LACEA 2001
MONTEVIDEO URUGUAY

Gustavo Gonzaga, joint with Sergio Firpo

From Public to Private: The Performance of Former Public Workers in the Brazilian Labor Market

From Public to Private:
The Performance of Former Public Workers in the Brazilian Labor Market*

Sergio Firpo

Department of Economics, UC Berkeley

firpo@econ.berkeley.edu

Gustavo Gonzaga

Department of Economics, PUC-Rio

gonzaga@econ.puc-rio.br

Preliminary and incomplete

April 2001

JEL Code: J45

Keywords: Public Sector Downsizing, Public-Private Wage Gap, Earnings Inequality

Abstract

This paper analyzes the placement in the private sector of a subset of Brazilian public-sector employees. This group left public employment in the mid-1990's through a voluntary dismissal program. This paper contrasts their wages before and after quitting the public sector, and compares both sets of wages to public and private sector wages for similar workers. We find that participants in this voluntary dismissal program suffered a significant reduction in average wage and an increase in wage dispersion. We test whether the reduction in average wage and the increase in wage dispersion is the expected outcome once one controls for observed characteristics, by means of counterfactual simulations. We also examine the importance of stigma effects and difficulties of private sector absorption due to specific human capital accumulation. A summary of new estimation procedures and results to be incorporated in the next version of the paper are included at the end of the paper, after the appendix.

* Firpo and Gonzaga are grateful to CAPES-Brazil and CNPq-Brazil, respectively, for financial support The authors thank Maurício Reis for help with the data set, and Ricardo Paes de Barros for intellectual inspiration. As usual, all remaining errors are ours.

I. Introduction

Mainly forced by a chronic public deficit problem, the Brazilian federal government exerted a great effort to reduce the number of public servants in the second half of the 1990's. One of the main government concerns was with the social welfare impact of such major layoffs. In fact, the design of an efficient public sector downsizing program and the evaluation of its social welfare effects depend heavily on an estimation of the capacity of the private sector to absorb former public workers.[1] Given the existing legal constraints on firing public workers and an overall opposition to public sector reform in Brazil, the federal government typically opted for voluntary job retrenchement schemes. Many state firms, in a step towards privatization, promoted the so-called Voluntary Dismissal Programs (*Programas de Demissão Voluntária*, VDP), which gave monetary incentives to workers who quit their jobs.[2]

In a developing country like Brazil, characterized by overstaffing, it is expected that voluntary separation programs would result in an adverse selection outcome. In general, it is expected that the best workers, those with higher reservation wages, opt to leave. On the other hand, the fact that these workers usually have very specific human capital, combined with a possible stigma effect, can reduce the probability of absorption by the private sector. Since all these variables (innate ability, specific human capital and the stigma effect) are unobservable to the econometrician, a method that considers not only how observable characteristics affect the expected wage changes of workers that join a VDP, but also leaves a role for worker unobservables, is desired.[3]

In 1995, *Rede Ferroviária Federal S.A.* (RFFSA), a huge public firm that had the monopoly on all Brazilian railways, adopted a VDP. From 1995 to the end of 1997 more than 7,000 workers joined the program, while many others opted for early retirement. A first survey with those former employees at RFFSA was conducted in March 1998.[4] It revealed that their reinsertion in the private labor market has not been a complete success: for those who were working during the month of the interview, 72% were earning less than at RFFSA and 80% changed occupation.

---

[1] See Rama (1999) for a survey of several studies on public sector downsizing.
[2] In fact, VDP's were also applied at the federal administrative level and by many states, targeting a variety of workers and locations. Particular emphasis, however, was given to state enterprises, since the government had a special interest in privatizing some of those firms and needed a fast reduction in their payroll. See Carneiro and Gill (1999) for a description of public sector downsizing programs in Brazil.
[3] Earlier attempts to estimate the wage loss expected to be suffered by workers in Brazil who joined a VDP focused on measures of the overall public-private wage gap, both gross and controlled by observed characteristics (see Foguel et al., 2000).
4 The details of that survey are described in section II and more fully in Amadeo, Camargo and Gonzaga (1998).

These results are in a certain way expected. It is a well-documented fact that Brazilian public employees earn on average more than workers at the private sector (see Foguel *et al.*, 2000). As mentioned above, unobservable factors might have also influenced this outcome. Program participants may have had only public sector experience, and thus, may have only accumulated public-sector-specific human capital. In fact, program participants had an average of 18 years of job tenure at RFFSA. Their previous experience in the public sector may also have stigmatized them.

A comparison between public and private sectors in Brazil and other Latin American countries reveals that wage inequality among workers in the private sector is typically significantly higher than among public employees (see Panizza & Qiang, 2000).. Therefore, it is not surprising that the earnings distribution widened among VDP participants who left the RFFSA.

In fact, for analyzing the changes in the earnings distributions of former RFFSA workers, it is necessary to first understand what happened to the earnings distributions of all workers in the Brazilian labor market. Only with this as a basis for comparison it is possible to investigate the reasons for both the wage decrease and the increased earnings inequality observed for former RFFSA employees.

The main task of this paper is to compare the outcomes of the former RFFSA employees with those in the public and private sectors in Brazil. In this paper, we use the RFFSA survey data set, in combination with information from national household surveys, to study the process of absorption by the private sector of RFFSA workers who joined a VDP. This is accomplished through the use of a decomposition analysis, based on micro-simulations, which attempts to identify the role of observable and unobservable characteristics, as well as their remuneration in the public and the private sector, in explaining the observed changes in the earning distributions of the former RFFSA employees.

This paper is organized as follows. The next section describes the data. Section III discusses the overall research methodology. Section IV presents and interprets the empirical results. Section V concludes.

II. Data description

Two data sets are used in this study. The first one, henceforth referred as REDE, contains information on more than 5,000 workers out of around 7,000 who left their jobs at RFFSA through the Voluntary Dismissal Program from 1995 to 1997. The survey was conducted in March 1998 and collected data on several labor market variables, including earnings, employment status, and duration of unemployment and employment spells. It has information on schooling levels, grouped into nine categories, which correspond to the following ranges of years of schooling: 0-3, 4, 5-7, 8, 9-10, 11, 12-14, 15, and more than 15.[5] REDE also contains official administrative information for all workers who joined the VDP, including their last wages at RFFSA.

The other data set, *Pesquisa Nacional de Amostra por Domicílio* (PNAD), is the annual Brazilian Household Survey, collected by IBGE. PNAD covers the whole country (except some rural areas) and it is the largest and most important Brazilian household survey, interviewing more than 75,000 households every year, which corresponds to about 300,000 individuals. PNAD collects data every September since 1976, with the exception of Census years and a couple of other years. It contains information on labor, demographic, educational and regional variables and has been widely used in many micro-econometric studies.

In this study, we use information from both data sets only for individuals in the age range of 25-64, with a paid main activity that requires at least 20 weekly hours of work. Two yearly editions of PNAD were selected, 1995 and 1997. These two years were chosen in order to approximately match both the timing of exit from RFFSA and the absorption of these workers by the private sector, as captured by the REDE survey.

Data on earnings correspond to positive monthly labor earnings on the main activity. This variable was properly deflated and evaluated at March 1998, the date of the REDE survey. In the PNAD data set, outliers (more than R$ 250,000 a month) are not considered in the analysis. In the REDE data set, we only use information about workers who switched to the private sector after quitting, as answered in the survey of March 1998.

Since REDE is restricted to some Federation States, only observations in the same geographical range are used in PNAD. In practice, this means that the North Region and some States in the Mid-West are not considered here. Finally, the classification at PNAD for a job at the public sector has the broadest sense, meaning that not only federal public servants are considered as working in the public sector but also any individual who declared to work at that sector with or without a formal labor contract.

Table 1 displays sample means for the most important variables in the two data sets, and the number of observations remaining after the application of the filters described above.

---

5 In fact, there are ten categories but the first two (illiteracy and less than four years of schooling) were merged into one group here. The categories corresponding to isolated years (4, 8, 11 and 15) reflect the attainment of some school degree in Brazil (elementary, medium, high school and college, respectively).

Table 1: Sample means and number of observations (REDE and PNAD)

Variables		REDE		PNAD			
		1995	1997	1995		1997	
		Public Sector	Private Sector	Public Sector	Private Sector	Public Sector	Private Sector
Earnings (R$)		967	757	672	412	617	428
Age		42.98	40.51	39.39	39.24	39.56	39.28
Years of Schooling	**0-3**	8.1%	4.9%	12.9%	34.7%	11.2%	31.6%
	4	10.1%	8.7%	11.0%	21.0%	9.2%	19.1%
	5-7	13.4%	13.4%	5.8%	11.7%	6.1%	13.3%
	8	14.5%	14.4%	7.9%	9.1%	8.0%	9.8%
	9-10	9.1%	9.5%	3.8%	3.3%	3.8%	3.7%
	11	22.2%	23.2%	27.5%	12.0%	28.2%	13.6%
	12-14	4.8%	5.9%	7.6%	2.1%	7.6%	2.3%
	15	13.7%	14.9%	16.4%	3.8%	17.5%	4.0%
	16 or more	4.0%	5.1%	7.2%	2.2%	8.6%	2.6%
		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Percentage of Men		89.5%	91.4%	44.5%	66.7%	44.9%	66.6%
# of observations		4103	1399	10565	58121	13121	73184

III. Overall Research Methodology

In this section, we describe our methodological approach. The section is divided into four sub-sections. In the first one, the regression model is presented describing the relationship between earnings, observable individual characteristics (such as age, gender and schooling) and unobservable characteristics. The second sub-section briefly discusses the method of comparing distributions using counterfactual micro-simulations. The third sub-section shows how the distribution parameters are estimated and discusses the hypotheses behind the use of regression residuals in the counterfactual earnings distributions. Finally, the last sub-section describes how the earnings distributions at different periods and sectors are compared.

III.1. The earnings equation

In order to estimate the effects of observable and non-observable individual characteristics on earnings,[6] the following model is specified:

(1) $\log(y_{i,t,s,p}) = w_{i,t,s,p} = \mathrm{h}_1(\boldsymbol{x}_{i,t,s,p}) + \mathrm{h}_2(\boldsymbol{z}_{i,t,s,p})$.

where w denotes log-earnings, $\boldsymbol{x}$ is a vector of observable characteristics and $\boldsymbol{z}$ is a vector of non-observable characteristics. This additively separable model holds by assumption for all individuals

[6] Henceforth, sometimes referred as "observables" and "unobservables".

i, at time t, at sector s (public and private), and for both available data sets p (REDE and PNAD). We also assume that these variables are i.i.d. and that:

(2) $w_{i,t,s,p} = \mathrm{h}_1(\boldsymbol{x}_{i,t,s,p}) + \boldsymbol{e}_{i,t,s,p}$; $\mathrm{E}[w_{t,s,p} \mid \boldsymbol{x}_{t,s,p}] = \mathrm{h}_1(\boldsymbol{x}_{t,s,p})$; and $\mathrm{Var}[w_{t,s,p} \mid \boldsymbol{x}_{t,s,p}] = \boldsymbol{s}^2_{t,s,p}$

i.e. that for a triple (t,s,p) every individual faces the same distribution, that the unobservable and observable characteristics are not correlated, and that the errors are homoscedastic.

The vector of observable individual characteristics, $\boldsymbol{x}$, consists of age, squared age, dummy variables for schooling groups, and a dummy variable for gender, which takes the value of 1 for males.[7] An intercept is also allowed in the model, meaning that the first entry in the vector $\boldsymbol{x}$ is a constant 1 for every observation for every triple (t,s,p). Imposing the restriction that the conditional mean is linear and additively separable on its arguments, we have:

(3) $\mathrm{h}_1(\boldsymbol{x}_{i,t,s,p}) = \boldsymbol{b}'\boldsymbol{x}_{i,t,s,p} = \boldsymbol{a}_{t,s,p} + \boldsymbol{b}_{1t,s,p}{*}a + \boldsymbol{b}_{2t,s,p}{*}a^2 + \boldsymbol{b}_{3t,s,p}{*}d_2 + \ldots + \boldsymbol{b}_{10t,s,p}{*}d_9 + \boldsymbol{b}_{11t,s,p}{*}g$

where $\boldsymbol{a}$ is the intercept and $\boldsymbol{b}$'s are the linear regression coefficients on age (a), squared age (a^2), dummy variables for schooling groups (d_2 to d_9) and gender (g).[8]

In practice, six regressions are estimated, since for the PNAD data set there are two sectors (public and private) at each period (1995 and 1997), while for the REDE data set workers are in the public sector (at RFFSA) in 1995 and in the private sector in 1997.

II.2. Comparing two distributions using counterfactual micro-simulations

In this sub-section, we show how we use the information from log-earnings equations to compare two distinct labor income distributions. Consider the log-earnings of an individual i at the triple (t_0,s_0,p_0). In order to facilitate notation, say $m_0 = (t_0,s_0,p_0)$. Then using (3)

(4) $w_{i,m0} = \boldsymbol{b}_{m0}'\boldsymbol{x}_{i,m0} + \boldsymbol{e}_{i,m0}$

Now define:

(5) $w_{\mathrm{SIM1}i,m0} = \boldsymbol{b}_{m1}'\boldsymbol{x}_{i,m0} + \boldsymbol{e}_{i,m0}$;

(6) $w_{\mathrm{SIM2}i,m0} = \boldsymbol{b}_{m1}'\boldsymbol{x}_{i,m0} + \boldsymbol{e}_{j,m1}$.

Equation (5) simulates what would be the earnings of an individual i at the triple $m_0 = (t_0,s_0,p_0)$ if she is paid $\boldsymbol{b}_{m1}$ instead of $\boldsymbol{b}_{m0}$ for her observable characteristics, where $\boldsymbol{b}_{m1}$ is the vector of the rates of return on human capital variables at the triple $m_1 = (t_1,s_1,p_1)$. This is usually referred

[7] As discussed in the previous section, the choice of schooling groups instead of a continuous schooling variable is made due to data collection constraints.

as a “price effect”. As this can be generalized for every individual *i* at the triple m_0, we come up with a “simulated” earnings distribution, with its own mean and inequality measures.

The same is true for equation (6), which “simulates” what would be the earnings of a worker *i* if she faces not only the “prices” of observable characteristics at the labor market described by m_1 but also receives the “prices” for her unobservable characteristics at that market, which is given by $\boldsymbol{e}_{j,m1}$. The implementation of that “simulated” labor earnings distribution is not as trivial as the former one, since it now involves not only the price effect, but also a complex choice of how the *e*’s should be allocated to each worker. We keep it to sub-section III.3 the description of the method of allocating the residuals used in this paper.

The previous equations allow us to compare the earnings distribution of workers at a labor market characterized by m_0 with workers at m_1, using the “simulated” earnings as intermediate steps. In order to see that note that using (4):

$$w_{i,m0} = \boldsymbol{b}_{m0}'\boldsymbol{x}_{i,m0} + \boldsymbol{e}_{i,m0} =$$

$$= \boldsymbol{b}_{m0}'\boldsymbol{x}_{i,m0} - \boldsymbol{b}_{m1}'\boldsymbol{x}_{i,m0} + \boldsymbol{b}_{m1}'\boldsymbol{x}_{i,m0} + \boldsymbol{e}_{i,m0} =$$

(7) $$w_{i,m0} = (\boldsymbol{b}_{m0} - \boldsymbol{b}_{m1})'\boldsymbol{x}_{i,m0} + w_{\text{SIM1}\ i,m0} =$$

$$= (\boldsymbol{b}_{m0} - \boldsymbol{b}_{m1})'\boldsymbol{x}_{i,m0} + \boldsymbol{b}_{m1}'(\boldsymbol{x}_{i,m0} - \boldsymbol{x}_{j,m1}) + \boldsymbol{b}_{m1}'\boldsymbol{x}_{j,m1} + \boldsymbol{e}_{i,m0} =$$

$$= w_{j,m1} + (\boldsymbol{b}_{m0} - \boldsymbol{b}_{m1})'\boldsymbol{x}_{i,m0} + \boldsymbol{b}_{m1}'(\boldsymbol{x}_{i,m0} - \boldsymbol{x}_{j,m1}) + (\boldsymbol{e}_{i,m0} - \boldsymbol{e}_{j,m1}).$$

Then,

(8) $$w_{i,m0} - w_{j,m1} = (\boldsymbol{b}_{m0} - \boldsymbol{b}_{m1})'\boldsymbol{x}_{i,m0} + \boldsymbol{b}_{m1}'(\boldsymbol{x}_{i,m0} - \boldsymbol{x}_{j,m1}) + (\boldsymbol{e}_{i,m0} - \boldsymbol{e}_{j,m1}) =$$

(8a) $$w_{i,m0} - w_{j,m1} = (\boldsymbol{b}_{m0} - \boldsymbol{b}_{m1})'\boldsymbol{x}_{i,m0} + \boldsymbol{b}_{m1}'(\boldsymbol{x}_{i,m0} - \boldsymbol{x}_{j,m1}) + (w_{\text{SIM1}\ i,m0} - w_{\text{SIM2}\ i,m0}).$$

Equation (7) shows the familiar “controlled individual wage gap” expression, now commonly used in the literature (see, for instance, Foguel *et al.*, 2000, and Panizza and Qiang, 2000). For every individual *i*, given her own observable characteristics, the controlled wage gap is the difference between her actual log-earnings and the log-earnings she would have obtained at labor market m_1, that is, $w_{i,m0} - w_{\text{SIM1}\ i,m0}$. According to (7), this controlled wage gap turns out to be just the difference in the coefficients of both regressions times the vector of the observable characteristics for the individual *i* at m_0.

Equations (8) and (8a) show how to decompose the “gross individual wage gap”, that is, the difference in log-earnings between two individuals, *i* and *j*, who are in two distinct labor markets,

[8] Notice that one schooling dummy variable is dropped (d_1) to allow for an intercept.

m_0 and m_1, respectively. Equation (8) reveals that this difference can be decomposed in three parts. The first one is the just mentioned controlled wage gap. The second component is the individual version for the so-called "composition effect": it is the part of the gross wage gap that reflects differences in the observable characteristics of two distinct workers. Finally, the last term reflects earnings differences that can be attributed to unobservable characteristics. As shown in the next sub-section, this last term reflects either qualitative differences in those unobservable characteristics or differences in their "prices" at each labor market, or both.

Now consider the first and second moments of the four distributions, w_{m0}, w_{m1}, $w_{\mathrm{SIM1}m0}$ and $w_{\mathrm{SIM2}m0}$. Note that we have dropped the subscripts i and j, making use of the i.i.d. assumption. After computing the first moment, the equation for the gross mean wage gap becomes:

(9) $\mathrm{E}[w_{m0}] - \mathrm{E}[w_{m1}] = (\boldsymbol{b}_{m0} - \boldsymbol{b}_{m1})'\mathrm{E}[\boldsymbol{x}_{m0}] + \boldsymbol{b}_{m1}'(\mathrm{E}[\boldsymbol{x}_{m0}] - \mathrm{E}[\boldsymbol{x}_{m1}])$.

The first term on the right hand side is the mean version for the controlled wage gap and the second term is the analogous composition effect. The differences between (9) and (8) are basically that (9) computes the wage gap for means of two different labor income distributions and because of that the third term in (8) vanishes, since $\mathrm{E}[\boldsymbol{e}_{m0}] = \mathrm{E}[\boldsymbol{e}_{m1}] = 0$.

The variance of log-earnings is one of the two measures of inequality used in this paper. Its population expression for equations (4), (5) and (6) are (under the homoscedasticity assumption):

(10) $\mathrm{Var}[w_{m0}] = \boldsymbol{b}_{m0}'\mathrm{Var}[\boldsymbol{x}_{m0}]\boldsymbol{b}_{m0} + \boldsymbol{s}^2_{m0}$

(10a) $\mathrm{Var}[w_{\mathrm{SIM1}m0}] = \boldsymbol{b}_{m1}'\mathrm{Var}[\boldsymbol{x}_{m0}]\boldsymbol{b}_{m1} + \boldsymbol{s}^2_{m0}$

(10b) $\mathrm{Var}[w_{\mathrm{SIM2}m0}] = \boldsymbol{b}_{m1}'\mathrm{Var}[\boldsymbol{x}_{m0}]\boldsymbol{b}_{m1} + \boldsymbol{s}^2_{m1}$

Equation (10) gives the traditional formula for the variance decomposition in a between effect ($\mathrm{Var}[\mathrm{E}[w \mid \boldsymbol{x}]] = \boldsymbol{b}'\mathrm{Var}[\boldsymbol{x}]\boldsymbol{b}$) and a within effect ($\boldsymbol{s}^2$), under the assumption that $\mathrm{E}[\boldsymbol{e} \mid x_k] = 0$ for $k = 1, \ldots, \mathrm{K}$; where K is the number of observable variables.

Equations (10a) and (10b) are analogous formulas to (10), giving the variances for the "simulated" log-earnings w_{SIM1} and w_{SIM2}. Similarly to the first moment equations, these two equations are used to decompose the effects of a difference in variances in three parts: a price effect on the observables; a composition effect; and a price effect on the unobservables. Therefore, the difference between (10a) and (10) gives the effect on inequality of a change in the returns of the observables. The difference between (10b) and (10a) gives the effects of a change in the returns of the unobservables. Finally, the difference between equation (10) with another group subscript, say

m_1, and equation (10b) gives the effects of a change in the variance of the observables (composition effect).

III.3. Estimation procedures

The earnings equations can be estimated by ordinary least squares, since (a) the regression function is assumed to be linear and separable; (b) the errors are assumed to be uncorrelated with regressors; and (c) identically and independent distributed.

For estimating changes in means and variances, one needs to have an estimate for $E[\exp(\boldsymbol{x})]$, the vector of the means for the observables. A natural choice is $E_N[\exp(\boldsymbol{x})]$, the sample means.[9] Note that exp(.) is used since earnings, $y = \exp(w)$, are the variable of interest.

The variance decomposition analysis described in the previous subsection has two steps. First, the actual and simulated distributions $w_{i,m0}$, $w_{j,m1}$, $w_{\text{SIM1}i,m0}$ and $w_{\text{SIM2}i,m0}$ are calculated. Then, the respective measures of inequality for each distribution are computed. These measures are the variance of the log-earnings (Var[w_m]) and the Theil-T index, which formula is described in the appendix.

The method of allocating the residuals in the simulated distributions, based on Juhn, Murphy and Pierce (1993), is as follows. First, the errors $\boldsymbol{e}$ are estimated by the regression residuals e. Second, for estimating $w_{\text{SIM2}i,m0}$ and its component $e_{j,m1}$, a crucial hypothesis is required.

Since an individual j from m_1 is not observed at m_0, there are several ways of designating the $e_{j,m1}$ that will be allocated to the individual i as her unobserved component in $w_{\text{SIM2}i,m0}$. The way used here is to sort each residual distribution in an ascending order and then match the residuals according to their ranking in those orderings. In other words, workers with residuals between the p^{th} and the $p+1^{\text{th}}$ percentiles of the residual distribution at m_0 are assigned the residuals between the p^{th} and the $p+1^{\text{th}}$ percentiles of the residual distribution m_1. As a practical matter, since the sample sizes typically differ, we rearrange the observations for each labor market m into 100 different groups. Each group corresponds to a quantile of the residual distribution at m and has received average values.

This method of allocating the residuals entails the hypothesis that the unobserved characteristics at m_1 have the same distribution at m_0. The only difference between the unobserved

---

[9] An important technical issue must be mentioned at this point. In fact, the procedure for estimating means is as follows. First the distributions $w_{i,m0}$, $w_{j,m1}$, $w_{\text{SIM1}i,m0}$ and $w_{\text{SIM2}i,m0}$ are calculated. Then, each one is exponentiated, that is, the earnings are recovered from log-earnings. Finally, means are computed. This procedure is equivalent to what is described above, but for the presumed small positive term $E[\exp(\boldsymbol{e})]$.

characteristics in two distinct labor markets is basically their "prices" at each market. To illustrate with an example, this assumption means that if ability is an unobservable variable,[10] ability is assumed to have the same distribution in both the public and private sectors. The only difference is that one sector remunerates better this unobserved ability than the other. Therefore, for the purpose of this study, computing differences between $w_{\mathrm{SIM1}m0}$ and $w_{\mathrm{SIM2}m0}$ is relevant in the sense that it allows measuring the role of specific labor environments and unmeasured characteristics on earnings inequality.

III.4. Decomposing changes in the earnings distribution

In the previous sub-sections, we have shown how to decompose the differences in means and of inequality measures of two distinct distributions in several components. Another insightful way of comparing differences between two parameters of distinct distributions is the differences-in-differences method. In this case, a third distribution is used and one uses the fact that a difference in means, $E[y_{m0}] - E[y_{m1}]$, is equal to the double difference $(E[y_{m0}] - E[y_{m2}]) - (E[y_{m1}] - E[y_{m2}])$. This method can be useful when the goal is to identify partial effects that otherwise would be confounded.

In this case, one can decompose the differences in parameters of two specific distributions into several distinct partial effects. In this paper, the main interest is the difference between the distributions (*1995, public, REDE*) and (*1997, private, REDE*). It is possible to decompose it into 4 steps. For example, one could compare the RFFSA employees with all workers in the public sector in 1995; then examine how the public sector evolved between 1995 and 1997; and finally compare the public and private sectors in. 1997. In fact, there are two ways of doing this decomposition, which are described below:

(11) $E[y_{1995,\ public,\ REDE}] - E[y_{1997,\ private,\ REDE}] =$

<u>Case 1:</u>

= $E[y_{1995,\ public,\ REDE}] - E[y_{1995,\ public,\ PNAD}]$ (step 1)

+ $E[y_{1995,\ public,\ PNAD}] - E[y_{1995,\ private,\ PNAD}]$ (step 2)

+ $E[y_{1995,\ private,\ PNAD}] - E[y_{1997,\ private,\ PNAD}]$ (step 3)

+ $E[y_{1997,\ private,\ PNAD}] - E[y_{1997,\ private,\ REDE}]$ (step 4)

<u>Case 2:</u>

= $E[y_{1995,\ public,\ REDE}] - E[y_{1995,\ public,\ PNAD}]$ (step 1, same as Case 1)

+ $E[y_{1995,\ public,\ PNAD}] - E[y_{1997,\ public,\ PNAD}]$ (step 2)

+ $E[y_{1997,\ public,\ PNAD}] - E[y_{1997,\ private,\ PNAD}]$ (step 3)

[10] Remember from assumption on the joint distribution of ***x*** and ***z*** that ability is uncorrelated with the regressors.

+ $E[y_{1997,\ private,\ PNAD}] - E[y_{1997,\ private,\ REDE}]$ (step 4, same as Case 1)

The difference between cases 1 and 2 is therefore in steps 2 and 3. In case 1, step 3 measures the temporal evolution of earnings on the private sector while step 2 compares the private and the public sector in 1995. In case 2, step 3 measures the difference between the public and the private sector in 1997 while step 2 captures the temporal evolution of earnings in the public sector.

IV. Estimation results

This section presents the results of the decomposition analysis. First, log-earnings regressions are estimated. Then, we compute each difference in means and in variances of the distributions in equation (11), including all 4 steps in the two cases considered. Finally, we decompose the differences in means into a controlled wage gap (a price effect on the observables) and a composition effect; and the differences in variances into a price effect on the observables, a composition effect, and a price effect on the unobservables.

Table 2 shows the log-earnings equation estimates for the six regressions described above. All coefficients have the expected sign and are significantly different from zero at the 5% level, with the exceptions of age and the dummy for 5-7 years of schooling in the (1997, private sector, REDE) regression.

Table 2: Log-Earnings regression results

Variables		REDE		PNAD			
		1995	1997	1995		1997	
		Public Sector	Private Sector	Public Sector	Private Sector	Public Sector	Private Sector
Intercept		3,89	4,54	2,13	2,60	2,75	2,90
		(0.18)	*(0.71)*	*(0.13)*	*(0.06)*	*(0.11)*	*(0.05)*
Age		0,08	0,03	0,10	0,09	0,08	0,08
		(0.01)	*(0.03)*	*(0.01)*	*(0.00)*	*(0.01)*	*(0.00)*
Squared Age (*10000)		-6,70	-2,61	-10,02	-9,65	-8,24	-8,23
		(0.94)	*(4.19)*	*(0.78)*	*(0.32)*	*(0.67)*	*(0.29)*
Years of Schooling	4	0,13	0,26	0,46	0,54	0,31	0,53
		(0.03)	*(0.13)*	*(0.03)*	*(0.01)*	*(0.03)*	*(0.01)*
	5-7	0,25	0,15	0,60	0,70	0,49	0,67
		(0.03)	*(0.12)*	*(0.04)*	*(0.01)*	*(0.03)*	*(0.01)*
	8	0,39	0,51	0,92	0,93	0,82	0,90
		(0.03)	*(0.02)*	*(0.03)*	*(0.01)*	*(0.03)*	*(0.01)*
	9-10	0,48	0,71	1,00	1,01	0,92	1,00
		(0.03)	*(0.13)*	*(0.04)*	*(0.02)*	*(0.04)*	*(0.02)*
	11	0,52	0,82	1,27	1,31	1,18	1,31
		(0.02)	*(0.12)*	*(0.03)*	*(0.01)*	*(0.02)*	*(0.01)*
	12-14	0,61	0,88	1,82	1,79	1,63	1,76
		(0.03)	*(0.14)*	*(0.03)*	*(0.02)*	*(0.03)*	*(0.02)*
	15	1,03	1,38	2,02	2,16	1,89	2,13
		(0.03)	*(0.12)*	*(0.03)*	*(0.02)*	*(0.02)*	*(0.02)*
	16 or more	1,30	1,84	2,30	2,41	2,20	2,38
		(0.04)	*(0.14)*	*(0.03)*	*(0.02)*	*(0.03)*	*(0.02)*
Dummy for Men		0,15	0,41	0,70	0,64	0,64	0,61
		(0.02)	*(0.09)*	*(0.01)*	*(0.01)*	*(0.01)*	*(0.01)*
# of observations		4103	1399	10565	58121	13121	73184
R square		0,46	0,22	0,52	0,41	0,51	0,40

Note: The dependent variable is log-earnings. Standard errors are in parentheses.

Table 3 presents the decomposition effects for the differences in means. The table reveals that the absorption by the private sector of those workers who left RFFSA through the VDP was painful. The first row of the table shows that those workers have their average wages reduced from monthly R$ 967 to R$ 757 and that decrease was mainly due to a price effect. The simulated average earnings of a worker who got a job in the private sector in 1997, but with observed average characteristics of those who left RFFSA in 1995, fell from R$ 967 to R$ 543. This corresponds to a controlled wage gap of R$ 425 a month, which means that the rates of return of those observed characteristics were much less valued in the private sector relative to RFFSA.

The composition effect, on the other hand, played a compensatory role of R$ 214. This means that the workers that got a job in the private sector were those with better-evaluated observable characteristics.

Table 3 also reports that RFFSA workers were better paid than other workers in the public sector in 1995. In this case, there is both a price effect and a composition effect, contributing 54% an 46%, respectively, in explaining the gross wage differential. In fact, as shown in Tables 1 and 2, the composition effect operates in favor of RFFSA workers not because of better productive characteristics but due to a combination of a higher percentage of men at RFFSA and a discrimination effect against women.

It can also be seen from Table 3 that the public-private wage gap fell dramatically between 1995 and 1997. The controlled wage differential of R$ 64 or approximately 10% of the average earnings in the public sector in 1995, decreased and became negative R$ 24 in 1997. This means that the wage premium paid at the public sector disappeared in this period.[11] As the same Table 3 reveals, this was due to a sharp real wage reduction in the public sector (R$ 89 decrease in the controlled wage gap) and not to significant better payments in the private sector (R$ 4 increase in the controlled wage gap).

Table 3: Average earnings

Case	Step	m_0	m_1	Y_{m0} (A) R$	Y_{m1} (B) R$	Y_{SIM1m0} (C) R$	Gross Wage Gap (A) – (B) R$	Controlled Wage Gap (A) – (C) R$	Composition Effect (C) – (B) R$
-	-	*1995, public, REDE*	*1997, private, REDE*	967	757	543	210	425	-214
1 and 2	1	*1995, public, REDE*	*1995, public, PNAD*	967	672	807	295	160	135
1	2	*1995, public, PNAD*	*1995, private, PNAD*	672	412	608	260	64	196
1	3	*1995, private, PNAD*	*1997, private, PNAD*	412	428	409	-15	4	-19
2	2	*1995, public, PNAD*	*1997, public, PNAD*	672	617	582	55	89	-35
2	3	*1997, public, PNAD*	*1997, private, PNAD*	617	428	641	190	-24	213
1 and 2	4	*1997, private, PNAD*	*1997, private, REDE*	428	757	454	-329	-27	-303

Finally, the last row of Table 3 reveals that workers who left RFFSA received, on average, R$ 329 a month more than workers in the private sector as a whole in 1997. This is mostly due to a composition effect, which contributes with 92% of the observed total wage gap. In fact, these workers have, on average, higher levels of observed human capital characteristics (like more years of schooling) than the private sector as a whole, besides having a larger proportion of males.

[11] Foguel *et al.* (2000) reported the most recent published evidence on still large public-private wage gaps in a study that uses data from PNAD 1995. The nominal wage freeze adopted by the government for most categories in the public sector since the beginning of 1995, combined with a 2-digit (declining) level annual inflation in 1995-96, put a heavy toll on public sector workers, eroding the public-private wage premium in less than two years.

Table 4 presents the results of the earnings inequality decompositions, for the two measures used: the variance of log-earnings and the Theil-T index. Since the variance of log-earnings is a measure that depends on the value of the mean of a distribution, we only comment the results based on the Theil-T indicator, which has the property of independence from the mean.

Table 4 reveals that the earnings distribution became much more uneven for those former RFFSA workers. The Theil index jumps from 0.11 to 0.39. The simulated distributions decomposition shows that a larger dispersion for the price of the unobservables is mostly responsible for this result. In fact, the observables price effect has an almost insignificant impact on the earnings distribution, increasing the Theil index from 0.11 to 0.12. The unobservables price effect, on the other hand, makes the inequality measure jump to 0.58.

Table 4: Earnings inequality

	Case	Step	m_0	m_1	Y_{m0}	Y_{m1}	$Y_{SIM1\,m0}$	$Y_{SIM2\,m0}$
Var(log)	-	-	*1995, public, REDE*	*1997, private, REDE*	0.15	0.74	0.16	0.76
	1 and 2	1	*1995, public, REDE*	*1995, public, PNAD*	0.15	0.57	0.20	0.61
	1	2	*1995, public, PNAD*	*1995, private, PNAD*	0.57	0.68	0.57	0.67
		3	*1995, private, PNAD*	*1997, private, PNAD*	0.68	0.69	0.68	0.69
	2	2	*1995, public, PNAD*	*1997, public, PNAD*	0.57	0.52	0.56	0.52
		3	*1997, public, PNAD*	*1997, private, PNAD*	0.52	0.69	0.52	0.67
	1 and 2	4	*1997, private, PNAD*	*1997, private, REDE*	0.69	0.74	0.67	0.74
Theil-T	-	-	*1995, public, REDE*	*1997, private, REDE*	0.11	0.39	0.12	0.58
	1 and 2	1	*1995, public, REDE*	*1995, public, PNAD*	0.11	0.26	0.16	0.45
	1	2	*1995, public, PNAD*	*1995, private, PNAD*	0.26	0.36	0.25	0.33
		3	*1995, private, PNAD*	*1997, private, PNAD*	0.36	0.35	0.36	0.36
	2	2	*1995, public, PNAD*	*1997, public, PNAD*	0.26	0.26	0.26	0.25
		3	*1997, public, PNAD*	*1997, private, PNAD*	0.26	0.35	0.26	0.35
	1 and 2	4	*1997, private, PNAD*	*1997, private, REDE*	0.35	0.39	0.35	0.41

In fact, workers at RFFSA were in an environment with very low levels of inequality even when compared with the public sector as a whole. As expected, this was also due to a small dispersion on payments to unobservables at RFFSA, since an individual firm probably values less unobservable dimensions than the entire private or public sector do.

Another interesting result reported in Table 4 is the lower public sector earnings inequality relative to the private sector, which is also observed in most countries. Note, however, that there was a sharp increase in the public sector's Theil index from 0.11 to 0.26 between 1995 and 1997, which might be reflecting a differentiated ability of some categories of public workers to obtain nominal wage increases (see footnote 10).

Finally, we notice that there is no big difference between the measures of earnings inequality among former RFFSA workers when compared to the private sector as a whole. In fact, this last result completes our picture of what resembles a full and fair absorption of former RFFSA workers by the private sector in Brazil, at least as given by the first two moments of simulated distributions. Not only average wages were typical of workers in the private sector with identical observed characteristics, but also wage dispersion was very similar.

V. Concluding remarks

The objective of this paper was to study the process of absorption by the private sector of RFFSA workers who joined a Voluntary Dismissal Program. The results presented clearly reveal that the move to the private sector of these former RFFSA workers caused a decline in average wages and an increase in wage dispersion. However, it is by no means a symptom of discrimination against former public employees at the private sector. In fact, it seems that the previously accumulated general human capital of these workers was decisive for their remuneration at the private sector, and that their lack of private sector-specific human capital did not prevent them from being fairly paid.

All results show that these workers were paid accordingly to their embodied observable characteristics and faced the same "prices" for those characteristics that any other worker in the private sector would face on average. Also, the increase in wage dispersion brought them to levels of earnings inequality that are not significantly different from that existing in the private sector as a whole.

As an aside conclusion, this study finds that the public-private sector wage gap rapidly eroded between 1995 and 1997. However, the inequality in both sectors has not changed over the selected years. It is also important to mention that in order to explain the gap in the inequality between the public and private sectors, the important factor that should be taken into consideration is the within-group component, or, in other words, the differences in the "prices" of the unobservable characteristics.

References

Amadeo, E., Camargo, J.M., & Gonzaga, G. (1998), "Projeto de Pesquisa da Situação Laboral dos Empregados que Aderiram ao Plano de Incentivo ao Desligamento da RFFSA, Nono Relatório de Atividades". PUC-Rio, Rio de Janeiro, June.

Bourguignon, F. (1979) "Decomposable Income Inequality Measures", *Econometrica*, 47(4): 901-920, July.

Carneiro, F. & Gill, I. (1999) "Public Sector Downsizing in Brazil", Texto para Discussão No.6, Série Economia, UCB, Brasília, August.

Foguel, M., Gill, I., Mendonça, R.& Barros, R. P. (2000), "The Public-Private Wage Gap in Brazil", *Revista Brasileira de Economia*, 54(4): 433-472, October- December.

Juhn, C., Murphy, K. and Pierce, B. (1993), "Wage Inequality and the Rise in Returns to Skill", *Journal of Political Economy*, 101(3): 410-442, June.

Panizza, U. and Qiang, Z.W. (2000), "Spoiled Bureaucrats and Exploited Women? Public Sector Premium and Gender Gap in Latin America", LACEA meetings, Rio de Janeiro, mimeo.

Rama, M. (1999) "Public Sector Downsizing: An Introduction", *The World Bank Economic Review*, 13(1): 1-22, January.

APPENDIX: The Theil Inequality Indicator

We present here the Theil-T index formula. Let X be the variable of interest and w the weights or the relative frequency of each value of X. Then, the X distribution has the following Theil-T index:

$$t(X) = \sum_{i=1}^{n} w_i \left(\frac{X_i}{\overline{X}}\right) \ln\left(\frac{X_i}{\overline{X}}\right) = \frac{\sum_{i=1}^{n} w_i X_i \ln(X_i)}{\sum_{i=1}^{n} w_i X_i} - \ln\left(\sum_{i=1}^{n} w_i X_i\right)$$

For further references on that inequality measure see, for instance, Bourguignon (1979).

A summary of the new estimation procedures and results to be incorporated in the next version of the paper (sessions III and IV):

III. Estimation Procedures

a. "Differences-in-differences" approach:

1st approach:

=	$E[y_{1995,\ public,\ REDE}] - E[y_{1995,\ public,\ PNAD}]$	(*within Public Sector, 1995*)
+	$E[y_{1995,\ public,\ PNAD}] - E[y_{1995,\ private,\ PNAD}]$	(*between sectors, 1995*)
+	$E[y_{1995,\ private,\ PNAD}] - E[y_{1997,\ private,\ PNAD}]$	(*temporal evolution, Private Sector*)
+	$E[y_{1997,\ private,\ PNAD}] - E[y_{1997,\ private,\ REDE}]$	(*within Private Sector, 1997*)

2nd approach:

=	$E[y_{1995,\ public,\ REDE}] - E[y_{1995,\ public,\ PNAD}]$	(*within Public Sector, 1995*)
+	$E[y_{1995,\ public,\ PNAD}] - E[y_{1997,\ public,\ PNAD}]$	(*temporal evolution, Public Sector*)
+	$E[y_{1997,\ public,\ PNAD}] - E[y_{1997,\ private,\ PNAD}]$	(*between sectors, 1997*)
+	$E[y_{1997,\ private,\ PNAD}] - E[y_{1997,\ private,\ REDE}]$	(*within Private Sector, 1997*)

b. Controlled Wage Gaps (or Average Treatment Effects, ATE):

Parametric Approaches:

i) Oaxaca Decomposition:

$$E[y_{m0}|\mathbf{X}] - E[y_{m1}|\mathbf{X}] = (\boldsymbol{\beta}_{m0} - \boldsymbol{\beta}_{m1})'\mathbf{x}_{m0} + \boldsymbol{\beta}_{m1}'(\mathbf{x}_{m0} - \mathbf{x}_{m1}) =$$

$$= -(\boldsymbol{\beta}_{m1} - \boldsymbol{\beta}_{m0})'\mathbf{x}_{m1} - \boldsymbol{\beta}_{m0}'(\mathbf{x}_{m1} - \mathbf{x}_{m0})$$

Then estimate the Controlled Wage Gap as:

i.: $\mathbf{X}_{\mathbf{m}0}(\hat{\boldsymbol{\beta}}_{\mathbf{m}0} - \hat{\boldsymbol{\beta}}_{\mathbf{m}1})$ or

ii. $-\mathbf{X}_{\mathbf{m}1}(\hat{\boldsymbol{\beta}}_{\mathbf{m}1} - \hat{\boldsymbol{\beta}}_{\mathbf{m}0})$

ii) Pooling $\mathbf{m_0}$ and $\mathbf{m_1}$ (estimating *gamma* coefficient):

$$\mathbf{E}[\mathbf{y} \mid \mathbf{X}, 1(\mathbf{m}_0)] = \boldsymbol{\gamma}\, 1(\mathbf{m}_0) + \mathbf{X}\boldsymbol{\beta}$$

Then estimate $\boldsymbol{\gamma}$ by OLS, for instance.

iii) Pooling $\mathbf{m_0}$ and $\mathbf{m_1}$ and using Propensity Score (estimating *rho* coefficient):

1[st] step: Estimate $\mathbf{E}[1(\mathbf{m}_0) \mid \mathbf{X}\,] = \Pr[1(\mathbf{m}_0) = 1] = \mathbf{F}(\mathbf{X}\boldsymbol{\delta})$, assuming some functional form for F(.) (for instance, logistic distribution). The fitted value, $\hat{\mathbf{p}}(\mathbf{X})$, is the **estimated P-score**.

2[nd] step: Estimate $\mathbf{E}[\mathbf{y} \mid \mathbf{X}, 1(\mathbf{m}_0)] = \boldsymbol{\rho}\, 1(\mathbf{m}_0) + \boldsymbol{\lambda}\hat{\mathbf{p}}(\mathbf{X})$

Semi-Parametric Approaches:

1) Stratifying on P-score:

1[st] step: Arrange m0 and m1 observations into J quantiles of $\hat{\mathbf{p}}(\mathbf{X})$;

2[nd] step: Verify balance of X within each quantile;

3[rd] step: Estimate ATE as:

i. $\sum_{\mathbf{j}=1}^{\mathbf{J}} \frac{\sum_{\mathbf{i}\in\mathbf{j}} 1_{\mathbf{i}}(\mathbf{m}0)}{\sum_{\mathbf{k}=1}^{\mathbf{J}}\sum_{\mathbf{i}\in\mathbf{k}} 1_{\mathbf{i}}(\mathbf{m}0)} [\overline{\mathbf{y}}_{\mathbf{mo}} - \overline{\mathbf{y}}_{\mathbf{m}1}]$ or

ii. $-\sum_{\mathbf{j}=1}^{\mathbf{J}} \frac{\sum_{\mathbf{i}\in\mathbf{j}} 1_{\mathbf{i}}(\mathbf{m}1)}{\sum_{\mathbf{k}=1}^{\mathbf{J}}\sum_{\mathbf{i}\in\mathbf{k}} 1_{\mathbf{i}}(\mathbf{m}1)} [\overline{\mathbf{y}}_{\mathbf{m}1} - \overline{\mathbf{y}}_{\mathbf{m}0}]$

2) Reweighing using P-score:

1[st] step: Estimate counterfactual distribution (see DiNardo, Fortin and Lemieux, 1996) using $\hat{\mathbf{p}}(\mathbf{X})$;

2[nd] step: Estimate ATE as:

i. $\overline{\mathbf{y}}_{\mathbf{m}0} - \overline{\mathbf{y}}^*_{\mathbf{m}1}$ or

ii. $-(\overline{\mathbf{y}}_{\mathbf{m}1} - \overline{\mathbf{y}}^*_{\mathbf{m}0})$

c. Actual and Counterfactual Measures of Dispersion:

1) "Oaxaca's style:

<u>Actual</u>: Var ($\mathbf{y}_{\mathbf{m}0}$), Var ($\mathbf{y}_{\mathbf{m}1}$)

<u>Counterfactual</u>: Var ($\mathbf{y}^*_{\mathbf{mo}}$), Var($\mathbf{y}^*_{\mathbf{m}1}$), where:

$$\mathbf{y}^*_{\mathbf{mo}} = \mathbf{X}_{\mathbf{mo}}\boldsymbol{\beta}_{\mathbf{m}1} + \boldsymbol{\varepsilon}_{\mathbf{mo}}$$

and

$$\mathbf{y}^*_{\mathbf{m}1} = \mathbf{X}_{\mathbf{m}1}\boldsymbol{\beta}_{\mathbf{m}0} + \boldsymbol{\varepsilon}_{\mathbf{m}1}$$

2) Dinardo, Fortin, Lemieux (1996):

<u>Actual</u>: Var ($\mathbf{y}_{\mathbf{m}0}$), Var ($\mathbf{y}_{\mathbf{m}1}$)

<u>Counterfactual</u>: Var ($\mathbf{y}^*_{\mathbf{mo}}$), Var($\mathbf{y}^*_{\mathbf{m}1}$), where:

$\mathbf{y}^*_{\mathbf{mo}}$ is the vector of individual log-earnings had $\mathbf{m}_1$ had the **X**'s attributes of $\mathbf{m}_0$ but the actual wage structure of $\mathbf{m}_1$. Analogy applies to $\mathbf{y}^*_{\mathbf{m}1}$

IV. New Results

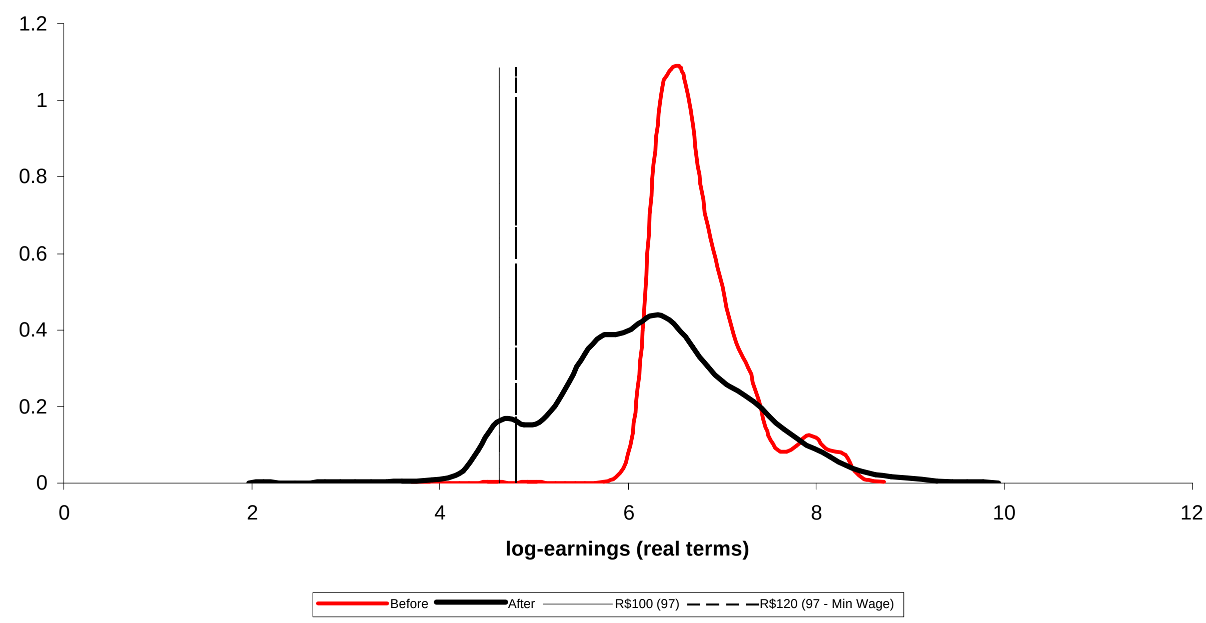
A. REDE: Public 1995 X Private 1997

m_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0*)	$(Y_{m0}/Y_{m1}) - 1$ (base: *m1*)	$y_{m0} - y_{m1}$
		Means				Wage Gaps		
1995, public, REDE	*1997, private, REDE*	994	841	6.76	6.25	0.15	0.18	0.51

Variance

m_0	m_1	Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	$\ln(Y^*m1)$ (base: *m0*)	$\ln(Y^*m0)$ (base: *m1*)
1995, public, REDE	*1997, private, REDE*	684	1138	0.24	0.93	0.34	0.83

Log-earnings density function - REDE Employees: before and after leaving public sector



B. 1995: REDE-Public X PNAD - Private

M_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0)*	$(Y_{m0}/Y_{m1})-1$ (base: *m1)*	$y_{m0}-y_{m1}$
		Means				Wage Gaps		
1995, public, REDE	***1995, private, PNAD***	994	567	6.76	5.72	0.43	0.75	1.03

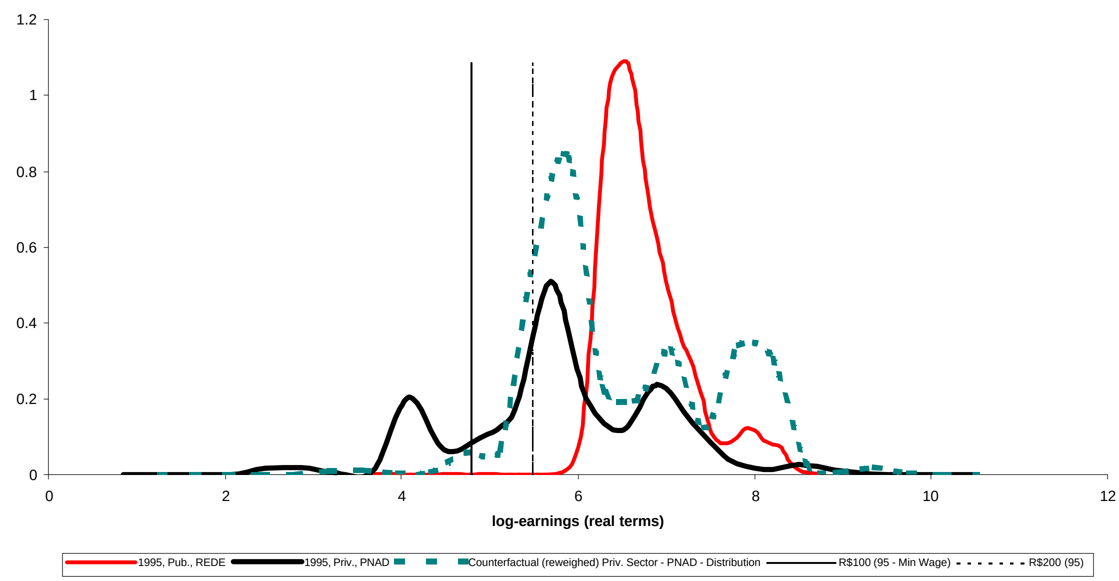
Controlled Wage Gaps

M_0	m_1	ln(Y*m0*)- ln(Y**m1*) (base: *m0)*	- (ln(Y**m0*) - ln(Y*m1*)) (base: *m1)*	gamma	rho	Stratifying base: *m0*	Stratifying base: *m1*	reweighing base: *m0*	reweighing base: *m1*
1995, public, REDE	***1995, private, PNAD***	0.26	0.71	0.27	0.23	0.23	-	0.18	0.60

Variance

Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	ln(Y**m1*) (base: *m0)*	ln(Y**m0*) (base: *m1)*	Reweighing (base: *m0)*	Reweighing (base: *m1)*
685	973	0.24	1.13	0.48	0.79	1.03	0.01

Log-earnings density function: Public REDE X Private PNAD in 1995



C. Within Public Sector (1995): REDE X PNAD

M_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0*)	$(Y_{m0}/Y_{m1}) - 1$ (base: *m1*)	$y_{m0} - y_{m1}$
		Means				Wage Gaps		
1995, public, REDE	*1995, public, PNAD*	994	750	6.76	6.08	0.25	0.33	0.68

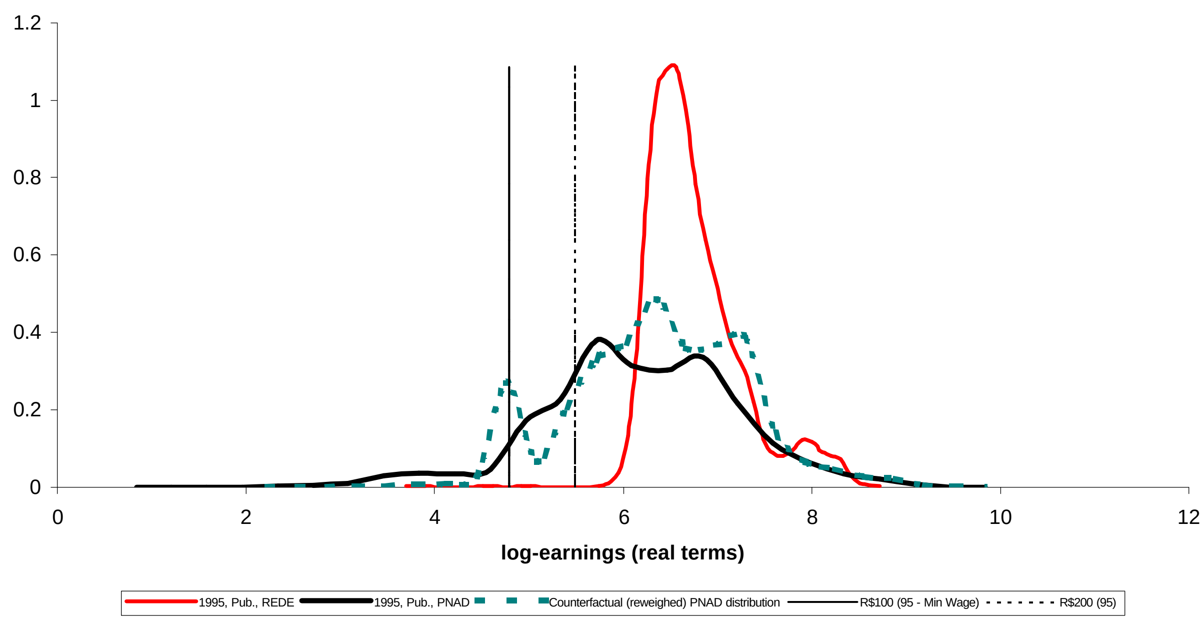
Controlled Wage Gaps

M_0	m_1	ln(Y*m0*)- ln(Y**m1*) (base: *m0*)	- (ln(Y**m0*) - ln(Y*m1*)) (base: *m1*)	gamma	rho	Stratifying base: *m0*	Stratifying base: *m1*	reweighing base: *m0*	reweighing base: *m1*
1995, public, REDE	*1995, public, PNAD*	0.29	0.54	0.29	0.26	0.23	0.55	0.30	0.38

Variance

Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	ln(Y**m1*) (base: *m0*)	Ln(Y**m0*) (base: *m1*)	Reweighing (base: *m0*)	Reweighing (base: *m1*)
684	972	0.24	1.12	0.46	0.72	0.80	0.14

Log-earnings density function: Public Sector 1995 - REDE X PNAD



D. Public X Private Sectors, 1995

M_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0)*	$(Y_{m0}/Y_{m1}) - 1$ (base: *m1)*	$y_{m0} - y_{m1}$
		Means				**Wage Gaps**		
1995, public, PNAD	***1995, private, PNAD***	750	567	6.08	5.72	0.24	0.32	0.36

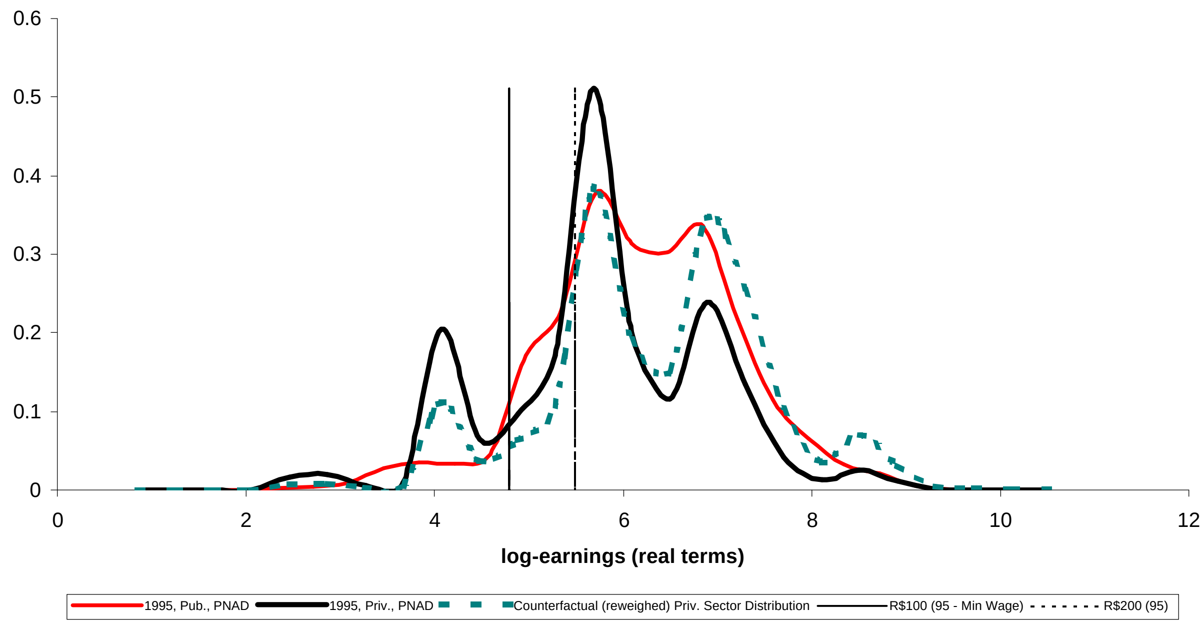
Controlled Wage Gaps

M_0	m_1	ln(Y*m0*)- ln(Y**m1*) (base: *m0)*	- (ln(Y**m0*) - ln(Y*m1*)) (base: *m1)*	gamma	rho	stratifying base: *m0*	stratifying base: *m1*	reweighing base: *m0*	reweighing base: *m1*
1995, public, PNAD	***1995, private, PNAD***	-0.06	-0.04	-0.06	-0.06	-0.06	-0.03	-0.17	-0.10

Variance

Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	ln(Y**m1*) (base: *m0)*	ln(Y**m0*) (base: *m1)*	Reweighing (base: *m0)*	Reweighing (base: *m1)*
972	973	1.12	1.13	1.13	1.13	1.33	0.99

Log-earnings density function: Public X Private sectors in 1995



E. Temporal Evolution of Public Sector, 1995-1997

M_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0*)	$(Y_{m0}/Y_{m1}) - 1$ (base: *m1*)	$y_{m0} - y_{m1}$
		Means				**Wage Gaps**		
1995, public, PNAD	*1997, public, PNAD*	750	766	6.08	6.15	-0.02	-0.02	-0.07

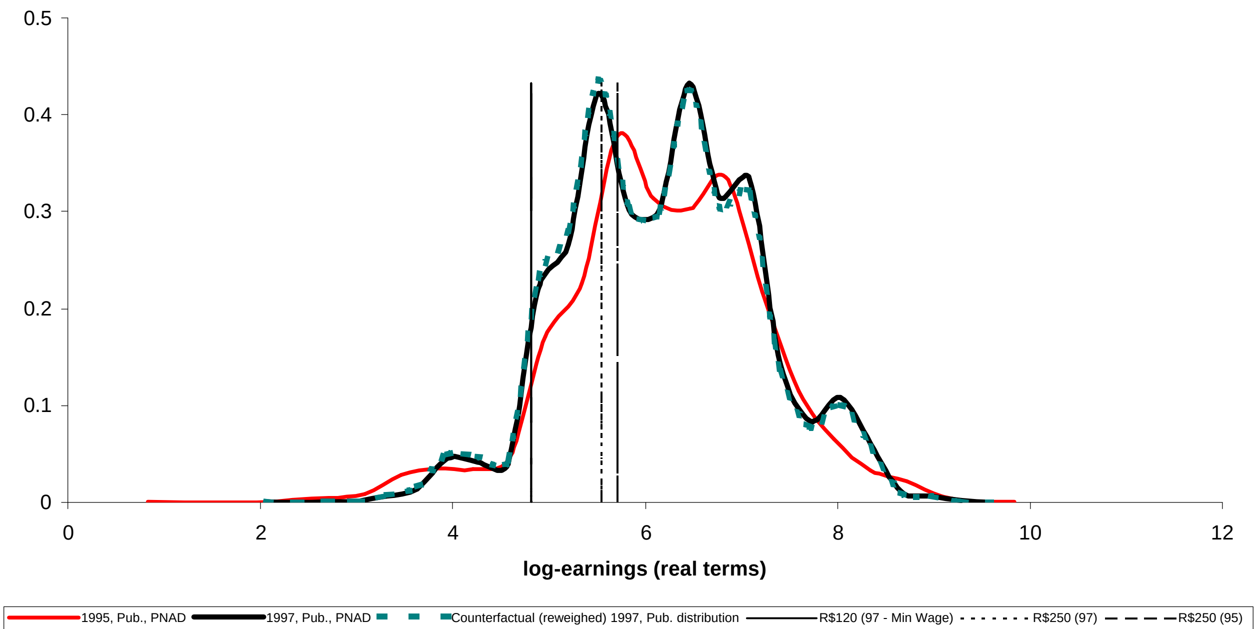
Controlled Wage Gaps

M_0	m_1	ln(Y*m0*)- ln(Y**m1*) (base: *m0*)	- (ln(Y**m0*) - ln(Y*m1*)) (base: *m1*)	gamma	rho	stratifying base: *m0*	stratifying base: *m1*	reweighing base: *m0*	reweighing base: *m1*
1995, public, PNAD	*1997, public, PNAD*	-0.01	-0.01	0.02	0.00	-0.01	-0.01	-0.03	-0.03

Variance

Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	ln(Y**m1*) (base: *m0*)	ln(Y**m0*) (base: *m1*)	Reweighing (base: *m0*)	Reweighing (base: *m1*)
972	923	1.12	1.00	1.07	1.04	1.00	1.11

Log-earnings density function: Public Sector - 1995 X 1997



F. Temporal Evolution of Private Sector, 1995-1997

M_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0)*	$(Y_{m0}/Y_{m1})-1$ (base: *m1)*	$y_{m0}-y_{m1}$
		Means				**Wage Gaps**		
1995, private, PNAD	*1997, private, PNAD*	567	577	5.72	5.74	-0.02	-0.02	-0.02

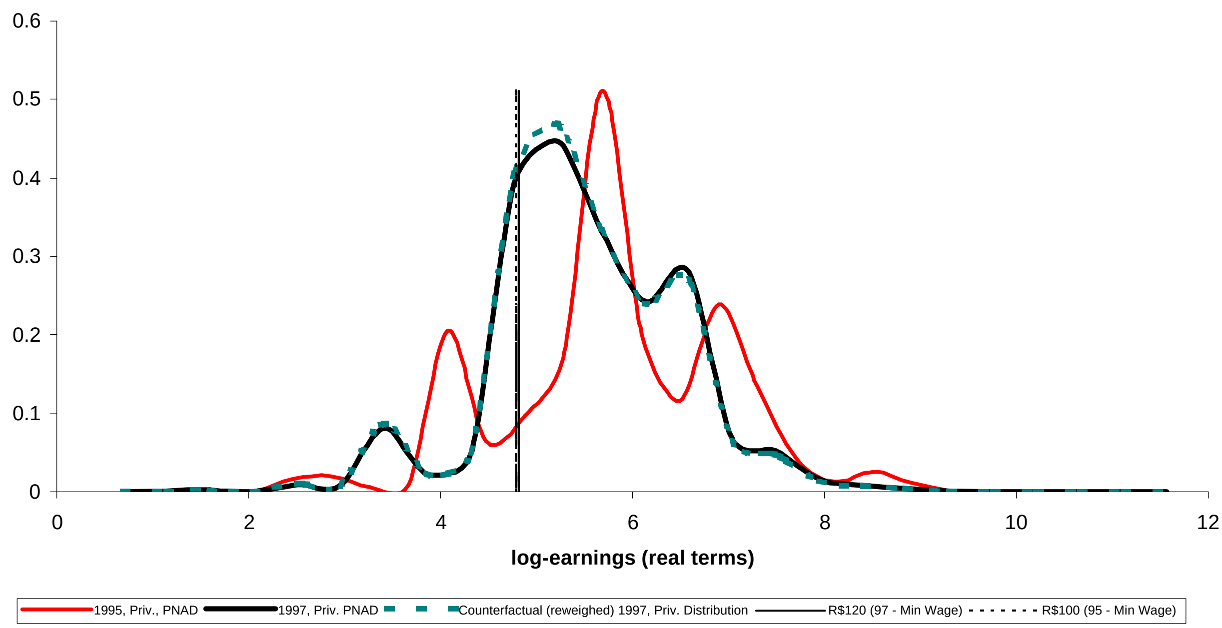
Controlled Wage Gaps

m_0	M_1	ln(Y*m0*)- ln(Y**m1*) (base: *m0)*	- (ln(Y**m0*) - ln(Y*m1*)) (base: *m1)*	gamma	rho	stratifying base: *m0*	stratifying base: *m1*	reweighing base: *m0*	reweighing base: *m1*
1995, private, PNAD	*1997, private, PNAD*	0.02	0.02	-0.01	0.02	0.02	0.02	0.03	0.03

Variance

Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	Ln(Y**m1*) (base: *m0)*	ln(Y**m0*) (base: *m1)*	Reweighing (base: *m0)*	Reweighing (base: *m1)*
973	1059	1.13	1.15	1.12	1.16	1.14	1.15

Log-earnings density function: Private Sector - 1995 X 1997



G. Public X Private Sectors, 1997

m_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0*)	$(Y_{m0}/Y_{m1}) - 1$ (base: *m1*)	$y_{m0} - y_{m1}$
		Means				Wage Gaps		
1997, public, PNAD	*1997, private, PNAD*	766	577	6.15	5.74	0.25	0.33	0.41

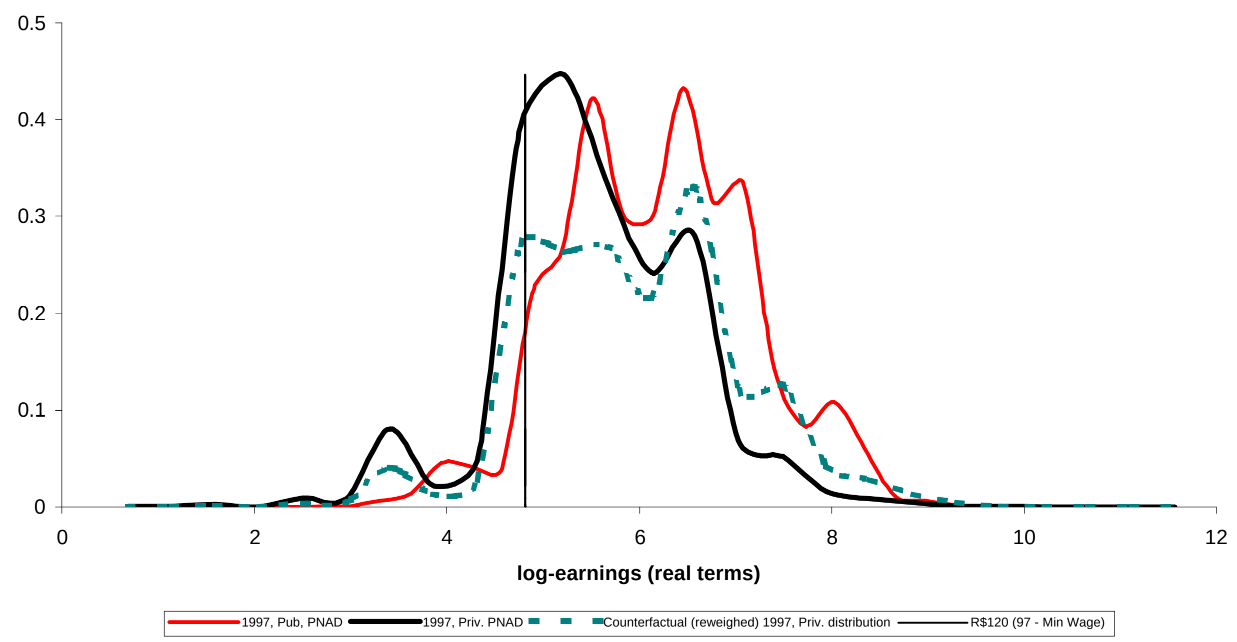
Controlled Wage Gaps

m_0	m_1	ln(Y*m0*)- ln(Y**m1*) (base: *m0*)	- (ln(Y**m0*) - ln(Y*m1*)) (base: *m1*)	gamma	rho	stratifying base: *m0*	stratifying base: *m1*	Reweighing base: *m0*	Reweighing base: *m1*
1997, public, PNAD	*1997, private, PNAD*	-0.04	-0.02	-0.03	-0.03	-0.04	-0.01	-0.11	-0.05

Variance

Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	ln(Y**m1*) (base: *m0*)	ln(Y**m0*) (base: *m1*)	Reweighing (base: *m0*)	Reweighing (base: *m1*)
923	1059	1.00	1.15	1.05	1.16	1.31	0.89

Log-earnings density function: Public X Private sectors in 1997



H. Within Private Sector (1997): PNAD X REDE

m_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0*)	$(Y_{m0}/Y_{m1}) - 1$ (base: *m1*)	$y_{m0} - y_{m1}$
		Means				Wage Gaps		
1997, private, PNAD	*1997, private, REDE*	577	841	5.74	6.25	-0.46	-0.31	-0.51

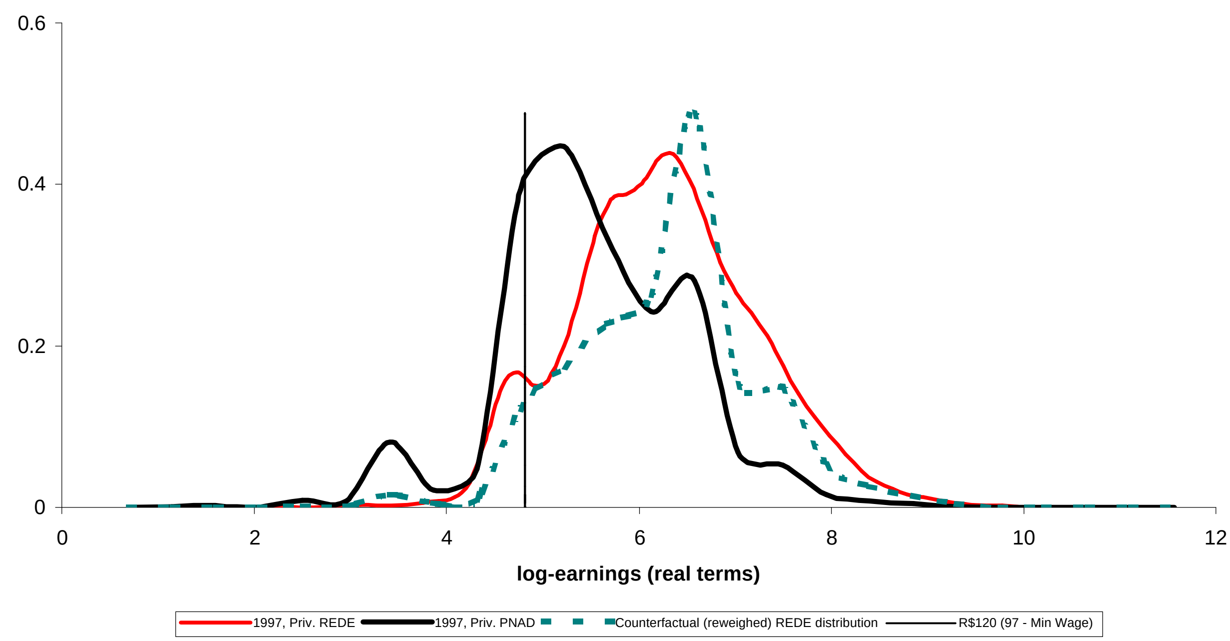
Controlled Wage Gaps

m_0	M_1	ln(Y*m0*)- ln(Y**m1*) (base: *m0*)	- (ln(Y**m0*) - ln(Y*m1*)) (base: *m1*)	gamma	rho	stratifying base: *m0*	stratifying base: *m1*	reweighing base: *m0*	reweighing base: *m1*
1997, private, PNAD	*1997, private, REDE*	-0.08	0.21	0.21	0.24	0.22	0.13	0.24	0.28

Variance

Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	Ln(Y**m1*) (base: *m0*)	ln(Y**m0*) (base: *m1*)	Reweighing (base: *m0*)	Reweighing (base: *m1*)
1059	1138	1.15	0.93	0.91	1.06	0.15	1.03

Log-earnings density function: REDE X PNAD in 1997



I. 1997: REDE-Private X PNAD - Public

M_0	m_1	Y_{m0} (R$)	Y_{m1} (R$)	y_{m0}	y_{m1}	$-(Y_{m1}/Y_{m0}-1)$ (base: *m0*)	$(Y_{m0}/Y_{m1}) - 1$ (base: *m1*)	$y_{m0} - y_{m1}$
		Means				**Wage Gaps**		
1995, public, REDE	*1995, private, PNAD*	841	766	6.25	6.15	0.09	0.10	0.10

Controlled Wage Gaps

M_0	m_1	ln(Y*m0*)- ln(Y**m1*) (base: *m0*)	- (ln(Y**m0*) - ln(Y*m1*)) (base: *m1*)	gamma	rho	Stratifying base: *m0*	Stratifying base: *m1*	reweighing base: *m0*	reweighing base: *m1*
1995, public, REDE	*1995, private, PNAD*	-0.19	-0.03	-0.19	-0.21	-0.20	-	-0.20	-0.47

Variance

Y_{m0}	Y_{m1}	$\ln(Y_{m0})=y_{mo}$	$\ln(Y_{m1})=y_{m1}$	ln(Y**m1*) (base: *m0*)	ln(Y**m0*) (base: *m1*)	Reweighing (base: *m0*)	Reweighing (base: *m1*)
1138	923	0.93	1.00	1.03	0.78	0.74	0.39

Log-earnings density function: Private REDE X Public PNAD in 1995