping
Educ	0.063 (0.002)**	0.063 (0.002)**	0.092 (0.002)**	0.092 (0.002)**
Educ17	0.579 (0.020)**	0.579 (0.025)**	0.467 (0.017)**	0.467 (0.018)**
Age	0.034 (0.004)**	0.034 (0.004)**	0.040 (0.004)**	0.040 (0.005)**
Age sq	−0.001 (<0.001)**	−0.001 (<0.001)**	−0.001 (<0.001)**	−0.001 (<0.001)**
N	0.018 (0.014)	0.018 (0.014)	−0.029 (0.021)	−0.029 (0.020)
NE	−0.237 (0.011)**	−0.237 (0.011)**	−0.267 (0.015)**	−0.267 (0.016)**
S	0.006 (0.019)	0.006 (0.019)	0.015 (0.012)	0.015 (0.013)
CW	0.116 (0.015)**	0.116 (0.015)**	0.111 (0.017)**	0.111 (0.017)**
Union	0.245 (0.015)**	0.245 (0.016)**	0.183 (0.015)**	0.183 (0.016)**
Urban	0.066 (0.020)**	0.066 (0.019)**	0.076 (0.026)**	0.076 (0.023)**
Constant	4.670 (0.085)**	4.670 (0.084)**	4.362 (0.095)**	4.362 (0.102)**
Observations	20700		20643	

Standard errors in parentheses * significant at 5%; ** significant at 1%

Tabela 5: Mincer-Heckman Stochastic Frontier Estimation Results

	Black Female Worker Log Wage			White Female Worker Log Wage		
	SE	Corrected SE	Bootstrapping	SE	Corrected SE	Bootstrapping
Educ	0.059 (0.002)**	0.059 (<0.001)**	0.059 (0.002)**	0.090 (0.002)**	0.090 (<0.001)**	0.090 (0.002)**
Educ17	0.589 (0.020)**	0.589 (<0.001)**	0.589 (0.023)**	0.472 (0.017)**	0.472 (<0.001)**	0.472 (0.019)**
Age	0.031 (0.004)**	0.031 (<0.001)**	0.031 (0.004)**	0.037 (0.004)**	0.037 (<0.001)**	0.037 (0.005)**
Age sq	< 0.001 (<0.001)**	< 0.001 (<0.001)**	< 0.001 (<0.001)**	< 0.001 (<0.001)**	< 0.001 (<0.001)**	< 0.001 (<0.001)**
N	0.028 (0.014)*	0.028 (<0.001)**	0.028 (0.015)	−0.024 (0.021)	−0.024 (<0.001)**	−0.024 (0.020)
NE	−0.231 (0.012)**	−0.231 (<0.001)**	−0.231 (0.012)**	−0.262 (0.015)**	−0.262 (<0.001)**	−0.262 (0.015)**
S	0.011 (0.019)	0.011 (<0.001)**	0.011 (0.017)	0.022 (0.013)	0.022 (<0.001)**	0.022 (0.012)
CW	0.115 (0.015)**	0.115 (<0.001)**	0.115 (0.015)**	0.112 (0.017)**	0.112 (<0.001)**	0.112 (0.017)**
Union	0.244 (0.015)**	0.244 (<0.001)**	0.244 (0.016)**	0.182 (0.015)**	0.182 (<0.001)**	0.182 (0.016)**
Urban	−0.025 (0.028)	−0.025 (<0.001)**	−0.025 (0.028)	0.010 (0.036)	0.010 (<0.001)**	0.010 (0.033)
Blackmillsratio	−0.171 (0.036)**	−0.171 (<0.001)**	−0.171 (0.034)**	-I	−	−
Whitemillsratio	-	-	-	−0.120 (0.044)**	−0.120 (<0.001)**	−0.120 (0.042)**
Constant	4.862 (0.093)**	4.862 (0.001)**	4.862 (0.090)**	4.500 (0.108)**	4.500 (0.001)**	4.500 (0.117)**
Observations	20700			20643		

Standard errors in parentheses * significant at 5%; ** significant at 1%

Tabela 6: Regressions Results for Inefficiency Wage

	Black Female Worker Inefficiency	White Female Worker Inefficiency
Educ	− 0.202 (0.025)**	− 0.001 (0.023)
Child<10	0.434 (0.135)**	0.382 (0.203)
Urban	− 1.004 (0.165)**	− 2.001 (0.382)**
Maid Formal	−22.476 (4.511)**	−12.193 (3.579)**
Maid Informal	0.299 (0.127)*	2.336 (0.441)**
Formal Labor	−17.183 (2.726)**	−11.429 (2.700)**
Age	− 0.057 (0.030)	0.018 (0.042)
Age sq	0.001 (0.001)**	0.001 (0.001)
Union	0.802 (0.167)**	− 0.725 (0.301)*
Children	0.057 (0.046)	0.132 (0.080)
Ind. Sector	2.429 (0.257)**	3.117 (0.578)**
Constant	− 0.972 (0.840)	− 5.771 (1.802)**
Observations	20700	20643

Standard errors in parentheses * significant at 5%; ** significant at 1%

Tabela 7: Oaxaca Decomposition by Increasing Efficient Wage

Efficient Wage Quantile	Pure Discrimination (%)
0.01	0.196
0.05	0.158
0.10	0.143
0.20	0.137
0.30	0.143
0.40	0.153
0.50	0.155
0.60	0.153
0.70	0.151
0.80	0.150
0.90	0.148

black women. Education contributes to a reduction in underpayment (in other words, it improves efficiency), but again its impact is stronger for a white worker than for a black worker. Legal employment ties contribute strongly to improve efficiency and to reduce underpayment for both types of laborer, black and white. The occupation of maid without legal ties, contributes strongly to the underpayment of women; however, its impact is stronger in case of black women.

Finally, the paper attempts to contribute to the methodology, to the extent in which it takes into account the selection bias due to the fact that many women do not participate in the labor market if the reservation wage is above the market wage. We used a two-stage Heckman correction in estimation of the Mincer Stochastic Frontier. Oaxaca type analysis was conducted using the Mincer Stochastic Frontier estimated parameters such that we could derive a measure of discrimination, excluding inefficient wages from the analysis.

The end result of our analysis is that reduction of underpayment can be achieved by traditional policies and enforcement of labor laws. However, it might happen that part of the inefficiency could be due to the fact that only part of the workers might be hired under the present labor laws and those excluded from the legal labor markets have to survive by accepting underpaid jobs, such as maids, without formal ties. But our results indicate that even if traditional social economic policies are able to reduce underpayment, the improvement would be of a higher magnitude for white female workers than for black female workers. More importantly, even if wage underpayment is eliminated, there still remains a wage differential that we have called pure discrimination, which can not be eliminated by traditional policies; it is very likely that some form of affirmative action policies, with mandatory enforcement, will be required to eliminate this type of pure discrimination in female labor markets in Brazil.

Referências Bibliográficas

Aigner, D., C., L. & P., S. (1977), 'Formulation and estimation of stochastic frontier reduction models', *Journal of Econometrics* **6**, 21–37.

Arcand, J. & d'Hombres, B. (2004), 'Racial discrimination in the brazilian labour market: Wage, employment and segregation effects.', *Journal of International Development* **16**, 1053–1066.

Barros, R., R., H. & R., M. A. (2000), 'Estabilidade inaceitável: desigualdade e pobreza no brasil', Desigualdade e pobreza no Brasil. IPEA, Rio de Janeiro.

Blinder, A. S. (1973), 'Wage discrimination: Reduced form and structural estimates', *Journal of Human Resources* **8**, 436–455.

Carvalho, Néri & Nascimento. (2006), 'Diferenciais de salários por raça e gênero no brasil: Aplicação dos procedimentos de oaxaca e heckman', Pesquisas Amostrais Complexas. Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística. Mimeo.

Dawson, P., T., H. & D., W. (2001), 'German wage underpayment: An investigation into labor market inefficiency and discrimination', *Vierteljahrshefte zur Wirtschaftsforschung* **70**(1), 107–114.

Ferreira, F. H. G. (2000), Os determinantes da desigualdade de renda no brasil: Luta de classes ou heterogeneidade educacional, Technical report, IPEA, Rio de Janeiro.

Freyre, G. (1933), *Casa-Grande & Senzala: Formação da Família Brasileira sob o Regime de Economia Patriarcal*, Maia & Schmidt.

Greene, W. (2002), *Econometric Analysis*, Prentice Hall.

Heckman, J. (n.d.), 'Sample selection bias as a specification error', *Econometrica* **47**(1), 153–161.

Henriques, R. (2001), 'Desigualdade racial no brasil: Evolução das condições de vida na década de 90', IPEA. Available in www.ipea.gov.br/pub/td/td_2001/td0807.pdf . Cited 25 July 2007.

Hofler, R. A. & Murphy, K. J. (1992), 'Underpaid and overworked: Measuring the effect of imperfect information on wages', *Economic Inquiry* **30**(3), 511–529.

Hofler, R. A. & Murphy, K. J. (1994), 'Estimating reservation wages of employed workers using a stochastic frontier', *Southern Economic Journal* **60**(4), 961–976.

Jondrow, J., Lovell, C. A. K., Materov, I. S. & Schmidt, P. (1982), 'On the estimation of technical inefficiency in the stochastic frontier production function model', *Journal of Econometrics* **19**, 233–238.

Kassouf, A. L. (1998), 'Wage gender discrimination and segmentation in the brazilian labor market', *Revista de Economia Aplicada* **2**(2), 243–269.

Kassouf, A. L. & Silvia, N. D. V. (2000), 'Mercados de trabalho formal e informal: Uma análise da discriminação e da segmentação', *Nova Economia* **10**(1), 41–77.

Killingsworth, M. & Heckman, J. (1986), Female labor supply: A survey, *in* 'Handbook of Labor Economics', Ashenfelter, O. and Layard.

Kumbhakar, S. C. & Lovell, C. A. K. (2003), *Stochastic Frontier Analysis*, University Press.

Lang, G. (2005), 'The difference between wages and wage potentials: Earnings disadvantages of immigrants in germany', *Journal of Economic Inequality* **3**(1), 21–42.

Matos, R. & Machado, A. F. (2006), 'Diferencial de rendimentos por cor e sexo no brasil (1987-2001)', *Econômica* **8**(1), 5–27.

McClure, K. G., Girma, P. B. & Hofler, R. A. (1998), 'International labor underpayment: a stochastic frontier comparison of canada and the united states', *Canadian Journal of Regional Science* **21**(1), 87–91.

Meeusen, W. & Van Den Broeck, J. (1977), 'Efficiency estimation from cobb-douglas production functions with composed error term', *International Economic Review* **18**, 435–444.

Mincer, J. (1970), 'The distribution of labor incomes: A survey with special reference to the human capital approach', *Journal of Economic Literature* .

Mincer, J. (1974), *Schooling*, *Experience and Earnings*, edição edn, Columbia University Press.

Murphy, K. & Topel, R. (2002), 'Estimation and inference in two step econometric models', *Journal of Business and Economics Statistics* **20**(1), 88–97.

Oaxaca, R. (1973), 'Male female wage differentials in urban labor markets', *International Economic Review* **14**(3), 693–709.

Prieto, C. G., Román, A. M. & Domínguez, C. P. (2003), 'Actual and potential returns to schooling in spain. estudios sobre la economía española, eee 157', Available at SSRN: http://ssrn.com/abstract=381060.

Robinson, M. D. & Wunnava, P. V. (1989), 'Measuring direct discrimination in labor markets using a frontier approach: Evidence from cps female earnings data', *Southern Economic Journal* **56**(1), 212–218.

Telles, E. (2004), *Race in Another America: The Significance of Skin Color in Brazil*, Princeton University Press.

Zuchi, J. (2006), 'Desigualdade de rendimentos entre brancos e negros nos setores público e privado', UNESP. Dissertation.

Zuchi, J. & Hoffman, R. (2004), 'Diferenças de renda associadas à cor', *Pesquisa & Debate* **15**(1), 107–129.