INTER-REGIONAL AND INTER-INDUSTRY WAGE DIFFERENTIALS WITH INDIVIDUAL HETEROGENEITY: ESTIMATES USING BRAZILIAN DATA[1]

Ricardo da Silva Freguglia[2]
University of São Paulo – Brazil and Federal University of Juiz de Fora – Brazil

Naercio Aquino Menezes-Filho
University of Sao Paulo – Brazil and IBMEC Sao Paulo – Brazil

July 11, 2007

RESUMO

Este artigo utiliza um longo painel de trabalhadores brasileiros a partir da RAIS para investigar o que ocorre com os diferenciais de salários entre os estados brasileiros, entre ramos de atividade e entre ocupações após controlarmos pelas características não observáveis, constantes no tempo, destes trabalhadores. Como uma parcela substancial desses trabalhadores mudou de região, atividade e/ou ocupação no período amostral, é possível examinar em que medida esses diferenciais ocorrem devido à concentração de trabalhadores com alta habilidade em alguns estados/setores/ocupações. Os resultados mostram que os sinais dos diferenciais persistem na maioria dos casos, mas que sua magnitude se reduz em até 10 vezes, após levarmos em conta os efeitos fixos dos trabalhadores.

Palavras-Chaves: Diferenciais de Salário, Indústria, região, ocupação, dados em painel

ABSTRACT

This paper uses a long panel of Brazilian workers from RAIS to investigate what happens to the state, industry and occupation wage differentials, after we control for the unobservable characteristics of workers that are fixed over time. Since a substantial share of workers changed region, industry and/or occupation in the sample period, it is possible to examine to what extent these differentials occur due to the concentration of high skill individuals in some states/sector/occupations. The results show that, while the sign of the differentials persist in most cases, their magnitude drops by up to 10 times, after we take into account the workers' fixed effects.

Keywords: Wage Differentials, Industry, Region, Ocupation, Panel Data

JEL: J24

ÁREA 12 - Economia do Trabalho

---

[1] Paper prepared for the XXXV National Meeting of Anpec-2007.

[2] I am grateful to seminar participants of the 2007 Annual Meeting of Midwest Economics Association, the EcoMod-2007 - International Conference on Policy Modeling, the Regional Economics Applications Laboratory, and the Center for Latin America and Caribbean Studies at the University of Illinois for their valuable comments. Financial support from CAPES Foundation, data support from the Labor Ministry of Brazil, and computational support from Nereus and FIPE are gratefully acknowledged.

1. Introduction

This paper explores the inter-region, inter-industry and inter-occupation wage differentials of Brazilian workers using a longitudinal (panel) data set of nine years (1995-2003), controlled by observable and non-observable characteristics of individuals. The main idea is to evaluate the role of unobservable individual heterogeneity as an explanation for the earning inequality. The inclusion of non-observable individual characteristics, such as ability and motivation, in the model was implemented by using the fixed effects method. Special attention was given to the accuracy in the estimation results, with the use of Krueger and Summers (1988) methodology, hereafter KS, improved by Haisken-DeNew and Schimidt (1997), HDS, in order to provide the exact estimation of standard errors, the correct measure of wage dispersion and its overall variability.

In this paper we use a large panel sample of workers from Labor Ministry of Brazil observed at nine points in time (1995-2003). Since the data permits to track the same workers over the years, we are in the unusual position of obtaining the individual salary before and after the migration process. The most important results show that the wages among the three patterns of comparison adopted in the course of this paper – region, industry and occupation – have lower differentials than the previous OLS results. A large amount of the wage variance assumed as inter-region, inter-industry and inter-occupation differentials are a consequence of non-observable differences among individuals which could not be removed by a simple OLS estimation. Although we still confirm the empirical regularity of persistent earning differences, the size and significance of wage differentials decrease considerably.

Our results corroborate the unobserved heterogeneity hypothesis, whose explanation is based on a complementary approach to competitive theories. Its central idea is that the observed wage premium received by workers contains a return on unmeasured attributes. However, disregarding the non-observable features, the estimated wage differentials are biased upwards. This study contributes to this topic by explicitly addressing the role of unobserved as an explanation of observed differentials among regions, industries and occupations as well.

This paper is organized as follows. Section 2 briefly reviews the theoretical explanations for the wage differences, discussing the possible role of unmeasured abilities in explaining wage differentials. In section 3, we present the methodology used to work on data. We describe the fixed effects model and the two-step approach of HDS to estimate the wage differences. Data from the Brazilian Labor Ministry – RAIS-Migra – and some preliminary evidence are presented in section 4. Section 5 movies on to provide our basic econometric results using fixed effects estimation compared with other concurrent models. We present evidence suggesting that inter-region, inter-industry and inter-occupation wage differentials are significantly explained by non-observed abilities. Section 6 offers concluding comments.

2. Brief review of the theoretical explanations for the wage differences structure

As emphasized in the labor economics literature, the pattern of wage differentials has a high stability level, and these differences hold in several countries with unequal institutions and structures (Gittleman and Wolff, 1993; Kahn, 1998). Moreover, wage differentials among workers persist, even after they are controlled by human capital, occupation and other variables (Krueger and Summers, 1988; Teal, 1996). This suggests that wage differentials are inherent to capitalist economies. In this regard, the support of simple competitive theories, which claim that changes in wage dispersion should naturally be transitory, is questionable.

One approach based on a complementary explanation to the competitive theories is the unobserved heterogeneity hypothesis. The central argument is that employers consider the measured abilities of workers, such as schooling and experience at work, when hiring and monitoring workers. However, an industry that pays relatively high wages to workers regardless of their human capital characteristics also considers the unmeasured skills of the individuals as an important issue. Employers providing higher wages and conducting a more accurate observation about workers than that of the researcher[3] assure that industry affiliated workers are evaluated as having higher unmeasured productivity attributes (Jackubson, 1991). The industry wage premium contains a return on these unmeasured attributes and the estimated inter-industry wage effect is biased upwards. Therefore, the inter-industry wage differentials can be explained by workers with various levels of measurable and unmeasurable abilities (Arbache, 2001).

Several studies (Murphy and Topel, 1987; Krueger and Summers, 1988; Gibbons and Katz, 1992; Keane, 1993; Shippen, 1999; Abowd *et al.*, 1999) have approached the non-observable heterogeneity in wage determination. However, its importance cannot be considered as representing a consensus view. KS, for example, present little evidence to support the importance of unobserved heterogeneity in the determination of industry pay. Gibbons and Katz (1992) provide an experiment that does not rule out an unobserved ability explanation, but another experiment is sympathetic to the KS thesis that true industry differentials exist across industries, even for identical workers. In contrast, Murphy and Topel (1987), Keane (1993), and Abowd *et al.* (1999) find that unobserved heterogeneity explains 66%, 84%, and approximately 90%, respectively, of the apparent differential in log-wages across industries. Carruth *et al.*(2004) reach a similar conclusion in a recent study based on industry switchers in the United Kingdom. They explicitly address the role of unobserved heterogeneity as an explanation of observed inter-industry differentials, finding that unmeasured abilities explain 90% of inter-industry wage differences.

This kind of investigation is even more interesting when the level of income inequality is high. That is the case of wage dispersion in developing countries, whose literature has been concentrated on the effects of both human capital variables (Corbo and Stelcner, 1983; Lam and Levinson, 1992; Psacharopoulos and Velez, 1992; Yamada, 1996) and different sources of segmentation associated with the institutional arrangements and structural characteristics of these economies on earnings (Harris and Todaro, 1970; Tokman, 1983; Castells and Portes, 1989; Lindauer and Sabot, 1983; Macedo, 1985; Fields and Wan, 1989; Teal, 1996; Morrison, 1994; Arbache and Carneiro, 1999). This wage dispersion pattern can also be seen in the Brazilian labor market, of which the high degree of wage inequality is a distinguishing aspect, particularly when compared with other countries that are at a similar development stage.

The true identification of the nature and the source of inter-region wage differentials in Brazil has been the target of several studies as a result of its great income inequality (Langoni, 1973; Bacha and Taylor, 1978; Barros and Mendonça, 1995; Cowell et al., 1996; Pinheiro and Ramos, 1994; Gatica et al, 1995). Arbache (2001) compared and tested the most prominent competitive and non-competitive theories on wage differentials using Brazilian data. His results confirm that unmeasured abilities and efficiency wage models seem to play a role in wage determination, while compensating differentials and transitory difference theories were found to be irrelevant to wage formation.

All things considered, the identification of the observed inter-industry wage differentials is important on the grounds of both research and policy, and also with regard to individual welfare. In order to expand this analysis to achieve a better understanding of this topic, our aim is to investigate the role of unobserved heterogeneity as an explanation of the observed inter-region, inter-industry, and inter-occupation differentials. We use the panel data approach considering the fact that, if the unobserved productivity differences are constant over time, wage differences can be estimated by a fixed-effects model. We also consider the HDS approach, showing explicitly that much of the observed cross-section inter-region/inter-industry/inter-occupation wage differentials can be accounted for by observed and unobserved individual-specific effects.

[3] It is important to stress that the econometrist cannot observe, by definition, the unmeasured abilities in the dataset. As a consequence, these productive characteristics of workers are not accounted for in ordinary econometric estimations.

3. Methodology

We implemented the analysis of wage differentials as described by HDS, expanding this study by including regressed log wages not only on a group of industry dummies, but on regional and occupational dummies as well. The HDS procedure improved the standard method first disseminated by KS.

In their seminal paper, KS consider regressed log wages on a group of *k*-1 industry dummies, in addition to a constant and other control variables, using standard cross-section wage equations. The kth dummy is the omitted reference dummy which corresponds to the constant term.[4] Further, they renormalize the estimated coefficients for industry differentials to express industry differentials as deviations from a hypothetical employment-share weighted mean, where n_j is the share of industry j, and K is the number of industries. Instead of calculating the standard errors of the renormalized coefficients, KS suggest approximating them by the standard errors of the coefficients in the original regression, and using the standard error of the constant term as an approximation for the standard error of the omitted industry.

HDS argue that the above procedure overstates both the standard error of renormalized coefficients and their variance. They also demonstrate empirically that the estimated standard errors vary considerably depending on the choice of omitted industry, irrespective of sample size. Such variation inevitably inhibits sensible economic interpretation of individual elements of the renormalized coefficient vector $\left(\hat{\varphi}_j^r\right)$ and the estimated summary measure of overall wage dispersion (SD(φ)). HDS also show that the overall measure of industry wage dispersion is always underestimated using the KS methodology. As described in equations (1) and (2), the HDS procedure provides economically sensible coefficients and their correct standard errors in a single regression step.

$$\ln w_{ij} = \beta X_i + \varphi Z_j + \varepsilon_{ij} \tag{1}$$

$$\hat{\varphi}_j^r = \hat{\varphi}_j - \sum_{j=1}^{K-1} \hat{\varphi}_j^* n_j \tag{2}$$

Therefore, using the two-step approach or, equivalently, the restricted least squares (RLS) procedure of HDS, all k dummy coefficients and standard errors are reported, i.e., all results calculated are independent of the choice of the reference category. This procedure corrects the problems of the KS methodology of overstated differential standard errors and understated overall dispersion. The coefficients are interpreted as percent-point deviations from the region/industry/occupation weighted average. An overall measure of dispersion is also reported (SD).

$$SD(\varphi) = \sqrt{\sum_j n_j \varphi_j^2 - \sum_j n_j \sigma_j^2} \tag{3}$$

First, we estimated the wage differentials equations to Brazilian workers using level regressions to each of the nine years (1995 to 2003) from our sample. The aim is to describe the wage differences at a first glance, permitting a further comparison with other models that consider pooled observations and person effects. The functional form of these regressions are based on the Mincerian equation (Mincer, 1974), expanded by a set of other explanatory variables:

$$\ln w_{it} = \alpha + \beta X_{it} + \delta R_{it} + \varphi I_{it} + \gamma O_{it} + \varepsilon_{it} \tag{4}$$

[4] Suits (1984) argues that this sort of specification may not be appropriate for economic interpretation in many cases, particularly when there is more than one set of dummy variables in the model, as is common when estimating inter-industry wage differentials (Lucifora, 1993; Arai, 1994; Gera and Grenier, 1994).

where ln w_{ij} is the natural logarithm of the real wage of worker i in year t, α is the constant term, X_i is a vector of control variables, including age, age squared, tenure, tenure squared, education level dummies and gender dummy, R_{it} is a vector of region dummies, I_{it} is a vector of industry dummies, O_{it} is a vector of occupation dummies, β, δ, φ and γ are vectors of parameters to be estimated, and ε_{it} is a random disturbance term.

Second, we found the best model for our data. Initially, we used pooled ordinary least squares including all years jointly, as a large cross-section. The main problem in this case is the possibility of correlation between ε_i and the explanatory variables, nullifying the following hypothesis:

$$E(\varepsilon i|Xi) = 0; E(\varepsilon i|Ri)=0; \quad E(\varepsilon i|Ii)=0; \quad E(\varepsilon i|Oi)=0 \tag{5}$$

Without adopting this hypothesis, the causality relationship cannot be maintained. As a result, an endogeneity[5] problem might appear, generating inconsistent and biased estimators. As we have a longitudinal panel data of workers, this endogeneity problem might be solved by considering the unobserved heterogeneity, i.e., using the random effects or the fixed effects methods. When considering the latter, the unobserved characteristics of the workers, that are constant through the years, are subsumed by the fixed effect (c_i):

$$\ln w_{it} = \alpha + \beta X_{it} + \delta R_{it} + \varphi I_{it} + \gamma O_{it} + c_i + \varepsilon_{it} \tag{6}$$

The identification hypothesis of the model in equation (6) requires that the correlation among R_i, I_i, O_i and ε_i is provided by an explanatory variable that does not vary among years. The estimation could be done by OLS or first differences if we include a dummy variable for each worker. However, it is important to highlight that some individuals need to change from a region/industry/occupation to another in order to identify the coefficients of wage differentials after the inclusion of fixed effects.

4. Data and Initial Evidence

Our empirical analysis on wage differentials is based on longitudinal micro-data from the Relatório Anual de Informações Sociais – Migração (RAIS-Migra), of the Labor Ministry of Brazil, for the nine years between 1995 and 2003. This data base is derived from RAIS, an annual administrative survey that makes available information to identify workers elegible to receive social benefits and to monitor the labor market. It also provides extensive employment coverage, about 22 million of workers a year, besides a rich source of economic information at the individual level.[6] In this regard, it can be considered as a labor census of formal employment. Informal employment and illegal activities are not recorded by the RAIS census.

The aim of RAIS-Migra is to follow longitudinally the professional course of workers by region, industry and occupational features in the labor market, providing the conduction of studies on migration of workers in the labor market. The investigation of the wage structure, which considers wages deflated by the IPCA (Índice de Preços ao Consumidor – Amplo), is carried out for the 26 Brazilian states plus the Federal District, the eight sectors of the economy classified by the Instituto Brasileiro de Geografia e Estatística (IBGE), and the six occupation categories based on the Brazilian Occupation Classification (CBO)[7]. The other independent variables are age, tenure (monthly), gender and nine educational levels, following the Labor Ministry classification. The micro-data sample is composed of workers between the

[5] In general, we do not have information about the worker potential to obtain wages. As a consequence, these unobservable characteristics – ability, motivation, creativity, etc. – might be correlated to other wage determinants, such as education, region, occupation, etc, nullifying the causal interpretation of the estimated coefficients. See Wooldridge (2002).

[6] See Arbache (2001) for further explanation about RAIS data base.

[7] (1) Scientifical, technical and artistical, (2) Legislative, executive, judiciary, public sector and directors, (3) Managerial, (4) Trade and services of tourism and embellishment, (5) Farming, forestry activities and fishing, (6) Bluecollars.

ages of 14 and 65, with non-zero monthly income. Initial evidence about the data used in this study are presented in Tables A1-A3 of the Appendix. These three tables are organized to show wage differentials and the other explanatory variables according to the three levels of analysis, i.e, regions, industries and occupations.

Due to the large number of observations available, a random sample was generated in order to run the wage differentials regressions. This sample draws 1% of the total number of individuals from the original data base. In this regard, the generated sample has 580,005 pooled and balanced observations, with 64,445 individuals by year. The percentage of movers among regions, industries and occupations is about 1% (4,353 individuals), 5% (21,418 individuals) and 8% (40,763 individuals), respectively

5. Results and discussion

Our results are organized in the following way. The preliminary analysis of the data is based with the level regressions on logarithm of real wages for each year between1995 and 2003. Each cross section has 64,445 individuals. The second sort of analysis is based on the comparison among three concurrent models: OLS, random effects and fixed effects. In this subsection, we account for inter-region wages differences. This approach is adopted in the first subsection and is followed as a benchmark in the further sections, 5.2 and 5.3, that consider the inter-industry and the inter-occupation differences.

5.1. Inter-region wage differences

5.1.1. Level regressions

Table 1 reports wage differentials among workers and the overall variability of these differences. The main contribution of this table is to describe the behavior pattern of wage differentials at a first glance. It can be seen that six regions might be underlined as prominent regions, i.e., those states in which the wage differences are above the average in all the years: DF, RR, SP, AP, RS and RO. A way to verify whether the wage differentials change significantly over the years is to observe the difference between the regions of higher and lower wage differences. For example, workers from the region that presents the largest wage differential – DF – earn 70% (in 2003) above the average wage.[8] In contrast, fourteen regions have wage differences consistently below the mean during the nine years: AC, PA, MA, PI, CE, RN, PB, PE, AL, SE, BA, MG, MS, MT, GO. We can illustrate this with the workers from the region of PB, whose wage differentials of 59% (in 2003) below the average wage figure are the lowest. Further, the results within each year do not change in important ways.

Another way to confirm whether the wage differentials change significantly over the years is to verify the overall variability of the wage differentials among regions. As we can see in table 1, this overall variability is between 25 and 30% in the nine years considered. Despite the decrease of this variability over the years, the differentials are still systematically large, indicating that the results within each year do not change significantly.

Some explanations to this wage differentials behavior are related to the existence of local labor markets. The worker skills are different according to the demand of each regional labor market. The same idea can be maintained when considering the transportation costs, which are unequal among regions (Haddad and Hewings, 1998).

---

[8] We follow the methodology proposed by Halvorsen and Palmquist (1980), where percentage differentials are calculated as [100 * (e^{δ}) – 1].

TABLE 1

Level Regression - Regions

Dependent variable: Logaritm of real wages																		
Region	**1995**		**1996**		**1997**		**1998**		**1999**		**2000**		**2001**		**2002**		**2003**	
RO	0.06	(0.039)	0.12*	(0.038)	0.16*	(0.038)	0.10*	(0.037)	0.18*	(0.037)	0.08**	(0.037)	0.08**	(0.036)	0.07**	(0.036)	0.11*	(0.036)
AC	-0.39*	(0.042)	-0.38*	(0.041)	-0.41*	(0.041)	-0.23*	(0.040)	-0.13*	(0.040)	-0.08**	(0.040)	-0.03	(0.038)	-0.05	(0.039)	-0.08***	0.04
AM	-0.18*	(0.026)	0.05**	(0.026)	0.07*	(0.026)	0.00	(0.025)	0.00	(0.025)	-0.05**	(0.025)	0.08*	(0.024)	0.02	(0.025)	-0.12*	0.02
RR	0.31*	(0.074)	0.23*	(0.073)	0.36*	(0.073)	0.15**	(0.072)	0.15**	(0.072)	0.05	(0.072)	0.06	(0.071)	0.10	(0.071)	0.26*	0.07
PA	-0.24*	(0.020)	-0.24*	(0.020)	-0.24*	(0.020)	-0.23*	(0.019)	-0.21*	(0.019)	-0.23*	(0.019)	-0.24*	(0.019)	-0.25*	(0.019)	-0.20*	0.02
AP	0.21*	(0.059)	0.01	(0.058)	0.49*	(0.057)	0.41*	(0.055)	0.02	(0.056)	0.04	(0.056)	0.25*	(0.054)	0.33*	(0.053)	0.49*	0.05
TO	0.06	(0.046)	-0.30*	(0.045)	-0.34*	(0.044)	-0.14*	(0.043)	-0.06	(0.043)	-0.09**	(0.043)	-0.11*	(0.041)	-0.15*	(0.042)	-0.13*	0.04
MA	-0.54*	(0.023)	-0.56*	(0.023)	-0.56*	(0.023)	-0.55*	(0.022)	-0.53*	(0.022)	-0.50*	(0.022)	-0.47*	(0.022)	-0.44*	(0.022)	-0.33*	0.02
PI	-0.81*	(0.028)	-0.52*	(0.028)	-0.49*	(0.028)	-0.26*	(0.027)	-0.47*	(0.027)	-0.46*	(0.027)	-0.47*	(0.026)	-0.38*	(0.026)	-0.51*	0.03
CE	-0.63*	(0.015)	-0.62*	(0.015)	-0.61*	(0.015)	-0.53*	(0.015)	-0.52*	(0.015)	-0.36*	(0.015)	-0.40*	(0.014)	-0.39*	(0.014)	-0.42*	0.01
RN	-0.73*	(0.026)	-0.71*	(0.026)	-0.63*	(0.026)	-0.58*	(0.025)	-0.52*	(0.025)	-0.54*	(0.025)	-0.49*	(0.024)	-0.46*	(0.024)	-0.35*	0.02
PB	-0.96*	(0.022)	-0.98*	(0.022)	-0.87*	(0.021)	-0.85*	(0.021)	-0.89*	(0.020)	-0.91*	(0.021)	-0.93*	(0.020)	-0.93*	(0.020)	-0.78*	0.02
PE	-0.37*	(0.015)	-0.39*	(0.015)	-0.35*	(0.015)	-0.35*	(0.014)	-0.33*	(0.014)	-0.34*	(0.014)	-0.33*	(0.014)	-0.38*	(0.014)	-0.36*	0.01
AL	-0.43*	(0.025)	-0.48*	(0.025)	-0.47*	(0.025)	-0.36*	(0.024)	-0.38*	(0.024)	-0.33*	(0.024)	-0.34*	(0.023)	-0.35*	(0.024)	-0.41*	0.02
SE	-0.50*	(0.026)	-0.56*	(0.026)	-0.51*	(0.026)	-0.44*	(0.025)	-0.43*	(0.025)	-0.43*	(0.025)	-0.40*	(0.024)	-0.40*	(0.024)	-0.43*	0.02
BA	-0.30*	(0.013)	-0.30*	(0.012)	-0.32*	(0.012)	-0.29*	(0.012)	-0.31*	(0.012)	-0.26*	(0.012)	-0.29*	(0.012)	-0.30*	(0.012)	-0.31*	0.01
MG	-0.05*	(0.008)	-0.07*	(0.008)	-0.06*	(0.008)	-0.10*	(0.007)	-0.08*	(0.007)	-0.10*	(0.007)	-0.09*	(0.007)	-0.08*	(0.007)	-0.11*	0.01
ES	0.06*	(0.020)	0.02	(0.020)	-0.03	(0.020)	-0.02	(0.019)	-0.02	(0.019)	-0.06*	(0.019)	-0.04**	(0.019)	-0.06*	(0.019)	-0.11*	0.02
RJ	-0.01	(0.008)	0.00	(0.008)	-0.04*	(0.008)	-0.03*	(0.008)	-0.05*	(0.008)	-0.04*	(0.008)	-0.01***	(0.007)	0.02**	(0.007)	-0.01	0.01
SP	0.28*	(0.004)	0.29*	(0.004)	0.28*	(0.004)	0.28*	(0.004)	0.28*	(0.004)	0.29*	(0.004)	0.27*	(0.004)	0.25*	(0.004)	0.24*	0.00
PR	0.04*	(0.010)	0.04*	(0.010)	0.05*	(0.010)	0.02**	(0.010)	-0.01	(0.010)	-0.03*	(0.010)	-0.05*	(0.010)	-0.05*	(0.010)	0.00	0.01
SC	0.09*	(0.013)	0.10*	(0.013)	0.09*	(0.013)	0.05*	(0.013)	0.03**	(0.012)	0.00	(0.012)	-0.01	(0.012)	-0.01	(0.012)	0.05*	0.01
RS	0.06*	(0.009)	0.07*	(0.009)	0.06*	(0.009)	0.06*	(0.009)	0.07*	(0.009)	0.05*	(0.009)	0.06*	(0.009)	0.08*	(0.009)	0.09*	0.01
MS	-0.15*	(0.027)	-0.17*	(0.026)	-0.12*	(0.026)	-0.17*	(0.025)	-0.13*	(0.025)	-0.14*	(0.025)	0.08*	(0.024)	-0.09*	(0.025)	-0.10*	0.02
MT	0.01	(0.028)	-0.09*	(0.028)	-0.05***	(0.028)	-0.07**	(0.027)	-0.05**	(0.027)	-0.08*	(0.027)	-0.07**	(0.026)	-0.04***	(0.027)	-0.08*	0.03
GO	-0.17*	(0.017)	-0.21*	(0.017)	-0.21*	(0.017)	-0.19*	(0.016)	-0.16*	(0.016)	-0.15*	(0.016)	-0.14*	(0.016)	-0.08*	(0.016)	-0.07*	0.02
DF	0.61*	(0.018)	0.55*	(0.018)	0.54*	(0.018)	0.48*	(0.017)	0.54*	(0.017)	0.50*	(0.018)	0.46*	(0.017)	0.45*	(0.017)	0.53*	0.02
SD(φ)	0.3012		0.2988		0.2885		0.2683		0.2722		0.2623		0.2569		0.2510		0.2450	
Observations	64,445		64,445		64,445		64,445		64,445		64,445		64,445		64,445		64,445	
R^2	0.4982		0.4945		0.4916		0.5021		0.5003		0.4919		0.4997		0.4943		0.4887	

Notes: The coefficients are the proportionate difference in wages between an employee in a given region and the average employee, following the KS methodology. The reported standard errors in parenthesis and the SD(φ) are based on HDS procedure. Independent variables are age, age squared, tenure, tenure squared, nine education dummies, gender dummy, eight industry dummies and six occupation dummies, besides a constant.

* Significant at 1% level.

** Significant at 5% level.

*** Significant at 10% level.

5.1.2. Panel Regressions

Regarding the wage differences among workers of the 27 Brazilian regions,[9] the results in the third column of table 2 show that there are only 18 significant coefficients at the 10% level when using the fixed effects regression. Whether we consider a higher level of significance, such as 1%, the number of significant coefficients falls to 14. This result contrasts with the OLS regression, in the first column, whose coefficients are all significant, except those for Parana state. Moreover, almost all the significant coefficients from the fixed effects regression show lower wage differences in comparison to the OLS coefficients. For instance, the first two states in the ranking of wage differences in the OLS regression were DF and SP, with coefficients of 68% and 31% above the average of all states. Using the fixed effects regression, these states drop to the 5th and 7th positions, respectively, with coefficients that are only 12% and 5% above the average.

TABLE 2
Estimated inter-region wage differentials

Dependent variable: Logaritm of real wages						
Region	**POLS**		**Random Effects**		**Fixed Effects**	
RO	0.1053 *	(0.0125)	0.0173	(0.0266)	0.0690	(0.0428)
AC	-0.1936 *	(0.0133)	-0.2745 *	(0.0330)	-0.0823	(0.0780)
AM	-0.0152 ***	(0.0084)	0.0132	(0.0191)	0.1267 **	(0.0332)
RR	0.1925 *	(0.0242)	0.2077 *	(0.0474)	0.2710 *	(0.0662)
PA	-0.2309 *	(0.0065)	-0.1811 *	(0.0142)	-0.0632 *	(0.0238)
AP	0.2523 *	(0.0186)	0.0714 ***	(0.0366)	-0.0982 ***	(0.0524)
TO	-0.1525 *	(0.0144)	-0.0016	(0.0296)	0.2686 *	(0.0436)
MA	-0.4937 *	(0.0074)	-0.4741 *	(0.0171)	-0.1583 *	(0.0320)
PI	-0.4839 *	(0.0090)	-0.5156 *	(0.0210)	-0.3369 *	(0.0403)
CE	-0.4949 *	(0.0050)	-0.4372 *	(0.0111)	-0.1647 *	(0.0192)
RN	-0.5532 *	(0.0084)	-0.4562 *	(0.0190)	-0.0973 *	(0.0331)
PB	-0.8949 *	(0.0069)	-0.5905 *	(0.0156)	-0.0939 *	(0.0283)
PE	-0.3554 *	(0.0048)	-0.3181 *	(0.0106)	-0.0194	(0.0187)
AL	-0.3928 *	(0.0081)	-0.3955 *	(0.0189)	-0.0868 **	(0.0364)
SE	-0.4509 *	(0.0084)	-0.4399 *	(0.0193)	-0.1196 *	(0.0357)
BA	-0.2960 *	(0.0040)	-0.2511 *	(0.0090)	-0.0143	(0.0154)
MG	-0.0821 *	(0.0025)	-0.0689 *	(0.0055)	-0.0435 *	(0.0096)
ES	-0.0250 *	(0.0065)	-0.0205	(0.0140)	-0.0091	(0.0220)
RJ	-0.0215 *	(0.0025)	0.0506 *	(0.0055)	0.0298 *	(0.0088)
SP	0.2730 *	(0.0013)	0.2177 *	(0.0030)	0.0484 *	(0.0053)
PR	-0.0009	(0.0033)	-0.0030	(0.0071)	-0.0245 **	(0.0113)
SC	0.0420 *	(0.0042)	0.0284 *	(0.0093)	-0.0252	(0.0158)
RS	0.0657 *	(0.0031)	0.0678 *	(0.0072)	0.0044	(0.0145)
MS	-0.1151 *	(0.0085)	-0.0395 **	(0.0180)	-0.0167	(0.0276)
MT	-0.0611 *	(0.0091)	-0.0155	(0.0182)	0.1544 *	(0.0263)
GO	-0.1550 *	(0.0055)	-0.1086 *	(0.0113)	0.0203	(0.0169)
DF	0.5186 *	(0.0058)	0.2663 *	(0.0097)	0.1140 *	(0.0120)
SD(φ)	0.2686		0.2167		0.0681	
Observations	580,005		580,005		580,005	
Individuals	64,445		64,445		64,445	
R^2	0.4947		-		-	
R^2 within	-	-	0.0891		0.0953	
R^2 between	-	-	0.4956		0.1603	
R^2 overall	-	-	0.4363		0.1097	
Hausman	-			chi2(59) = 53,810.81		
Breusch-Pagan	-			chi2(1) = 1,300,000		

Notes: The coefficients are the proportionate difference in wages between an employee in a given region and the average employee, following the KS methodology. The reported standard errors in parenthesis are based on HDS procedure. Independe variables are age, age squared, tenure, tenure squared, nine education dummies, gender dummy, years dummies, eight industry dummies and six occupation dummies, besides a constant.; * Significant at 1% level; ** 5% level; *** 10% level.

[9] Although there are 26 states plus the Federal District, we will consider all regions as states in order to make easier the analysis.

Comparing the regions of higher and lower wage differences between the OLS regression and the fixed effects regression, we can note that the magnitude of these differentials is significantly reduced. First, observe the coefficients of workers from the DF and PB regions in the OLS regression. They are +68% and -59% compared to the average, respectively. Second, observe and compare the coefficients from the extreme regions of the fixed effects regression: now, RR, with +31%, and PI, with -29% compared to the average. This considerable reduction in the range of wage differentials provides evidence in favor the unmeasured abilities theory. These results also corroborate the previous studies about the large Brazilian inequality in earnings.[10]

Another important result obtained is the overall variability of the wage differentials. While the employment-weighted adjusted standard deviation of the raw region log differentials - SD(δ) - is 0.27 in the OLS regression, the fixed effect regression exhibits an overall variability of only 0.07. These two results show that there is a large part of inter-region wage differences that can be explained by the individual heterogeneity. When the individual fixed effects are controlled, the inter-regional wage differentials persist, but lose importance. As a consequence, the differences among states have a more limited explanation of wage differences than the results found in prior literature on wage determination.

It is important to note that the fixed effects coefficients estimated represent the wage differential reached by workers who migrate among regions. In this regard, the inter-region wage differences could be explained by non-observed individual abilities which vary according to the regional "affiliation", i.e., the region they are working in. However, the point here is to show that such specific differences are of a smaller degree than previously thought.

5.2. Inter-industry wage differences

5.2.1. Level regressions

Table 3 reports inter-industry wage differentials among workers and the overall variability of these differences. The pattern of wage differentials among industries presents five sectors above the average wage and the other three below the average wage. The first set of industries includes public utilities, mining, manufacturing, services, and construction, whereas the second set incorporates the public sector, farming and trade. The contrast concerning inter-industry wage differences between these two sets of industries can be exemplified by the comparison of the sector of public utilities to the public sector. While in the first sector the wage differences are, for example, 67% above the average in 2003, in the second the wage differences are 19% below the same average.

The overall variability of the wage differentials among industries is between 17 and 23% in the nine years considered[11]. In spite of the decrease of this variability between 2000 and 2002, the differentials are still systematically large. This indicates that the results within each year do not change in important ways.

[10] It can also be seen that the lowest wage differences compared to the average are concentrated in the states of Northeast region, particularly in the PI, CE, MA, SE, AP, RN, PB and AL states.

11 Arbache (1999), using individual data from the National Household Survey for Brazilian manufacturing in 1995, found an overall variability of 7%.

TABLE 3

Level Regression - Industries

Dependent variable: Logarithm of real wages									
Industry	1995	1996	1997	1998	1999	2000	2001	2002	2003
Pub. Administration	-0.25* (0.004)	-0.26* (0.004)	-0.25* (0.004)	-0.24* (0.004)	-0.23* (0.004)	-0.20* (0.004)	-0.18* (0.004)	-0.18* (0.004)	-0.21* (0.004)
Farming	-0.11* (0.022)	-0.12* (0.021)	-0.09* (0.022)	-0.03 (0.021)	-0.06* (0.022)	-0.11* (0.022)	-0.07* (0.022)	-0.08* (0.022)	-0.07* (0.022)
Trade	-0.03* (0.009)	-0.02** (0.009)	-0.05* (0.009)	-0.04* (0.009)	-0.02*** (0.009)	-0.06* (0.009)	-0.07* (0.009)	-0.06* (0.009)	-0.05* (0.009)
Construction	0.10* (0.020)	0.11* (0.020)	0.12* (0.021)	0.16* (0.020)	0.13* (0.021)	0.10* (0.020)	0.08* (0.020)	0.08* (0.021)	0.11* (0.021)
Mining	0.40* (0.044)	0.43* (0.044)	0.35* (0.043)	0.36* (0.039)	0.37* (0.041)	0.38* (0.039)	0.40* (0.038)	0.38* (0.039)	0.42* (0.038)
Manufacturing	0.30* (0.007)	0.29* (0.007)	0.28* (0.007)	0.24* (0.007)	0.24* (0.007)	0.23* (0.007)	0.23* (0.007)	0.24* (0.007)	0.28* (0.006)
Public Utilities	0.46* (0.018)	0.49* (0.017)	0.53* (0.017)	0.54* (0.017)	0.47* (0.016)	0.48* (0.017)	0.50* (0.016)	0.50* (0.016)	0.51* (0.016)
Services	0.11* (0.004)	0.13* (0.004)	0.13* (0.004)	0.14* (0.004)	0.13* (0.004)	0.11* (0.004)	0.08* (0.004)	0.08* (0.004)	0.09* (0.004)
SD(φ)	0.2180	0.2269	0.2226	0.2123	0.2039	0.1862	0.1765	0.1749	0.2002
Observations	64,445	64,445	64,445	64,445	64,445	64,445	64,445	64,445	64,445
R^2	0.4982	0.4945	0.4916	0.5021	0.5003	0.4919	0.4997	0.4943	0.4887

Notes: See notes of table 1. Independent variables are age, age squared, tenure, tenure squared, nine education dummies, gender dummy, twenty seven region dummies and six occupation dummies, besides a constant.

5.2.2. Panel Regressions

The inter-industry wage differences reported in table 4 have a pattern similar to that verified among Brazilian regions. The estimated coefficients using the POLS regression are all significant, exhibiting an impressive variation range between the extreme values. While the group of public management is characterized by wages 20% below the average of all industries, the public utilities has workers earning 65% above the average of all industries. However, when considering the fixed effects regression, only six estimated coefficients are significant. Besides that, all exhibited differences are lower than that from POLS regression and the variation range is considerably reduced. For example, the range between farming, the group that has the least differentials compared to the mean (-4.2%) in the fixed effects regression, and the public utilities, which has the larger coefficient (9.4%), is substantially reduced in comparison to the POLS regression.

The overall variability, measured by SD(φ), reduces significantly: 0.20 in OLS regression to 0.04 in the fixed effects regression. Therefore, the individual heterogeneity can account for a large part of inter-industry wage differences. The inter-industry wage differentials persist after controlling for individual-specific fixed effects, but lose importance. These outcomes are similar to that reached by Carruth *et al.* (2004) using data from the UK. They found an overall variability of 0.02 when taking into account the fixed effects, showing that differences among industries have a more limited explanation of wage differences.

TABLE 4
Estimated inter-industry wage differentials

Dependent variable: Logaritm of real wages						
Industry	**POLS**		**Random Effects**		**Fixed Effects**	
Pub. Administration	-0.2203 *	(0.0013)	-0.0610 *	(0.0018)	-0.0166 *	(0.0021)
Farming	-0.0871 *	(0.0073)	-0.0846 *	(0.0067)	-0.0426 *	(0.0070)
Trade	-0.0443 *	(0.0030)	-0.0364 *	(0.0035)	-0.0407 *	(0.0038)
Construction	0.1110 *	(0.0069)	0.0075	(0.0063)	-0.0228	(0.0064)
Mining	0.3876 *	(0.0135)	0.0877 *	(0.0146)	0.0068	(0.0153)
Manufacturing	0.2591 *	(0.0022)	0.1086 *	(0.0029)	0.0701 *	(0.0032)
Public Utilities	0.4989 *	(0.0056)	0.2471 *	(0.0083)	0.0895 *	(0.0095)
Services	0.1101 *	(0.0014)	0.0138 *	(0.0017)	-0.0098 *	(0.0018)
SD(φ)	0.2021		0.0728		0.0367	
Observations	580,005		580,005		580,005	
Individuals	64,445		64,445		64,445	
R^2	0.4947		-		-	
R^2 within	-	-	0.0891		0.0953	
R^2 between	-	-	0.4956		0.1603	
R^2 overall	-	-	0.4363		0.1097	
Hausman	-		chi2(59) = 53,810.81			
Breusch-Pagan	-		chi2(1) = 1,300,000			

Notes: See notes of table 2. Independent variables are age, age squared, tenure, tenure squared, nine education dummies, gender dummy, years dummies, six occupation dummies and 27 region dummies, besides a constant.

5.3. Inter-occupation wage differences

5.3.1. Level regressions

Table 5 reports estimated inter-occupation wage differentials among workers and the overall variability of these differences based on level regressions. The pattern of wage differentials among occupations indicates three groups above the average wage and the other three below the average wage. The first set of occupations includes more qualified occupations, whereas the second set incorporates less qualified occupations. The contrast concerning inter-occupation wage differences can be highlighted by the comparison of the occupation 2, which involves workers from legislative, executive, judiciary, public sector and directors, and the occupation 5, which includes workers from farming, forestry activities and fishing. These categories present the extreme wage differentials. While the former shows, for example, differences around 51% above the average wage in 2003, the latter has 32% below the same average.

The overall variability of the wage differentials among occupations increases along the years, changing from 14% in 1995 to 18% in 2003. This indicates that the results within each year do not change significantly.

5.3.2. Panel Regressions

The inter-occupation wage differences reported in table 6 have a pattern similar to that verified among Brazilian regions. The estimated coefficients using the POLS regression are all significant, exhibiting a large variation between the extreme values. While the fifth occupational group is characterized by wages 33% below the average of all occupations, the second occupational group has workers earning 51% (on average) above the average of all occupations. On the other hand, when considering the fixed effects regression, none of the estimated coefficients are significant. Moreover, all exhibited differences are lower than that from POLS regression and variation range is considerably reduced. For example, the range between occupation 5, the group that has the least differentials compared to the mean (-4.4%) in the fixed effects regression, and occupation 2, which has the larger coefficient (4.7%), is substantially reduced in comparison to the POLS regression.

The overall variability measured by SD(γ) is reduced significantly: 0.16 in OLS regression to 0.02 in the fixed effects regression. The most important evidence of the wage differentials behavior is that the wage differences persist, but lose importance. Therefore, a large part of inter-occupation wage differences can be explained by the individual heterogeneity. In other words, differences among occupations have a more limited explanation of wage differences than the previous results that disregarded the individual fixed effects.

Although the regional differences are higher than the industry differences and also the occupation differences, the general idea is exactly the same; the unobserved individual abilities have an important role in the explanation of wage differences. As a result, the region, industry and occupation “affiliation” have a lower power of explanation in accounting for the wage differentials among workers. Then, a significant proportion of wage differentials which was previously ascribed to region, industry and occupation differences is actually a result of non-observable individual attributes that could not be removed using an estimation of pooled ordinary least squares.

TABLE 5

Level Regression - Occupations

Dependent variable: Logarithm of real wages

Occupation	1995	1996	1997	1998	1999	2000	2001	2002	2003
1	0.02* (0.007)	0.05* (0.006)	0.06* (0.006)	0.09* (0.006)	0.09* (0.006)	0.06* (0.006)	0.08* (0.006)	0.08* (0.006)	0.12* (0.005)
2	0.38* (0.013)	0.36* (0.012)	0.31* (0.012)	0.42* (0.011)	0.38* (0.011)	0.43* (0.011)	0.56* (0.012)	0.55* (0.011)	0.41* (0.010)
3	0.11* (0.005)	0.10* (0.005)	0.10* (0.005)	0.08* (0.004)	0.08* (0.004)	0.07* (0.004)	0.07* (0.004)	0.06* (0.004)	0.01** (0.005)
4	-0.17* (0.006)	-0.18* (0.006)	-0.18* (0.006)	-0.20* (0.006)	-0.19* (0.006)	-0.18* (0.006)	-0.20* (0.006)	-0.20* (0.006)	-0.23* (0.006)
5	-0.35* (0.023)	-0.37* (0.023)	-0.40* (0.023)	-0.46* (0.023)	-0.42* (0.023)	-0.40* (0.023)	-0.42* (0.023)	-0.41* (0.023)	-0.39* (0.022)
6	-0.04* (0.006)	-0.04* (0.006)	-0.03* (0.006)	-0.05* (0.006)	-0.05* (0.006)	-0.05* (0.006)	-0.07* (0.006)	-0.07* (0.006)	-0.08* (0.006)
SD(φ)	0.1395	0.1446	0.1409	0.1654	0.1553	0.1576	0.1772	0.1817	0.1700
Observations	64,445	64,445	64,445	64,445	64,445	64,445	64,445	64,445	64,445
R^2	0.4982	0.4945	0.4916	0.5021	0.5003	0.4919	0.4997	0.4943	0.4887

Notes: See notes of table 1. Independent variables are age, age squared, tenure, tenure squared, nine education dummies, gender dummy, twenty seven region dummies and eight industry dummies, besides a constant.

The occupations are described in section 4.

TABLE 6
Estimated inter-occupation wage differentials

Dependent variable: Logaritm of real wages								
Occupation	**POLS**			**Random Effects**			**Fixed Effects**	
Occupation 1	0.0763	*	(0.0020)	0.0523		(0.0018)	0.0167	(0.0019)
Occupation 2	0.4134	*	(0.0038)	0.1042	*	(0.0029)	0.0427	(0.0030)
Occupation 3	0.0785	*	(0.0015)	0.0338		(0.0014)	0.0042	(0.0015)
Occupation 4	-0.1942	*	(0.0019)	-0.0789	*	(0.0019)	-0.0303	(0.0019)
Occupation 5	-0.4008	*	(0.0076)	-0.1752	*	(0.0077)	-0.0479	(0.0083)
Occupation 6	-0.0543	*	(0.0021)	-0.0279		(0.0020)	0.0002	(0.0021)
SD(φ)	0.1575			0.0610			0.0201	
Observations	580,005			580,005			580,005	
Individuals	64,445			64,445			64,445	
R^2	0.4947			-			-	
R^2 within	-			0.0891			0.0953	
R^2 between	-			0.4956			0.1603	
R^2 overall	-			0.4363			0.1097	
Hausman	-			chi2(59) = 53,810.81				
Breusch-Pagan	-			chi2(1) = 1,300,000				

Notes: See notes of table 2. Independent variables are age, age squared, tenure, tenure squared, nine education dummies, gender dummy, years dummies, eight industry dummies and 27 region dummies, besides a constant.

6. Conclusion

The aim of this paper was to verify the inter-region, inter-industry and inter-occupation wage differentials of Brazilian workers over nine years (1995-2003), controlled by observable and non-observable characteristics of individuals. The inclusion of non-observable individual characteristics, such as ability, motivation, etc., in the model, was implemented by using the fixed effects method. We also considered the methodology of KS HDS in order to provide estimates of wage dispersion. By using the HDS approach, we were able to provide more accuracy in the estimation results, such as the exact estimation of standard errors, the correct measure of wage dispersion and its overall variability.

The most important results show that the wages among the three patterns of comparison adopted in the course of this paper – region, industry and occupation – have lower differentials than the previous OLS results. A large amount of the wage variance assumed as inter-region, inter-industry and inter-occupation differentials is a consequence of non-observable differences among individuals that could not be removed by a simple OLS estimation. Although we still confirm the empirical regularity of persistent earning differences, the size and significance of wage differentials decrease considerably. Our results also confirm, in fact, that the procedure of KS provides overestimation of standard errors and a biased estimation of the measure of wage dispersion.

7. References

Abowd, J. M., Kramarz, F. (1999). 'The analysis of labor markets using matched employer-employee data'. *In*: Ashenfelter, O. e Card, D, eds., Handbook of labor economics (Elsevier Science B. V.), pp. 2629-2711.

Abowd, J. M., Kramarz, F. and Margolis, D. N. (1999). 'High wage workers and high wage firms', Econometrica, Vol. 67, pp. 251-333.

Akerlof, G. (1984). 'Gift exchange and efficiency wages: four views', American Economic Review, Vol. 74, pp. 79-83.

Arai, M. (1994) 'An empirical analysis of wage dispersion and efficiency wages', Scandinavian Journal of Economics, Vol. 96, No. 1, pp. 31-50.

Arbache, J. S. (2001) 'Wage differentials in Brazil: theory and evidence'. Journal of Development Studies, Vol. 38, pp. 691-714.

Arbache, J. S., Carneiro, F. G. (1999). 'Unions and interindustry wage differentials', World Development, Vol. 27, No. 10, pp. 1875-83.

Arbache, J. S. (1999). 'A comparison of different estimates of inter-industry wage differentials: the case of Brazil', Applied Economics Letters, Vol. 6, pp. 67-71.

Bacha, E., Taylor, L. (1978). 'Brazilian income distribution in the 60's: facts, model results and controversy'. The Journal of Development Studies, Vol. 14, No. 3, pp. 271-97.

Barros, R. P., Mendonça, R. (1995) "A Evolução do Bem-Estar e da desigualdade no Brasil", Revista Brasileira de Economia, vol.49, pp.329-52.

Bulow, J. I., Summers, L. H. (1986) 'A theory of dual labor market with application to industrial policy, discrimination, and Keynesian unemployment'. Journal of Labor Economics, Vol. 4, No. 3, pp. 376-414.

Carruth, A., Collier, W., Dickerson, A. (2004) 'Inter-industry wage differences and individual heterogeneity.' Oxford Bulletin of Economics and Statistics, Oxford, Vol. 66, No. 5.

Castells, M., Portes, A. (1989) 'World underneath: the origins, dynamics and effects of the informal economy', in A. Portes, M. Castells and L. A. Benton (eds.), The informal economy: studies in advanced and less developed countries. Baltimore, MD: Johns Hopkins University Press.

Corbo, V., Stelcner, M. (1983) 'Earnings determination and labor markets: gran Santiago – Chile – 1978', Journal of Development Economics, Vol. 12, No. 5.1-2, pp. 251-66.

Cowell, F. A., Ferreira, F. H. G., Litchfield, J. (1996). 'Income distribution in Brazil 1981-1990: parametric and non-parametric approaches'. Discussion paper No. DARP 21, London School of Economics.

Dickens, W. T. and Katz, L. F. (1987). Inter-industry Wage Differences and Theories of Wage Determination, NBER Working Paper n. 2271.

Dickens, W. T. (1986). Wages, Employment and the Threat of Collective Action by Workers, NBER Working Paper no. 1856.

Fields, G, Wan, H. (1989) 'Wage-setting institutions and economic growth', World Development, Vol. 17, No. 9, pp. 1471-83.

Gatica, J., Mizala, A., Romaguera, P. (1995) Ínterindustry wage differentials in Brazil'. Economic Development and Cultural Change, Vol. 43, No. 2, pp. 315-31.

Gera, S. and Grenier, G. (1994) 'Interindustry wage differentials and efficiency wages: some Canadian evidence.' Canadian Journal of Economics and Statistics, 27, 81–100.

Gibbons, R. and Katz, L. (1992). 'Does unmeasured ability explain inter-industry wage differentials?', Review of Economic Studies, Vol. 59, pp. 515-535.

Gittleman, M. and Wolff, E. N. (1993). 'International comparisons of inter-industry wage differentials', Review of Income and Wealth, Vol. 39, pp. 295-312.

Groshen, E. L. (1991) 'Five reasons why wages vary among employers'. Industrial Relations, Vol. 30, No. 3, pp. 351-81.

Haddad, E. A. and Hewings, G. J. D. (1998). 'Transportation costs and regional development: an interregional CGE analysis', 38th European Congress of the Regional Science Association Vienna, Austria.

Haisken-DeNew, J. P. and Schmidt, C. M. (1997). 'Inter-industry and inter-region differentials: mechanics and interpretation', Review of Economics and Statistics, Vol. 79, pp. 516-521.

Halvorsen, R. and Palmquist, R. (1980). 'The interpretation of dummy variables in semi-logarithmic equations', American Economic Review, Vol. 40, pp. 474-475.

Harris, J. R., Todaro, M. P. (1970) 'Migration unemployment and development: a two sector analysis', American Economic Review, Vol. 60, March, pp. 126-42.

Jackubson, G. (1991) 'Estimation and testing of the union wage effect using panel data', The Review of Economic Studies, Vol. 58, No. 5, October, pp. 971-991.

Kahn, L. M. (1998). 'Collective bargaining and the inter-industry wage structure: international evidence', Economica, Vol. 65, pp. 507-534.

Keane, M. P. (1993). 'Individual heterogeneity and interindustry wage differentials', Journal of Human Resources, Vol. 28, pp. 134-161.

Krueger, A. B. and Summers, L. H. (1987). 'Reflection on the inter-

industry wage structure', in Lang K. and Leonard J. (eds), Unemployment and the Structure of Labor Markets, Basil Blackwell, Oxford.

Krueger, A. B. and Summers, L. H. (1988). 'Efficiency wages and the inter-industry wage structure', Econometrica, 56, pp. 259-293.

Lam, D., Levinson, D. (1992) 'Age, experience, and schooling: decomposing earnings inequality in the United States and Brazil', Social Inquiry, Vol. 62. No. 2, pp. 221-45.

Langoni, C. G. (1973). 'Distribuição de renda e crescimento econômico no Brasil', Rio de Janeiro, Editora Expressão e Cultura.

Lindauer, D., Sabot, R. (1983). 'The public/private wage differential in a poor urban economy', Journal of Development Economics, Vol. 12. No. 5. 1-2. pp. 137-52.

Lindbeck, A. and Snower, D. J. (1986). 'Wage setting, unemployment and insider-outsider relations', American Economic Review, Vol. 76, pp. 235-239.

Lucifora, C. (1993) 'Inter-industry and occupational wage differentials in Italy', Applied Economics, 25, 1113–24.

Macedo, R. (1985). "Diferenciais de salários entre empresas estatais e privadas no Brasil', Revista Brasileira de Economia. Vol. 39, No. 4, pp. 437-48.

Mincer, J. (1974) 'Schooling, experience and earnings', New York: National Bureau for Economic Research.

Morrison, A.R. (1994). 'Are institutions or economic rents responsible for interindustry wage differentials?' World Development, Vol. 22, No. 3, pp. 355-68.

Murphy, K. M. and Topel, R. H. (1987). 'Unemployment, risk, and earnings: testing for equalizing wage differences in the labor market', in Lang K. and Leonard J. (eds), Unemployment and the Structure of Labor Markets, Basil Blackwell, Oxford.

Pinheiro, A. C., Ramos, L. (1994). 'Inter-industry wage differentials and earning inequality. Estudios de Economia. Vol. 21, No. 1, pp. 79-111.

Psacharopoulos, G., Velez, E. (1992). 'Schooling, ability, and earnings in Colombia, 1988. Economic Letters, Vol. 4, pp. 181-5.

Relatório Anual de Informações Sociais – Migração (RAIS-Migra), of the Labor Ministry of Brazil.

Shapiro, C. and Stiglitz, J. E. (1984). ‘Equilibrium unemployment as a worker discipline device’, American Economic Review, Vol. 74, pp. 433-444.

Shippen, B. S. (1999). ‘Unmeasured skills in inter-industry wage differentials: evidence from the apparel industry’, Journal of Labor Research, Vol. 20, pp. 161-169.

Solow, R. M. (1985). ‘Insiders and outsiders in wage determination’, Scandinavian Journal of Economics, Vol. 87, pp. 411-428.

Stiglitz, J. E. (1985). ‘Equilibrium wage distribution’, Economic Journal, Vol. 95, pp. 595-618.

Suits, D.B. (1984) Dummy variables: mechanics and interpretations, Review of Economics and Statistics, 66, 177–80.

Teal, F. (1996). ‘The size and sources of economic rents in developing country manufacturing labor market’. Economic Journal. Vol. 106, No. 473, pp. 963-76.

Tokman, V. E. (1983). ‘The influence of the urban informal sector on economic inequality’. In F. Stewart (ed.). Work, income and inequality. Payments System in the third world. London: Macmillan Press.

Weiss, A. (1980). ‘Job queues and layoffs in labour markets with flexible wages’, Journal of Political Economy, Vol. 88, pp. 526-538.

Wooldridge, J. M. (2002) ‘Econometric analysis of cross-section and panel data’, Massachusetts: MIT.

Yamada, G. (1996). ‘Urban informal employment and self-employment in developing countries: theory and evidence’. Economic Development and Cultural Change, Vol. 44, No. 2, pp. 289-314.

8. Apendix

TABLE A1

Definitions and summary statistics by region

Variable/Region	RO	AC	AM	RR	PA	AP	TO	MA	PI	CE	RN	PB	PE	AL
Education Level														
Illiterate	1.33	1.82	1.44	1.90	2.80	0.32	1.59	9.20	1.33	1.50	1.95	6.96	6.14	8.93
	[37]	[44]	[86]	[14]	[284]	[4]	[33]	[726]	[71]	[258]	[117]	[638]	[1,145]	[588]
Incomplete 1st.	9.12	17.90	7.40	10.99	5.57	4.88	5.46	5.95	12.38	8.33	7.75	8.72	13.73	14.74
elementary level	[253]	[432]	[443]	[81]	[565]	[61]	[113]	[469]	[660]	[1,430]	[466]	[799]	[2,561]	[970]
1st. Elementary level	15.90	2.44	6.48	4.34	8.28	2.56	4.15	3.47	4.67	6.16	6.69	6.23	7.38	7.51
	[441]	[59]	[388]	[32]	[840]	[32]	[86]	[274]	[249]	[1,058]	[402]	[571]	[1,377]	[494]
Incomplete 2nd.	14.06	11.27	9.15	9.91	10.66	7.76	7.58	3.31	4.48	7.90	9.50	7.45	10.21	8.52
elementary level	[390]	[272]	[548]	[73]	[1,081]	[97]	[157]	[261]	[239]	[1,357]	[571]	[683]	[1,904]	[561]
2nd. Elementary level	13.48	15.78	10.77	19.81	12.51	16.48	9.76	12.06	22.10	14.68	11.88	7.03	10.31	10.82
	[374]	[381]	[645]	[146]	[1,269]	[206]	[202]	[951]	[1,178]	[2,520]	[714]	[644]	[1,923]	[712]
Incomplete	6.52	2.57	5.58	4.75	6.61	3.84	5.56	3.64	7.26	4.20	7.29	3.15	5.49	4.38
medium school	[181]	[62]	[334]	[35]	[670]	[48]	[115]	[287]	[387]	[722]	[438]	[289]	[1,023]	[288]
Medium school	30.57	31.11	42.15	25.78	32.14	46.24	37.68	49.68	34.80	37.00	35.47	15.11	27.46	27.53
	[848]	[751]	[2,525]	[190]	[3,260]	[578]	[780]	[3,919]	[1,855]	[6,354]	[2,132]	[1,385]	[5,120]	[1,812]
Incomplete higher	1.84	8.33	3.94	1.09	2.36	1.28	1.64	1.99	4.65	2.59	2.33	2.22	3.85	4.10
degree	[51]	[201]	[236]	[8]	[239]	[16]	[34]	[157]	[248]	[444]	[140]	[203]	[718]	[270]
Higher degree	7.17	8.78	13.11	21.44	19.08	16.64	26.57	10.70	8.31	17.63	17.15	43.13	15.42	13.48
	[199]	[212]	[785]	[158]	[1,935]	[208]	[550]	[844]	[443]	[3,028]	[1,031]	[3,952]	[2,875]	[887]
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	[2,774]	[2,414]	[5,990]	[737]	[10,143]	[1,250]	[2,070]	[7,888]	[5,330]	[17,171]	[6,011]	[9,164]	[18,646]	[6,582]
Gender														
Female	48.09	68.64	51.99	51.42	49.40	49.84	60.82	57.61	61.76	52.76	50.91	51.54	44.88	45.76
	[1,334]	[1,657]	[3,114]	[379]	[5,011]	[623]	[1,259]	[4,544]	[3,292]	[9,060]	[3,060]	[4,723]	[8,369]	[3,012]
Male	51.91	31.36	48.01	48.58	50.60	50.16	39.18	42.39	38.24	47.24	49.09	48.46	55.12	54.24
	[1,440]	[757]	[2,876]	[358]	[5,132]	[627]	[811]	[3,344]	[2,038]	[8,111]	[2,951]	[4,441]	[10,277]	[3,570]
Tenure(1)	144.96	138.72	121.60	159.15	131.03	133.67	107.36	141.78	173.55	148.09	153.51	149.73	130.26	143.96
	(81.44)	(70.70)	(83.73)	(70.96)	(83.82)	(82.74)	(76.17)	(83.44)	(99.54)	(92.54)	(97.40)	(83.48)	(89.04)	(87.52)
Age	40.08	39.25	39.13	41.29	39.96	38.87	37.68	40.30	41.60	40.29	39.16	40.77	39.30	40.09
	(9.25)	(9.58)	(9.36)	(8.60)	(8.65)	(8.37)	(8.57)	(8.54)	(8.95)	(9.56)	(8.36)	(8.95)	(9.23)	(9.27)
Real monthly wage(2)	1495.186	927.8484	1371.488	1445.142	1295.822	2021.695	1211.11	862.6585	876.4914	1037.49	949.5711	853.8357	1107.416	966.7505
	(2044.22)	(1239.49)	(1846.29)	(1040.10)	(1974.28)	(2096.69)	(1576.11)	(1309.87)	(1218.84)	(1593.97)	(1423.39)	(1357.52)	(1821.21)	(1367.34)

TABLE A1 (continued)

Definitions and summary statistics by region

Variable/Region	SE	BA	MG	ES	RJ	SP	PR	SC	RS	MS	MT	GO	DF	Total
Education Level														
Illiterate	3.34	2.48	1.34	1.44	1.11	1.09	1.07	0.58	0.86	1.81	2.28	1.93	1.11	**1.72**
	[203]	[631]	[858]	[143]	[678]	[1,919]	[387]	[136]	[370]	[107]	[117]	[272]	[139]	[10,005]
Incomplete 1st.	10.70	9.97	7.12	6.59	5.48	6.14	7.90	4.28	5.59	8.48	13.15	8.61	4.74	**7.10**
elementary level	[650]	[2,536]	[4,556]	[655]	[3,353]	[10,834]	[2,870]	[1,012]	[2,413]	[502]	[675]	[1,211]	[592]	[41,162]
1st. Elementary level	6.16	7.97	14.36	11.79	12.67	12.56	12.86	15.26	8.68	10.32	9.84	12.87	6.00	**11.15**
	[374]	[2,029]	[9,191]	[1,172]	[7,749]	[22,171]	[4,673]	[3,607]	[3,748]	[611]	[505]	[1,811]	[749]	[64,693]
Incomplete 2nd.	12.54	8.29	12.73	11.06	9.20	12.14	11.47	12.78	17.78	11.38	15.20	13.38	8.61	**11.49**
elementary level	[762]	[2,110]	[8,148]	[1,099]	[5,629]	[21,431]	[4,168]	[3,021]	[7,675]	[674]	[780]	[1,882]	[1,075]	[66,648]
2nd. Elementary level	9.22	10.75	12.16	13.92	15.09	15.16	13.16	17.11	12.70	12.21	13.09	10.24	13.11	**13.64**
	[560]	[2,736]	[7,783]	[1,384]	[9,231]	[26,759]	[4,782]	[4,046]	[5,483]	[723]	[672]	[1,440]	[1,637]	[79,101]
Incomplete	5.93	4.70	5.23	6.51	5.49	6.46	6.92	6.72	6.51	6.89	11.42	13.52	4.77	**6.14**
medium school	[360]	[1,195]	[3,344]	[647]	[3,359]	[11,404]	[2,513]	[1,589]	[2,812]	[408]	[586]	[1,902]	[595]	[35,593]
Medium school	31.54	39.19	25.86	30.29	27.28	22.38	23.77	23.97	22.97	20.95	17.85	20.60	32.37	**26.28**
	[1,916]	[9,974]	[16,547]	[3,011]	[16,687]	[39,487]	[8,635]	[5,666]	[9,916]	[1,241]	[916]	[2,898]	[4,042]	[152,445]
Incomplete higher	3.75	2.37	2.80	3.41	4.11	5.21	4.86	3.87	6.32	4.96	2.44	2.60	4.54	**4.21**
degree	[228]	[604]	[1,795]	[339]	[2,512]	[9,189]	[1,766]	[916]	[2,728]	[294]	[125]	[366]	[567]	[24,394]
Higher degree	16.82	14.28	18.39	14.98	19.56	18.86	17.99	15.43	18.59	23.01	14.73	16.24	24.75	**18.27**
	[1,022]	[3,633]	[11,771]	[1,489]	[11,965]	[33,273]	[6,536]	[3,648]	[8,026]	[1,363]	[756]	[2,285]	[3,090]	[105,964]
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	[6,075]	[25,448]	[63,993]	[9,939]	[61,163]	[176,467]	[36,330]	[23,641]	[43,171]	[5,923]	[5,132]	[14,067]	[12,486]	[580,005]
Gender														
Female	54.67	52.47	45.26	45.43	43.60	39.79	43.68	39.59	46.05	42.16	49.32	51.43	41.86	**44.67**
	[3,321]	[13,353]	[28,962]	[4,515]	[26,669]	[70,225]	[15,869]	[9,360]	[19,881]	[2,497]	[2,531]	[7,234]	[5,227]	[259,081]
Male	45.33	47.53	54.74	54.57	56.40	60.21	56.32	60.41	53.95	57.84	50.68	48.57	58.14	**55.33**
	[2,754]	[12,095]	[35,031]	[5,424]	[34,494]	[106,242]	[20,461]	[14,281]	[23,290]	[3,426]	[2,601]	[6,833]	[7,259]	[320,924]
Tenure(1)	141.66	136.92	115.92	113.87	123.43	100.84	106.02	103.59	112.08	116.26	113.66	126.40	129.04	**116.55**
	(82.55)	(90.63)	(80.15)	(83.40)	(89.97)	(76.82)	(76.63)	(78.67)	(81.58)	(78.14)	(81.55)	(84.02)	(90.02)	(83.80)
Age	38.94	39.92	38.82	38.63	39.67	37.56	37.64	36.15	38.02	38.49	38.62	39.39	38.99	**38.53**
	(8.49)	(8.89)	(9.27)	(8.88)	(9.32)	(9.48)	(9.26)	(9.17)	(9.16)	(9.14)	(9.43)	(9.13)	(8.87)	(9.32)
Real monthly wage(2)	950.7726	1111.126	1289.744	1406.86	1572.917	1765.278	1367.511	1348.745	1498.409	1310.614	1052.761	1145.271	2947.275	**1,637.90**
	(1401.51)	(1568.00)	(1649.56)	(1741.61)	(2011.44)	(1926.81)	(1812.80)	(1594.31)	(1839.88)	(1686.76)	(1287.95)	(1643.90)	(3414.61)	(1876.94)

Source: Data from Labor Ministry of Brazil - RAISMIGRA-1995-2003.

Notes: Percentual exhibited to cathegorical variables (observations number in brackets); mean values exhibited to continuous variables (standard deviations in parenthesis).

(1) In months; (2) Real Income in Reais deflated by IPCA.

TABLE A2
Definitions and summary statistics by industry

Variable/Industry	Public Administration	Farming	Trade	Construction	Mining	Manufacturing	Public Utilities	Services	Total
Education Level									
Illiterate	1.92	7.57	0.83	2.92	0.71	1.43	1.02	1.35	**1.72**
	[4,313]	[1,191]	[394]	[268]	[17]	[1,410]	[142]	[2,270]	[10,005]
Incomplete 1st.	6.64	30.16	3.80	16.53	10.93	7.36	9.72	5.53	**7.10**
elementary level	[14,912]	[4,747]	[1,815]	[1,519]	[261]	[7,271]	[1,354]	[9,283]	[41,162]
1st. Elementary level	8.12	27.13	9.40	21.78	15.12	15.41	9.66	11.19	**11.15**
	[18,239]	[4,269]	[4,489]	[2,002]	[361]	[15,212]	[1,346]	[18,775]	[64,693]
Incomplete 2nd.	7.37	11.99	16.18	14.20	10.89	18.71	11.09	11.28	**11.49**
elementary level	[16,542]	[1,887]	[7,725]	[1,305]	[260]	[18,469]	[1,545]	[18,915]	[66,648]
2nd. Elementary level	11.00	7.79	21.44	13.81	13.74	17.66	11.34	13.31	**13.64**
	[24,689]	[1,226]	[10,241]	[1,269]	[328]	[17,440]	[1,580]	[22,328]	[79,101]
Incomplete	3.93	2.68	12.46	4.39	6.83	7.76	6.72	6.70	**6.14**
medium school	[8,820]	[422]	[5,951]	[403]	[163]	[7,661]	[936]	[11,237]	[35,593]
Medium school	31.38	7.31	26.47	14.58	29.19	19.63	28.32	25.54	**26.28**
	[70,450]	[1,151]	[12,643]	[1,340]	[697]	[19,382]	[3,944]	[42,838]	[152,445]
Incomplete higher	3.81	1.00	3.33	2.61	2.76	4.01	4.39	5.49	**4.21**
degree	[8,563]	[157]	[1,589]	[240]	[66]	[3,964]	[611]	[9,204]	[24,394]
Higher degree	25.83	4.37	6.09	9.18	9.84	8.03	17.74	19.62	**18.27**
	[57,986]	[688]	[2,910]	[844]	[235]	[7,924]	[2,471]	[32,906]	[105,964]
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	**100.00**
	[224,514]	[15,738]	[47,757]	[9,190]	[2,388]	[98,733]	[13,929]	[167,756]	[570,191]
Gender									
Female	63.98	11.4	33.28	12.09	6.37	20.93	15.15	43.93	**44.67**
	[143,649]	[1,794]	[15,895]	[1,111]	[152]	[20,667]	[2,110]	[73,703]	[259,081]
Male	36.02	88.6	66.72	87.91	93.63	79.07	84.85	56.07	**55.33**
	[80,865]	[13,944]	[31,862]	[8,079]	[2,236]	[78,066]	[11,819]	[94,053]	[320,924]
Tenure[(1)]	148.7536	96.3928	71.02402	78.28474	105.4386	90.14226	158.5426	102.606	**116.55**
	(83.07)	(79.15)	(60.79)	(76.97)	(77.30)	(71.36)	(87.51)	(80.72)	(83.80)
Age	41.32584	38.04658	34.24725	38.54831	37.5536	35.27873	40.22794	37.83176	**38.53**
	(8.73)	(10.34)	(9.30)	(9.46)	(8.39)	(8.88)	(7.99)	(9.14)	(9.32)
Real monthly wage[(2)]	1511.68	836.9946	1149.461	1471.176	2083.704	1791.918	2838.322	1833.484	**1,637.90**
	(1922.53)	(1107.95)	(1352.00)	(1581.23)	(2073.00)	(1766.49)	(2315.17)	(1963.41)	(1876.94)

Source: Data from Labor Ministry of Brazil - RAISMIGRA-1995-2003.

Notes: Percentual exhibited to cathegorical variables (observations number in brackets); mean values exhibited to continuous variables (standard deviations in parenthesis).

(1) In months; (2) Real Income in Reais deflated by IPCA.

TABLE A3
Definitions and summary statistics by occupation

Variable/Occupation	Occupation 1	Occupation 2	Occupation 3	Occupation 4	Occupation 5	Occupation 6	Total
Education Level							
Illiterate	0.40	1.21	0.87	2.81	10.99	2.16	**1.72**
	[509]	[393]	[1,368]	[3,233]	[1,639]	[2,863]	[10,005]
Incomplete 1st.	0.95	1.57	3.15	12.63	35.87	11.01	**7.10**
elementary level	[1,216]	[511]	[4,964]	[14,518]	[5,352]	[14,601]	[41,162]
1st. Elementary level	1.54	3.06	5.35	18.28	27.98	21.20	**11.15**
	[1,965]	[993]	[8,421]	[21,022]	[4,175]	[28,117]	[64,693]
Incomplete 2nd.	2.04	3.45	7.19	17.55	12.05	22.33	**11.49**
elementary level	[2,608]	[1,122]	[11,324]	[20,186]	[1,798]	[29,610]	[66,648]
2nd. Elementary level	5.08	5.03	13.10	19.88	6.48	20.01	**13.64**
	[6,476]	[1,635]	[20,628]	[22,862]	[967]	[26,533]	[79,101]
Incomplete	3.37	3.35	7.92	7.11	1.88	6.99	**6.14**
medium school	[4,302]	[1,088]	[12,466]	[8,181]	[280]	[9,276]	[35,593]
Medium school	34.48	21.44	39.75	17.61	3.16	13.72	**26.28**
	[43,987]	[6,967]	[62,579]	[20,247]	[471]	[18,194]	[152,445]
Incomplete higher	6.27	5.86	7.04	1.71	0.23	1.06	**4.21**
degree	[7,996]	[1,904]	[11,083]	[1,972]	[34]	[1,405]	[24,394]
Higher degree	45.86	55.03	15.62	2.41	1.36	1.52	**18.27**
	[58,499]	[17,883]	[24,591]	[2,772]	[203]	[2,016]	[105,964]
Total	100	100	100	100	100	100	**100**
	[127,558]	[32,496]	[157,424]	[114,993]	[14,919]	[132,615]	[570,191]
Gender							
Female	65.90	47.25	53.36	47.30	7.37	15.22	**44.67**
	[84,059]	[15,354]	[83,994]	[54,388]	[1,100]	[20,186]	[259,081]
Male	34.10	52.75	46.64	52.70	92.63	84.78	**55.33**
	[43,499]	[17,142]	[73,430]	[60,605]	[13,819]	[112,429]	[320,924]
Tenure(1)	139.12	148.22	126.44	95.57	95.58	95.89	**116.55**
	(86.16)	(92.85)	(86.23)	(72.85)	(81.02)	(74.74)	(83.80)
Age	39.52	41.48	37.75	39.00	38.25	37.40	**38.53**
	(8.57)	(8.29)	(9.17)	(9.88)	(10.62)	(9.49)	(9.32)
Real monthly wage(2)	2158.55	3710.41	1771.38	882.32	591.73	1243.69	**1,637.90**
	(2070.50)	(3688.72)	(1897.03)	(1001.69)	(480.43)	(1031.17)	(1876.94)

Source: Data from Labor Ministry of Brazil - RAISMIGRA-1995-2003.

Notes: Percentual exhibited to cathegorical variables (observations number in brackets); mean values exhibited to continuous variables (standard deviations in parenthesis).

(1) In months; (2) Real Income in Reais deflated by IPCA.