FUNDAÇÃO
GETULIO VARGAS
EPGE
Escola de Pós-Graduação
em Economia

Ensaios Econômicos

Escola de
Pós Graduação
em Economia
da Fundação
Getulio Vargas

Nº 532 ISSN 0104-8910

Regional or educational disparities? A counterfactual exercise

Angelo José Mont'Alverne
Pedro Cavalcanti Ferreira
Márcio Antônio Salvato

Março de 2004

Regional or Educational Disparities? A Counterfactual Exercise*

Angelo José Mont'Alverne Duarte † Pedro Cavalcanti Ferreira ‡

Márcio Antônio Salvato§

December 2, 2003

Abstract

This work investigates the impact of schooling on income distribution in states/regions of Brazil. Using a semi-parametric model, discussed in DiNardo, Fortin & Lemieux (1996), we measure how much income differences between the Northeast and Southeast regions – the country's poorest and richest – and between the states of Ceará and São Paulo in those regions – can be explained by differences in schooling levels of the resident population. Using data from the National Household Survey (PNAD), we construct counterfactual densities by reweighting the distribution of the poorest region/state by the schooling profile of the richest. We conclude that: (i) more than 50% of the income difference is explained by the difference in schooling; (ii) the highest deciles of the income distribution gain more from an increase in schooling, closely approaching the wage distribution of the richest region/state; and (iii) an increase in schooling, holding the wage structure constant, aggravates the wage disparity in the poorest regions/states.

JEL Classification Codes: C14; I20 e J31.

*We gratefully acknowledge the comments and suggestions of Luiz Henrique Braido, Luiz Renato Lima, Carlos Eugênio da Costa, Samuel Pessôa, Marcelo Fernandes, Renato Flôres and the participants in a seminar at EPGE/FGV. All remaining errors are ours. We thank CNPq, Banco Central do Brasil and PUC.Minas for financial support.

†EPGE/FGV and Banco Central do Brasil, e-mail: aduarte@fgvmail.br

‡EPGE/FGV, e-mail: ferreira@fgv.br

§EPGE/FGV, PUC.Minas and CEPE, e-mail: salvato@fgvmail.br

1 Introduction

Disparities in income between regions of Brazil, notably between the Southeast and the North and Northeast, have been a longstanding subject for debate among academics, policymakers and members of the political class. As far back as the mid-19th century, Emperor Dom Pedro II remarked that he would sell his crown to feed the hungry northeasterners. Over this century and a half, many diagnoses have been conducted and programs implemented trying to diminish this inequality. Government policies intensified and became systematic starting in the 1950s. Following the lead of the economist Celso Furtado, the document "Uma Política para o Desenvolvimento do Nordeste" (" A Policy to Develop the Northeast") was published, which suggested industrialization as a way to reduce income disparities between the Northeast and Southeast. This document inspired the creation of two agencies – SUDENE (Northeast Development Superintendency) and SUDAM (Amazon Development Superintendency) – along with two regional development banks, the BNB and BASA. The 1988 Federal Constitution chipped in by determining that of 3% of revenues from taxes on income and manufactured products must be allocated to programs to finance the productive sector in the North, Northeast and Midwest regions.

Despite the large number of governmental organs and programs at the federal and state levels devoted to regional economic development, government policies implemented over the past 50 years have mainly focused on the same instruments: tax incentives and public credits to subsidize private initiative, along with direct government spending on infrastructure projects. It is estimated that between 1989 and 2002, the constitutional funds for regional development spent some US$ 10 billion. Although successful in some respects, such as accelerating industrialization of the Northeast and North, these policies have not managed to transform social indicators by reducing the rate of poverty and changing the distribution of income in these regions.

In the past decade, a new approach has emerged to the problem of regional inequality in Brazil. It is based on the thesis that low per capita income in the North and Northeast regions is related to the concentration of individuals with low schooling levels (human capital) and low physical capital, which together limit their ability to earn. In this form, reducing regional inequality cannot be divorced from combating poverty, so what is needed is a policy stressing education and professional training and programs for access to credit.

Pessoa (2000) suggests there are two facets to the problem of unequal income: the first and most important involves the difference in per capita income between the regions; and the second, less significant, refers to the spatial distribution of production. According to the author, supposing perfect labor mobility between the regions, there can only be an income differential between them if the characteristics of the individuals in the regions differs. In this way, the development policies based on subsidies and accumulation of physical capital adopted in Brazil have always tried to focus on the first problem but in reality are more suited to addressing the second facet. Barros (1993), Barros and Mendonça (1995) and Barros, Camargo and Mendonça (1997) strengthen the thesis of the importance individuals' characteristics, particularly of education level, in determining income differences, based on empirical evidence.

Bourguignon, Ferreira and Leite (2002) investigated the differences between income distributions in Brazil, the United States and Mexico using an extension of Oaxaca (1973) and of Binder (1973), which consists of counterfactual distributions constructed by replacing the original values of the distribution parameters with the distribution of another country. This allows measuring the effect on a country's income distribution if some characteristic of that country's individuals, represented

by an income distribution parameter, were equal to that of another country. The authors conclude that inequalities in human capital qualifications and transfers explain nearly 2/3 of the difference in inequality levels between Brazil and the United States.

In this work we analyze the effects of an alteration in the educational characteristics of a region[1] on its wage distribution. We use a non-parametric approach presented in the work of DiNardo, Fortin and Lemieux (1996), in which they analyzed the effects of institutional and labor-market factors, such as the minimum wage level and degree of unionization, on wage distribution in the United States. In this methodology, the effects on wages of a determined characteristic of the population or factor that influences its behavior are measured by applying a non-parametric method of estimating the density function, and reweighting the samples by the characteristic/factor of interest. This leads to constructing what is called a counterfactual distribution, which can be compared with the original wage distribution of the population. In our case, we work with the distribution of income from labor (referred to simply as income for economy), since we have no data on minimum wages for Brazil.[2]

This approach has the same foundation as the decomposition of Oaxaca (1973)[3], which is based on the construction of counterfactuals, but unlike the original idea – which worked only with the means of the distributions – we analyze the entire distribution. The methodology we employ allows us to visualize the income density function clearly and to see the changes that would occur in this distribution if there were any changes in the population's educational conditions, maintaining the original income distribution structure in place. In this sense, instead of focusing the study of income and educational differential on some index of inequality, such as the coefficients of Gini and of Theil, or the variance of the logarithm of income, we intend to derive our conclusions by observing the entire income distribution, using measures of distribution differences and the Komogorov-Smirnov distribution inequality test.

Construction of the counterfactual distributions is done by reweighting the sample according to some characteristic of interest. In other words, to study, as in our case, the effects of schooling level on income, we estimate the income distribution by reweighting the samples available in such a way that they become part of the profile of a population with the desired schooling profile. In this form, we can obtain what the income distribution of a region would be if it had the schooling characteristics of another region, maintaining its original income structure. It should be clear that there is a limitation to this methodology, since it only considers partial effects and cannot analyze the effects in general equilibrium. Nevertheless, this approach will be quite useful to answer questions of the type: What would be the income distribution in the state of Ceará if its education conditions were similar to those of the

[1] By region we mean a determined geographic area, which can be a municipality, state, set of states, geographic region or a country, about whose inhabitants we want to study a certain characteristic.

[2] In Brazil, worker pay is rarely computed in terms of an hourly wage, but rather in relation to a monthly salary. The government establishes the minimum salary (nowadays adjusted yearly) for full-time employees, and most job earnings, as well as social security benefits, are set in multiples of this minimum. So, although speaking of a minimum wage for Brazil is in the strictest sense a misnomer, the expression minimum wage can be employed without causing any distortion. However, based on the characteristics measured in our database, as indicated, we generally speak of income, meaning job income, and in some instances use minimum salary when the situation specifically warrants it.

[3] The approach of Oaxaca is generally restricted to comparing means. When the distributions are unimodal, symmetric and have similar variances, this procedure is well conceived to analyze the structure of wages. Nevertheless, one cannot expect these conditions to be valid when comparing wage distributions of different regions. This is the reason for choosing here an estimation of the entire density function.

state of São Paulo?

To answer that question, for example, we measure the change in the income distribution in Ceará and the Northeast Region if the resident population in these regions had the same schooling as those in São Paulo and the Southeast Region.[4]

In the following section we present some facts that illustrate the problem of inequality in Brazil. In Section 3 we present the data used in our analysis. In Section 4 we discuss in some detail the methodology for arriving at the weighted kernel density estimators and constructing the counterfactual densities, besides presenting some parametric and non-parametric measures used to compare the estimated densities. The results are presented in Section 5,and we conclude in Section 6.

2 Stylized Facts

The inequality among Brazilian regions and states can be seen both from indicators of the well-being of the population and their level of income. The difference between the Human Development Indices (HDI) in the regions of interest went down between 1991 and 2000. The distance between HDIs in the Northeast and Southeast, for example, which was 0.16 in 1991, decreased to 0.12 in 2000 (Table 1)[5]. Nevertheless, the relative positions of the regions have not changed since the indices were computed for the first time in 1970. Besides this, the relative position of the states also did not change much in the same period, inasmuch as the nine Northeastern states remained among the 11 units of the federation[6] with the worst HDIs throughout the interval 1971-1990. The disparity in income, measured by the coefficients of Theil and Gini[7], worsened in all Brazilian regions in the 1970s and 80s, and improved slightly in the 90s. This decline was more pronounced in the North and Northeast regions, making them even more unequal when compared with the Southeast (Table 2).

Table 1
Human Development Index - HDI

Human Development Index - HDI	Year	Brazil	Regions				
			Midwest	Northeast	North	Southeast	South
HDI-Education	1991	0.745	0.778	0.606	0.705	0.812	0.804
	2000	0.849	0.877	0.762	0.818	0.886	0.895
HDI-Lifespan	1991	0.668	0.682	0.587	0.637	0.709	0.715
	2000	0.731	0.747	0.669	0.706	0.759	0.776
HDI-Income	1991	0.681	0.699	0.564	0.614	0.732	0.689
	2000	0.723	0.747	0.614	0.640	0.768	0.747
HDI-Total	1991	0.698	0.720	0.586	0.652	0.751	0.736
	2000	0.768	0.790	0.682	0.722	0.805	0.806

Source: Authors

[4] As shall be seen in the next section, in fact we do not consider the entire Southeast Region.

[5] We calculated the HDIs in the regions following the methodology of the PNUD.

[6] The Federal District (DF), location of the nation's capital, Brasília, is counted along with the states as a unit of the federation. From here on we will simply use the term state unless specifically speaking of the DF.

[7] The value of both coefficients varies between 0, meaning no inequality, and 1, when it is at its maximum.

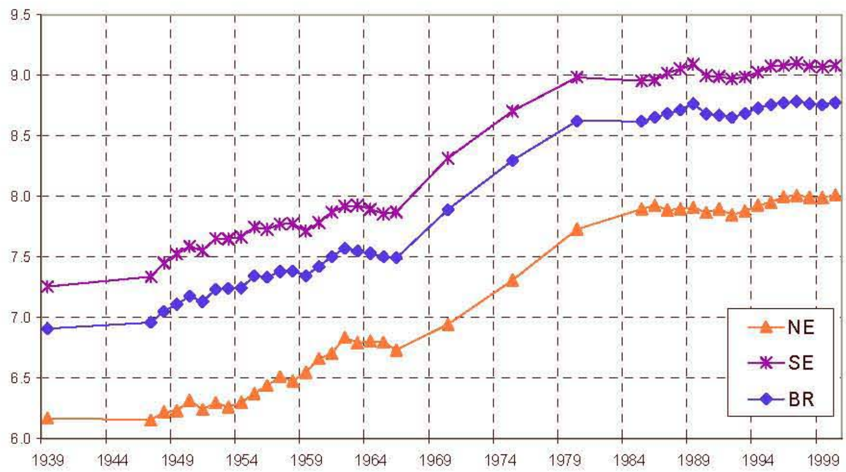
Table 2
Coefficients of Theil and Gini of Labor Income

	Theil			Gini		
	1970	1980	1991	1981	1990	2001
North	0.44	0.56	0.72	0.51	0.58	0.57
Northeast	0.57	0.65	0.78	0.57	0.63	0.60
Midwest	0.55	0.66	0.70	0.58	0.61	0.60
Southeast	0.61	0.60	0.66	0.56	0.58	0.57
South	0.53	0.58	0.63	0.54	0.58	0.55
Brazil	**0.68**	**0.70**	**0.78**	**0.58**	**0.61**	**0.60**

Source: IPEA

In terms of per capita GDP, the regional differences have remained the same since the 1940s (Graph 1). In fact, the relative participation of the Northeast in national GDP went down, from 16.7% in 1939 to 14.8% in 1960, and again to 13.1% in 2000.[8]

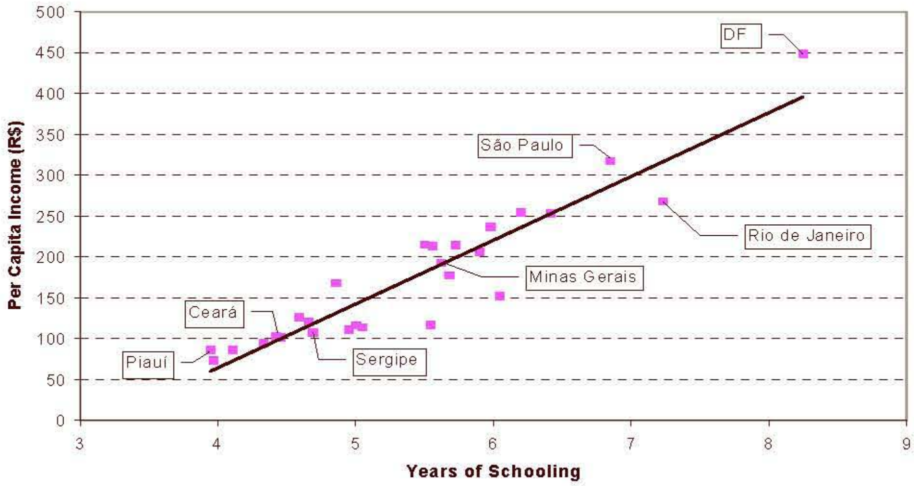
Graph 1
Per Capita Income (in log)



The data from the 2000 census show that there is a positive linear relation between the per capita income in the states and the average schooling levels of the population older than 25 (Graph 2), which corroborates the thesis that the regional income disparities reflect the differences in human capital of the people in those regions.

[8] GDP and population data are from the IBGE.

Graph 2
Per Capita Income x Average Schooling



3 Data

We used data from the National Household Survey (PNAD) for 1999 (conducted by the Brazilian Institute of Geography and Statistics – IBGE), with the weighting obtained from the results of the 2000 census. The results of the PNAD are presented both in terms of individuals and households. In the case of the individual study, which is of our interest, it gives information on general characteristics (current state of residence, age, sex, race, color, etc.)[Are you sure both race and color are included?], migration characteristics (birth state, state of residence 5 years before the study, etc.), education, work and income of the subjects. We give each individual a weight, which translates into how much persons with his/her characteristics represent in relation to the overall population.

The sample used consists of the individuals with positive income from labor in the reference month of the survey (September 1999), with known schooling level and working at least 40 hours a week. We excluded those working less than this to try to make income comparisons uniform and approximate the income measurements used here to those of wages, since this latter variable would be the ideal for this type of study (see footnote 2 above).

A total of 352,393 people were interviewed throughout the country, with 67,111 being less than 10 years old at the time, and for whom no questions were asked about labor and income. Of the 285,282 remaining people, only 173,634 were economically active[9], and of these 3,265 did not state their income and/or schooling, leaving 170,369 people. From this sample for Brazil at large, we extracted 5 subsamples according to the state of residence: Ceará State (CE); São Paulo State (SP), Northeast Region (NE); Southeast Region, excluding the state of Espírito Santo (SE1); and Southeast Region excluding the states of Espírito Santo and São Paulo (SE2). [10]

We chose the Northeast and Southeast regions because they have the lowest and highest respective per capita incomes of the five Brazilian regions. We chose

[9] The economically active population is made up of those occupied or who are actively looking for work.

[10] SE2 comprises the states of Rio de Janeiro and Minas Gerais, respectively the second and third wealthiest in the country.

the states of Ceará and São Paulo, in the Northeast and Southeast, respectively, because Ceará, despite being one of the poorest states in the country, is pointed out as an example of success in implementing development policies based on attracting private industrial investments, and São Paulo because it is the richest state in the country. We excluded Espírto Santo because it has a much less developed and industrialized economy than the rest of the Southeast states, to the point of its also being the target of pubic policies to fight regional inequalities (it is even included in the SUDENE area). We also decided to consider a sub-sample without the state of São Paulo, because it is to far ahead of all other Brazilian states in per capita income and level of industrialization. Moreover, as we shall see presently, São Paulo is the only Brazilian state in which the value of the minimum salary, which is determined exogenously by the federal government, is not a strong restriction on the distribution of work income. In the other states the mode of the labor income distribution is exactly the value of the minimum salary.

The following tables present the sample used by regions/state of interest and the main descriptive statistics, which were calculated considering the weighting of each individual.

Table 3 - Number of Interview Subjects per Region/State

	Brazil	NE	CE	SE1	SE2	SP
Persons interviewed	352,393	113,902	22,124	110,558	67,890	42,668
IAP with known income and schooling	170,369	53,210	10,163	53,162	32,658	20,504
Persons with no labor income	37,808	14,718	2,588	9,530	6,061	3,469
Persons working less than 40 hours/week	28,501	10,237	1,867	8,251	5,486	2,765
Sample size	104,060	28,255	5,708	35,381	21,111	14,270

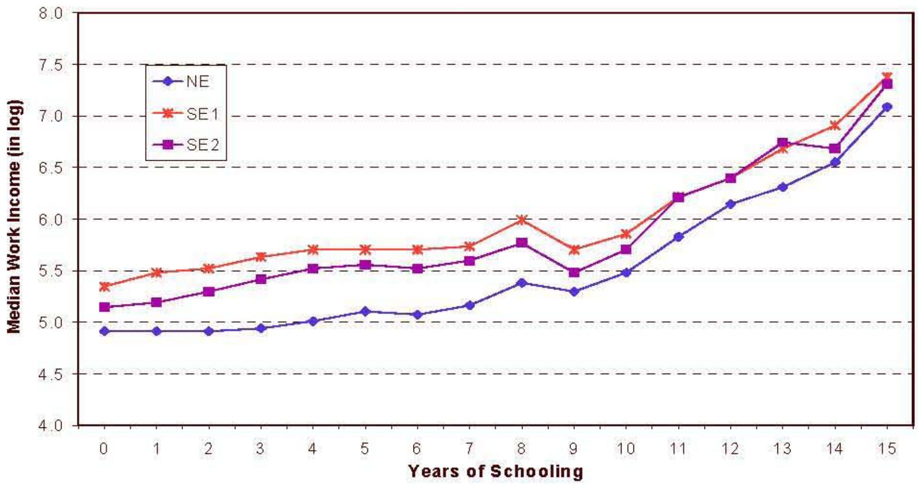
Table 4 - Descriptive Statistics of the Selected Sample

(in R$)	Brazil	NE	CE	SE1	SE2	SP
Income from labor						
mean	572	358	338	676	554	775
standard deviation	905	704	697	977	866	1.055
Gini Coeff.	0.55	0.57	0.58	0.52	0.53	0.50
Log labor income						
mean	5.81	5.3	5.21	6.04	5.83	6.22
standard deviation	0.96	0.95	0.98	0.9	0.9	0.85
Schooling (years)	6.4	4.6	4.5	7.3	6.9	7.8
Gini Coeff.	0.32	0.42	0.42	0.29	0.30	0.27

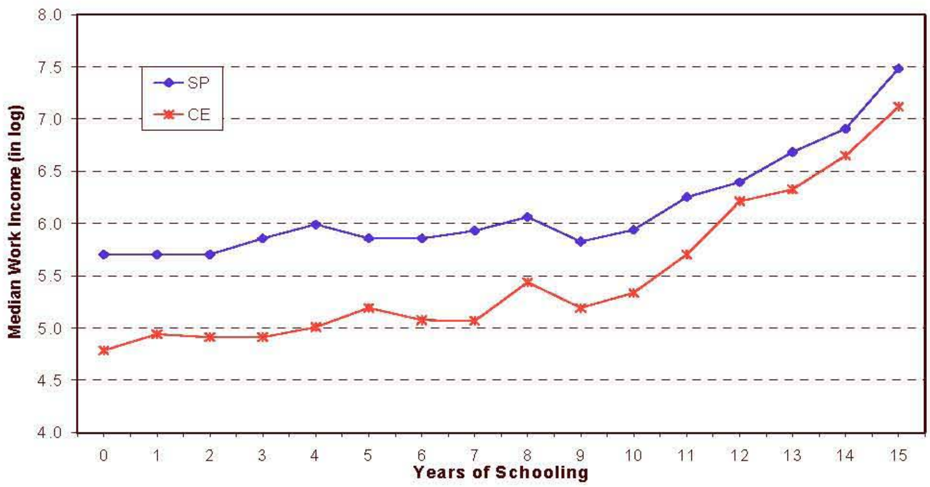
Table 4 shows that the higher the average income, the less is the income and schooling inequality, measured here by the Pearson variation coefficient and the coefficient of Gini. This permits us to raise the hypothesis that the inequality of income and schooling are strongly correlated, and in turn smaller in the richest regions/states. The average schooling in the poorest regions is nearly 3 years less than in the richest ones. Indeed, we can see that income is directly proportional to schooling, which strengthens the hypothesis that the income differential can be explained by the difference in schooling levels.

We computed for each region/state the median of the logarithm of income by level of schooling. The schooling x income curves, also known as Mincer curves, are plotted in Graphs 3(a) and 3(b). One can see that the marginal return of schooling is rising for schooling levels above 10 years, unlike observed in developed countries with high levels of schooling among the general population.

Graph 3(a) – Schooling x Income



Graph 3(b) – Schooling x Income



Although we have data on a very large number of interview subjects, as can be seen in the last line in Table 3, the amount of information contained in these interviews is quite a bit lower than it appears. Income, besides being collected as a monthly integer value, was always reported in multiples of R$ 10 or the minimum salary (R$ 136 at the time), and many figures are even stated in multiples of R$ 100. The number of income levels was between 480 for CE and 1040 for SE1.

3.1 Methodology

We used a semiparametric model to construct the counterfactual density functions. These counterfactual densities were estimated from a sample generated by taking as a base the original sample and changing the attribute we wished to study, to see the impact of schooling on the distribution of income. The method comprises two steps: the first, parametric, which involves constructing the reweighting functions; and the second, non-parametric, which consists of estimating, based on kernel functions, the density functions, as proposed by Rosenblatt (1956) and Parzen (1962).

Let f_n be the density kernel estimator of density f, whose support is the variable w, based on a random sample of size n, $\{W_1, W_2..., W_n\}$, with weighting $\theta_1, ..., \theta_n$, respectively, and where $\sum_i \theta_i = 1$. We then have:

$$\hat{f}_h(w) = \sum_{i=1}^{n} \frac{\theta_i}{h} K\left(\frac{w - W_i}{h}\right) \tag{1}$$

where h is the window and $K(\cdot)$ is the kernel function. The most used kernels are the uniform, Gaussian and that of Espanechnikov, with the choice among them being an ad-hoc decision of the econometrician, who should keep in mind the nature of the variable whose density is being estimated. In this work, following the suggestions of DiNardo, Fortin and Lemieux (1996) and Butcher and DiNardo (1998), we adopted the Gaussian kernel and worked with the logarithm of income to reduce the problem of asymmetry.

The choice of the window is an important point in estimating the kernel density, since there is a tradeoff between bias (difference between the estimated and real distribution) and variance: larger windows result in greater bias and less variance and vice-versa. There are various methods to select the window automatically, among them the cross validation and plug-in methods[11]. However, these methods are not adequate for data with the characteristics we are working with, since they are censured by intervals (grouped). The cross validation methods, for example, as pointed out by Silverman (1986), tend to generate inadequate results, $h = 0$. Thus, we used visual selection, as detailed in Butcher and DiNardo (1998): we began with a very narrow window (low smoothing), $h = 0.05$, and enlarged it until we obtained a smooth distribution, which turned out to occur at $h = 0.12$. We can justify this procedure of starting with a small window and increasing it by the belief that the human eye is better at smoothing than in the contrary sense (Butcher and DiNardo (1998)). Underlying this method we are adopting the hypothesis that the labor income distribution, and consequently the marginal productivity of labor, is smooth, which appears very reasonable given the large sample population. The choice of the lower bound of the windows that produce smooth distributions indicates that we prioritized bias over variance.

We estimated the counterfactual densities as proposed by DiNardo, Fortin and Limieux (1996), where reweighting functions are chosen for the sample. One can consider that each observation of the sample is a vector (w, z), where w represents wages (a continuous variable) and z the characteristics of each individual (we considered only education, measured in years of schooling). Hence, we have joint density distributions $F(w, z)$ for income and education. The density of income for region 1, $f_{R1}(w)$, can be written as the integral of the income density conditional on the level of schooling of the individuals, $f(w/z)$, over the distribution of education, $F(z)$:

[11] See Park and Marron (1990) and Sheater and Jones (1991) as examples.

$$
\begin{aligned}
f_{R1}(w) &= \int_{z\in\Omega_z} dF(w,z|E_{w,z}=R1) \\
&= \int_{z\in\Omega_z} f(w|z,E_w=R1)dF(z|E_z=R1) \\
&\equiv f(w;E_w=R1,E_z=R1)
\end{aligned} \tag{2}
$$

where Ω_z is the domain of the set of attributes, E_z represents the region from where the education distribution is considered and E_w represents the region from where the income distribution is considered. To study counterfactuals, we are interested in modifying the structure of characteristics and so we define $f(w;E_w=R1,E_z=R1)$ as the real income density of region 1 and $f(w;E_w=R1,E_z=R2)$ as the income density of region 1 if the education distribution were that of region 2 in the same period.

Under the hypothesis that the income structure of region 1, $f(w|z,E_w=R1)$, does not depend on the distribution of education in region 2, we can write the hypothetical density $f(w;E_w=R1,E_z=R2)$ as:

$$
\begin{aligned}
f(w;E_w &= R1,E_z=R2) = \int f(w|z,E_w=R1)dF(z|E_z=R2) \\
&\equiv \int f(w|z,E_w=R1)\Psi_z(z)dF(z|E_z=R1)
\end{aligned} \tag{3}
$$

where $\Psi_z(z)$ is a reweighting function defined by

$$
\Psi_z(z) \equiv dF(z|E_z=R2)/dF(z|E_z=R1) \tag{4}
$$

Equation (3) defines the income density in region 1 that would prevail if the educational conditions were similar to those of region 2, and as can be seen, is identical to the definition in (2), except for the reweighting function $\Psi_z(z)$. In reality, the problem of estimating the desired counterfactual density function reduces to calculating the appropriate weightings. So, we can estimate counterfactual density functions using the method of weighted kernel density estimators where we use a new weighter that contains an estimate for $\Psi_z(z)$. Thus, we have,

$$
\hat{f}(w;E_w=R1,E_z=R2) = \sum_{i\in S_{R1}} \frac{\theta_i}{h}\hat{\Psi}_z(z_i)K\left(\frac{w-W_i}{h}\right) \tag{5}
$$

where $\sum_{i\in S_{R1}} \theta_i\hat{\Psi}_z(z_i)=1$ and S_{R1} is the set of indices from the sample of individuals of region 1. The differences observed between the real density of region 1 and the hypothetical density created represents the effect of a change in the education distribution.

One can see that by applying a rule of Bayes on (4), this quotient can be written as

$$
\Psi_z(z) = \frac{\Pr(E_z=R2|z)}{\Pr(E_z=R1|z)}\frac{\Pr(E_z=R1)}{\Pr(E_z=R2)}. \tag{6}
$$

Since the level of education is a discreet variable that assumes a finite number of values, the estimation of $\Psi_z(z)$ by a probit model is equivalent to a simple counting.

There are various forms, both parametric and non-parametric, to compare the density estimates (real and counterfactual) and measure the difference between them. One can simply take the difference between the counterfactual and real densities, obtaining a complete description of the changes in the income distribution

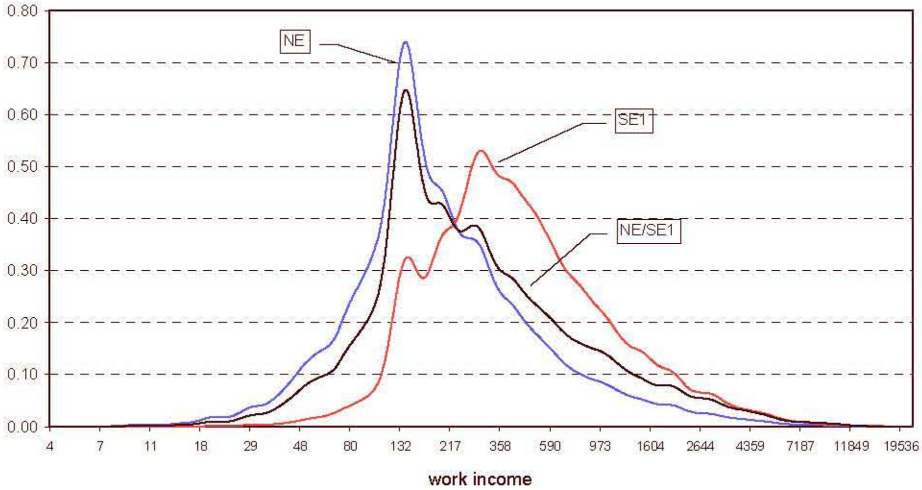
from reweighting by the schooling. Other methods of comparing estimated densities try to reduce their differences into a single number: the distance of Kullbach-Leibler, of Sibson, of Chernoff, difference between the standard deviations, differences between percentiles, differences between the percentile differentials (10-90, 10-50, 25-75, 5-95). [why does only this one not add to 100?] All these measures were used in this work. Besides this, we performed the Komogorov-Smirnov test of density inequalities.

4 Results

We applied the methodology described in the previous section and estimated the real distributions of the logarithm of income for the 5 sub-samples and for the full sample (Brazil), as presented in Section 3. Then, we estimated the counterfactual distributions for the Northeast and for Ceará, reweighting the samples by the schooling characteristics of the Southeast Regions (SE1 and SE2) and the state of São Paulo. In this section we report and comment on the results for the state of Ceará, reweighted by the schooling of São Paulo, and for the Northeast Region, reweighted by the schooling of the Southeast Region (SE1 and SE2). The results for the other cases are presented in the appendix.

Each of the graphs (4-6) below shows three densities: one counterfactual density, the real density from which this originated, and the real density of the region used in the reweighting[12]. The horizontal axis of the graphs is logarithmic.

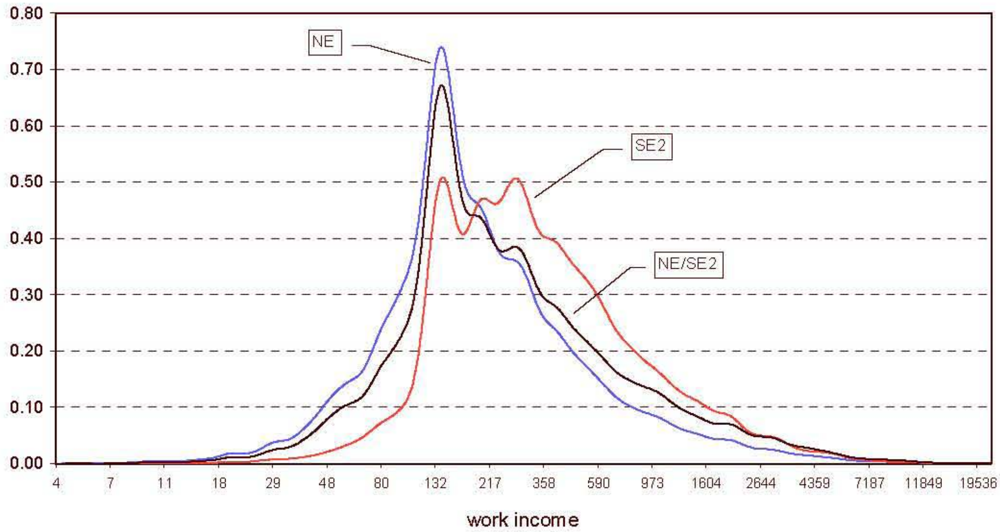
Graph 4
Real densities for SE1 and NE and counterfactual densities for NE with schooling of SE1



[12]All the money values are in Brazilian currency, the real (plural reais).

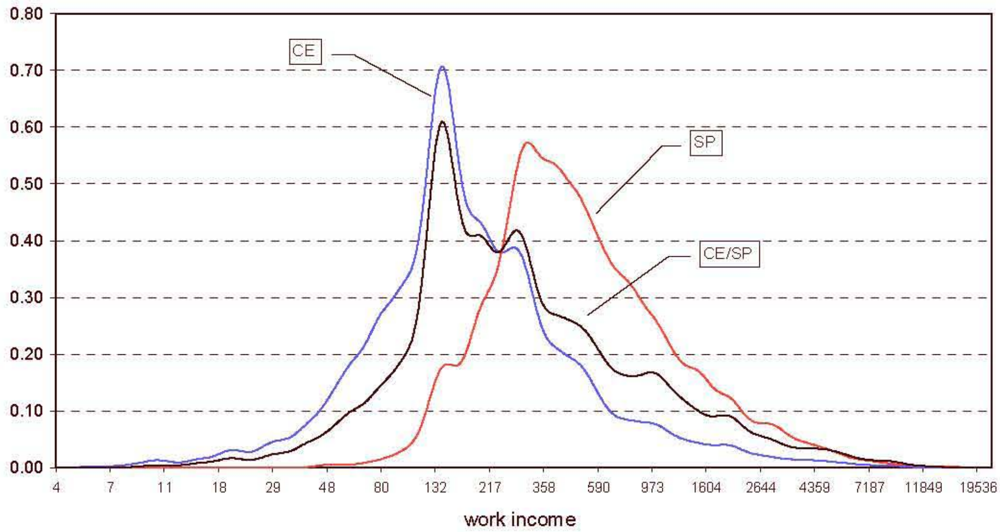
Graph 5

Real densities for SE2 and NE and counterfactual densities for NE with schooling of SE2



Graph 6

Real densities for SP and CE and counterfactual densities for CE with schooling of SP



We tested the equality between the original empirical distributions and their counterfactuals using the Kolmogorov-Smirnov test, in which the test statistic is the maximum of the difference between the accumulated densities: $\max |F_1(w) - F_2(w)|$. We rejected the null hypothesis that the densities are equal at 1% significance in all cases, showing that the change in the schooling profile alters the income distribution.

It is evident that the distributions[13] are sensitive to the differences in schooling and that the degree of sensitivity depends on the income percentile considered. The value of the minimum salary, R$ 136 at the time of the study, is the determining

[13] Read from the graphs:
NE/SE1: counterfactual of the Northeast using schooling of the Southeast (MG+RJ+SP).
NE/SE2: counterfactual of the Northeast using schooling of the Southeast (MF+RJ).
CE/SP: counterfactual of Ceará using the schooling of São Paulo.

factor in job income of workers in the Northeast, Ceará, and the Southeast (SE2) [SE1?], in view of the enormous concentration around this value in the distribution. The income distribution for the Southeast (SE2) is bimodal, with one of the modes corresponding to the minimum salary value and the other equal to the mode of the distribution for São Paulo, about R$ 300. The distributions for SP and SE1 are quite similar, which shows the significance of São Paulo on the aggregate of the Southeast Region. We can state, then, that in all states of the Northeast and Southeast, excepting São Paulo (SE2), the value of the minimum salary has a strong impact on the distribution of job income.

Comparing the tails of the counterfactual distributions with the original distributions, we see that they are quite close, except for a rightward translation in the former ones. This is due to the direct proportionality between schooling and income together with the reweighting of the sample, whose effects on the income distribution are similar to what would be obtained if we had added a schooling constant for each individual. One can observe that the central areas of the income distributions, which contain exactly the concentrations of most individuals and those whose income is near one minimum salary, change very little after reweighting. This behavior is explained by the fact that the wage structure is not altered in this process, overlapping [overpowering/overcoming?] the effect of the increase in schooling resulting from the same process. The mode, for example, remains equal to one minimum salary in all cases.

Table 5 presents some notable points (percentiles, means, standard deviations and percentile differences) of the estimated, real and counterfactual distributions. The support of [for?] the estimated distributions is the logarithm of income, which we limited for computational effects to the interval [1,10] with steps of 0.01. Based on these distributions, we constructed distributions, with the support being the level of income, simply taking the exponential of each point of the estimated distributions and renormalizing. We present in Tables ^{a} 1 and ^{a} 2 (in the appendix) the notable points of all the distributions constructed and the ratios of the percentiles of the estimated distributions.

Table 5 - Notable Points of the Estimated Income Distributions

	Estimated Distributions (*)								
	Brazil	CE	SP	NE	SE1	SE2	CE/SP	NE/SE1	NE/SE2
Percentile									
10	116	61	187	70	144	125	87	89	84
20	147	90	250	101	196	153	125	124	119
30	189	117	302	124	250	192	147	144	140
40	237	136	358	141	299	235	183	174	164
50	293	158	428	166	365	284	233	219	204
60	365	194	523	202	450	358	296	281	262
70	469	247	672	262	572	459	412	384	351
80	665	334	916	365	796	639	639	578	523
90	1,130	567	1,510	639	1,339	1,086	1,200	1,086	982
Mean	572	338	775	358	676	553	584	534	493
SD	905	697	1,056	704	977	855	1,091	687	929
Differences of Percentiles									
10-90	1,014	506	1,323	569	1,195	961	1,113	997	898
10-50	177	97	242	96	221	159	145	130	120
50-90	837	409	1,082	473	974	801	967	867	778
20-80	518	244	666	264	600	486	514	454	404

(*) The values presented were extracted from the estimated (in log), taking the exponential of each point.
The means and standard deviations were calculated directly from the reweighted sample with the income in level.

Table 6 - Income Ratio in the Percentiles

Percentile	CE x SP	CE/SP x SP	NE x SE1	NE x SE2	NE/SE1 x SE1	NE/SE2 x SE2
10	33%	47%	49%	56%	62%	67%
20	36%	50%	52%	66%	63%	78%
30	39%	49%	50%	64%	58%	73%
40	38%	51%	47%	60%	58%	70%
50	37%	54%	45%	58%	60%	72%
60	37%	57%	45%	57%	63%	73%
70	37%	61%	46%	57%	67%	76%
80	36%	70%	46%	57%	73%	82%
90	38%	79%	48%	59%	81%	90%

Comparing the original distributions of Ceará and São Paulo, one can see that the income of the former is approximately 1/3 that of the latter in all the percentiles (Table 6). On the other hand, in comparing the distribution of Ceará reweighted by the education of São Paulo with the original São Paulo distribution, we can see two effects: (i) there has been a gain in the income distribution of Ceará in all percentiles, and (ii) the gain increased as the income level (percentile) went up[14]. The growing marginal gain from schooling, illustrated in Graph 3, justifies the second effect, since individuals with more schooling, and consequently greater incomes, in receiving a " gain" in schooling by the reweighting effect, will have their incomes raised in a greater proportion than those with less schooling, whose income evolves little after an increase in schooling.

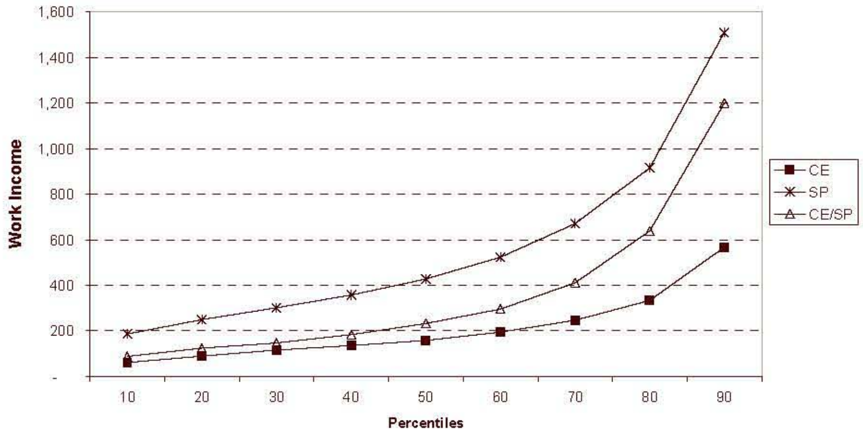
Comparing the Northeast Region with the Southeast, one can see that the first's income is between 45% and 50% of the second's for the SE1 sample, and between 55% and 65% for the SE2 sample. After the reweighting, the same effects can be seen as when we compared the distribution Ceará with that of São Paulo. The effect of the reweighting is directly proportional to the difference between the original distributions, so it is greater in the case of CE x SP and less for NE x SE2, for all the percentiles. An interesting comparison is that the average income of the Northeast reweighted by the education of the Southeast (SE1) is 93% of the average Brazilian income. Here, like in the CE X SP comparison, the fact that schooling presents a rising marginal return justifies a greater income gain in the higher deciles.

The absolute gains are significant in all cases: in no decile is the income gain with reweighting less that 13% of the original income, and it reaches 70% in the ninth decile for the CE x SP comparison. The most relevant is that the income of Northeasterners is never less that about 60% of that of Southeasterners for the SE1 sample and 67% for the SE2, and that in the upper deciles it comes quite near the income of Southeast residents if the two groups have the same level of education. The income in the ninth decile of the NE reweighted by the schooling of the SE2 region reaches 90% of the SE2 income.

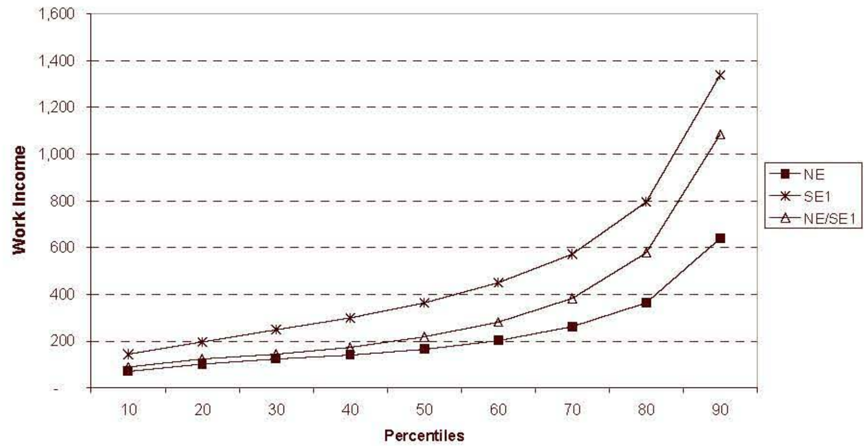
This fact is better represented in Graphs 7 to 9 (and A.4 to A.6 of the appendix), where the counterfactual and the income distribution of the region/state that was used as a base for the reweighting approach each other at the higher income levels.

[14]For the first decile, the income of Ceará went from 33% to 47% of that of São Paulo, and for the ninth decile it went from 38% to 79%.

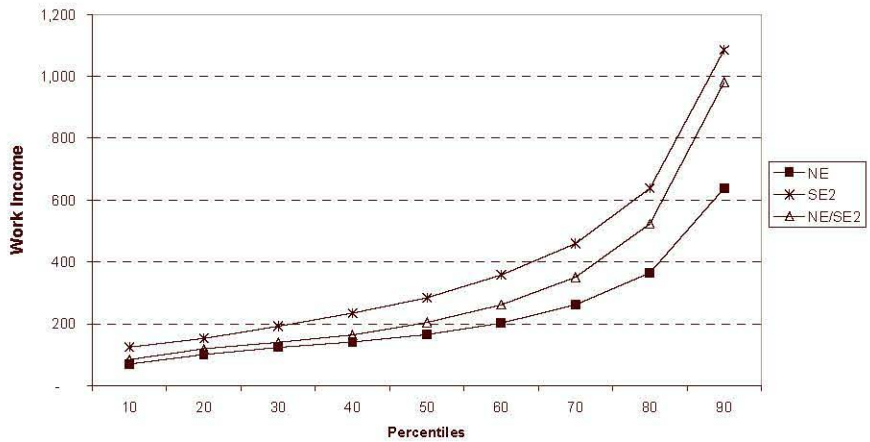
Graph 7 – Evolution of Income by Percentile
for CE, SP and CE/SP



Graph 8 – Evolution of Income by Percentile
for NE, SE1 and NE/SE1

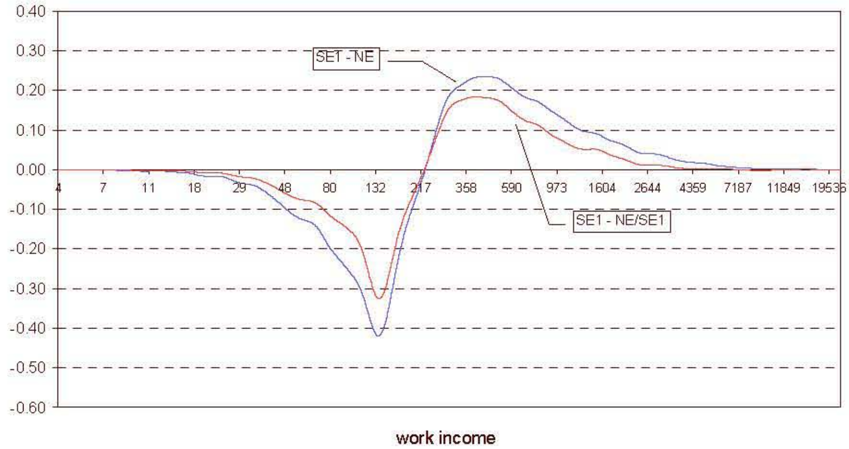


Graph 9 - Evolution of Income by Percentile
for NE, SE2 and NE/SE2

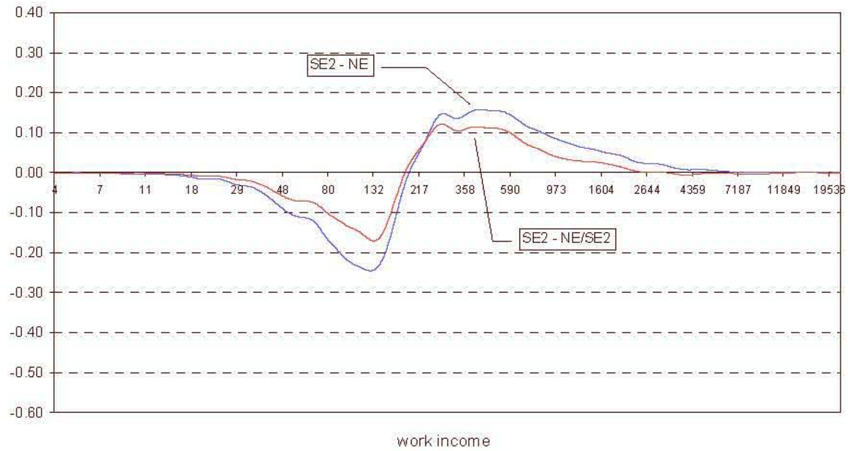


In Graphs 10 to 12 below, we present the difference between the two real distributions and the difference between the counterfactual distribution and the real one whence it came[15]. These graphs allow us to visualize how much the income distributions approach each other after the reweighting by schooling.

Graph 10 – Difference of the Distributions (SE1 x NE)

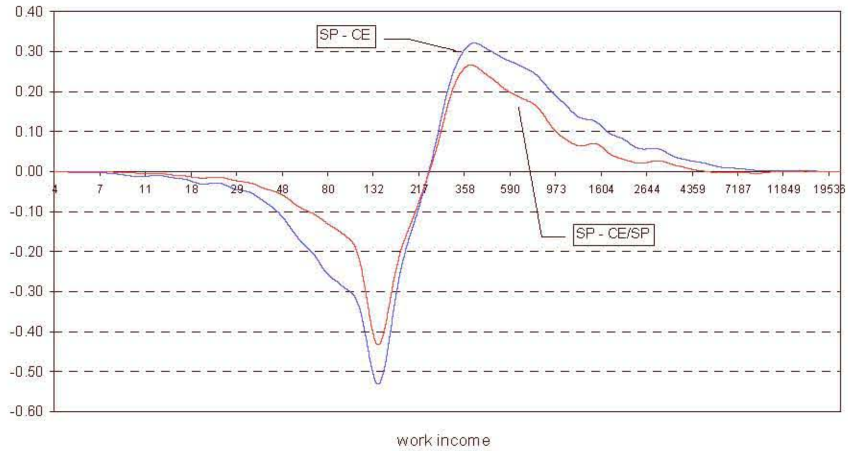


Graph 11 – Difference of the Distributions (SE2 x NE)



[15] For each level of income, we took the difference (counterfactual – real), i.e., f(w;.....).

Graph 12 – Difference of the Distributions (SP x CE)



We can see that the differences are negative at the lower income levels and positive for the higher ones, which denotes that the preponderant effect is the difference of the average between the distributions rather than the difference of the dispersion. If this latter effect had prevailed, we would have seen positive differences in the tails and negative ones in the center of the support of the distributions.

We computed the distance between the distributions, before and after the reweighting, by the parametric measure known as the Kullbach-Leibler distance (J)[16]. The results are shown in Table 7 below. Basically what this measurement does is represent the area between two distributions where each area element (corresponding to each w) is weighted by the difference in density between the distributions. The effect of the weighting is to introduce a non-linear relationship between the distance between the distributions, for each w, and its contribution to J. We find that more than 55% of the distances between Ceará and São Paulo and between the Northeast and Southeast regions (SE1 and SE2) are explained by schooling – 55.3% of the distance between São Paulo and Ceará and 55.0% (57.3%) of the distance between the Northeast and region SE1 (SE2). This shows that more than half the income difference between the poorest region/state and the richest is due to the schooling differences of the population. We also computed the distance between the distributions by the metrics of Chernoff[17] and Sibson[18], suggested in Krzanowski (2003), and obtained similar results, which are shown in Table 9.

[16] The Kullbach-Leibler distance, J, is a measure of the divergence between two distributions .....
[17] The Chernoff distance between the distributions f1 and f2 is defined as – log...
[18] The Sibson distance between the distributions f1 and f2 is defined as – $\frac{1}{2}$ f...., where f(x) ...

Table 7 – Kullback-Leibler Distances

Distance between the distributions								
	CE	NE	CE/SP	CE/SE1	CE/SE2	NE/SP	NE/SE1	NE/SE2
SP	1.5901	1.3826	0.7110			0.6700		
SE1	0.9608	0.8031		0.3969			0.3611	
SE2	0.5396	0.4127			0.2145			0.1761

% of the distance between the distributions explained by education			
SP - CE	55.3%	SP - NE	51.5%
SE1 - CE	58.7%	SE1 - NE	55.0%
SE2 - CE	60.3%	SE2 - NE	57.3%

We calculated the entropy coefficients of Theil and the concentration coefficients of Gini for the estimated distributions. These coefficients are measures of dispersion and are widely used as measures of inequality when applied to income distributions. The results are presented in Table 8 below (and in Table A.4 in the appendix). In sum, we observed an increase in income inequality when we reweighted the sample. This is due to the fact that the rise in income is not uniform along the entire distribution, there being greater gains for the higher income levels.

Table 8 – Gini and Theil Coefficients

		Gini Coefficient	Theil Coefficient
Brazil		0.551	0.607
Regions	CE	0.582	0.768
	NE	0.574	0.731
	SE1	0.523	0.539
	SE2	0.534	0.573
	SP	0.505	0.498
	CE/SP	0.622	0.797
	NE/SE1	0.608	0.766
	NE/SE2	0.604	0.768

Remark: SE1 = MG + RJ + SP

SE2 = MG + RJ

We measured the effect of the reweighting by schooling by computing the percentage variation for some metrics, calculated based on the estimated distributions. The results are reported in Table 9 below. In terms of average income, or considering the Kullbach-Leibler distance, schooling was able to explain more than 55% of the income difference, with this index reaching 70% when reweighting the Northeast by the schooling of region SE2. The reweighting caused two effects: (i) the measures of central tendency converged, and (ii) the measures of dispersion increased, signaling that the income inequality in the poorest regions would increase if they had the schooling profile of the richest, maintaining the same wage structure. In other words, the Northeast with the schooling profile of the Southeast would be richer, but more unequal. The effect referred to in item (ii) arises, in part, from the

fact that our exercise did not account for the change in the wage structure from the change in schooling, i.e., in general equilibrium this phenomenon would be observed partly or totally.

Table 9 - Percentage Explained by Schooling

Metric	SP x CE	SE1 x NE	SE2 x NE
Mean	56.3%	55.3%	69.2%
Median	27.8%	26.9%	32.6%
Theil Coefficient	-10.7%	-18.0%	-23.4%
Gini Coefficient	-51.9%	-66.9%	-75.0%
Kullbach-Leibler Distance	55.3%	55.0%	57.3%
Chernoff Distance	56.0%	55.5%	57.5%
Sibson Distance	52.5%	53.6%	56.5%

We investigated the robustness of the results to the choice of window size by re-estimating all the distributions with $h = 0.09$ (–25% of the original h) and $h = 0.15$ (+25%), and comparing the results against those obtained with the original value of h. We found that the changes in the results did not compromise them, even without observing a direct relation between the size variation of the window and the variation of the indicators presented in Tables 5 through 9.

It should be pointed out that the portion of the income difference explained by schooling is perhaps greater than shown in this work, since the cost of living is different in the regions. Services and nontradable goods are cheaper in the poorer regions, reflecting exactly the difference in wage levels, since these items tend to be more labor-intensive. The effect of correction by purchasing power in each region would be less in the lower income levels (consuming fewer services) and greater for individuals with higher incomes (consuming more services), accentuating the effect shown in Graphs 7 to 9.

5 Conclusion

In this work we sought to identify how much the income differential between the Northeast and Southeast regions and between the states of Ceará and São Paulo is explained by the differences in schooling levels of their populations. We used a semiparametric model to constrict counterfactual density functions, reweighting the individuals of the base region/state by the education distribution of the region of comparison. We estimated the distributions of real and counterfactual income of the state of Ceará and the Northeast Regions reweighted by the schooling levels of the Southeast Region and the state of São Paulo.

We found that: (i) the income dispersion is greater in the distributions with lower means, i.e., the income inequality, which is enormous in all regions, is greater in the poorer regions; (ii) more than 55% of the difference in income between the Northeast and Southeast regions and between the states of São Paulo and Ceará, when measured by the Kullbach-Leibler distance or in terms of average income, is due to the difference in schooling; (iii) the reweighting by schooling increased the average income of the Northeast by nearly 50% and by more than 70% for Ceará; (iv) the higher the income percentile considered, the greater was the contribution of the difference in schooling for the difference in income; and (v) the income dispersion in the poorest regions increased when we gave them the schooling levels of the richest regions, while maintaining the wage profile of the region.

There are various factors that may be determining the income difference not explained by the schooling differential, among them the life expectancy of the people, ethnic factors, the age structure of the population, quality of infrastructure, presence/absence of development stimuli and historical factors. A natural extension of this work would be, applying the method of counterfactuals, to decompose the income differential into some of these factors besides schooling, seeking to explain a greater portion of the income differential. The relative importance of the above factors against the cost of eliminating them could well be of great value in orienting public policies to combat " regional inequality" . Other interesting extensions would be: (i) reweighting the income distribution of the Northeast Region by the schooling of Brazil without the North Region[19]; and (ii) reweighting the income distribution of all regions by the schooling profile of Brazil as a whole, permitting determination of the size of regional inequalities controlling for schooling.

References

[1] Barros, Ricardo Paes de (1993). "Regional Disparities in Education Within Brazil: the Role of Quality of Education", Texto para Discussão nº 311, IPEA.

[2] Barros, Ricardo Paes de & Mendonça, Rosane (1995). "Os Determinantes da Desigualdade no Brasil", Texto para Discussão nº 377, IPEA.

[3] Barros, Ricardo Paes de, Camargo, José Márcio & Mendonça, Rosane (1997). "A Estrutura do Desemprego no Brasil", Texto para Discussão nº 478, IPEA.

[4] Blinder, Alan S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources*, 8, 436-455.

[5] Bourguignon, François, Ferreira, Francisco & Leite, Phillippe (2002). "Beyond Oaxaca-Blinder: Accounting for Differences in Household Income Distributions Across Countries", mimeo.

[6] DiNardo, John, Fortin, Nicole M. & Lemieux, Thomas (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach", *Econometrica*, Vol. 64, No. 5, 1001-1044.

[7] Butcher, Kristin F. & DiNardo, John (1998). "The Immigrant and Native-Born Wage Distributions: Evidence from United States Censuses", *NBER Working Paper Series 6630.*

[8] Krzanowski, W. J. (2003). "Non-parametric estimation of distance between groups", *Journal of Applied Statistics*, Vol. 30, No. 7, 743-750.

[9] Oaxaca, R. (1973). "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review*, 14, 693-709.

[10] Park, B. U. & Marron, J. S. (1990). "Comparision of Data-driven Bandwidth Selectors", *Journal of American Statistical Association,* 85, 66-72.

[11] Parzen, E. (1962). "On Estimation of a Probability Density Function and Mode", *The Annals of Mathematical Statistics*, 33, 1065-1076.

[12] Pessoa, Samuel (2000). "Existe um Problema de Desigualdade Regional no Brasil?", mimeo.

[19] This because the PNAD does not interview the rural population, which is very significant in the North Region.

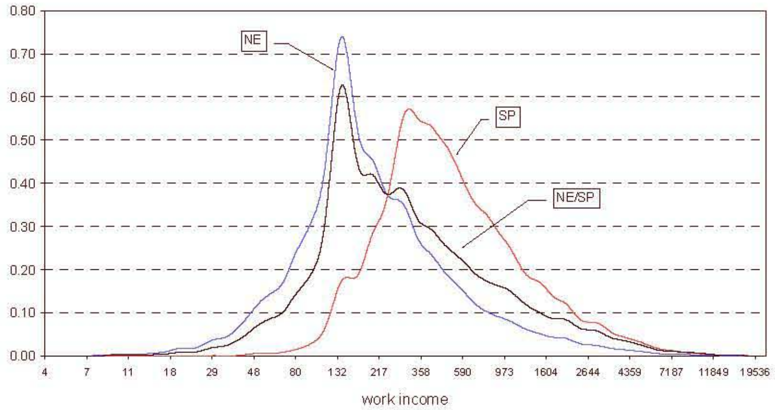
[13] Rosenblatt, M. (1956). "Remarks on Some Non-parametric Estimates of a Density Function", *The Annals of Mathematical Statistics*, 27, 832-837.

[14] Sheater, S. J. & Jones, M. C. (1991). "A Reliable Data-based Bandwidth Selection Method for Kernel Density Estimation", *Journal of Royal Statistical Society,* 53, 683-690.

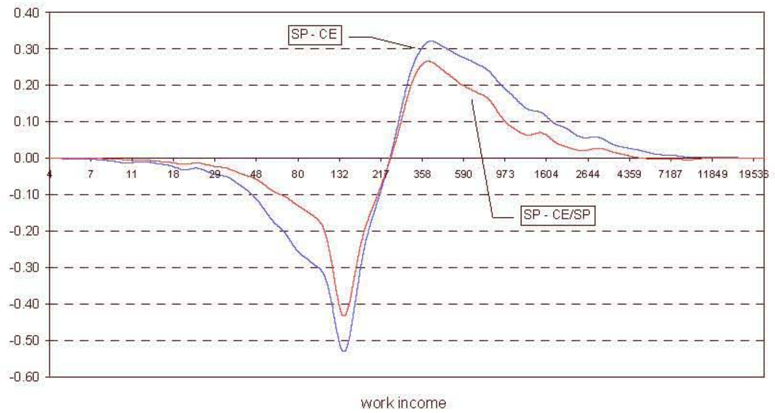
[15] Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall, 1986.

A Tables and Figures

Graph A.1
Real densities for SP and NE and counterfactual for NE with schooling of SP

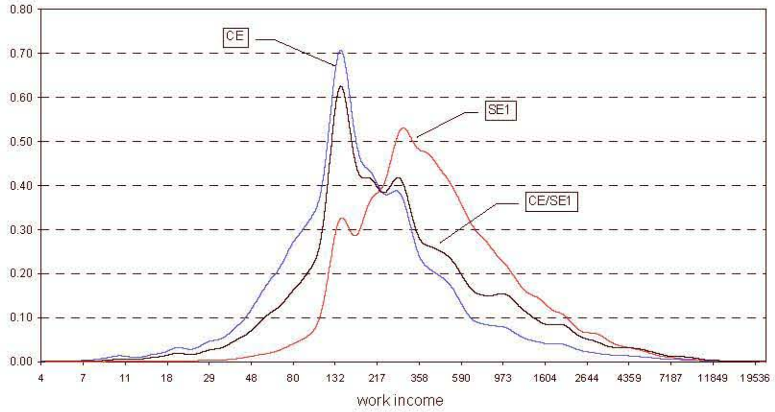


Differences between the distributions of SP and NE (real and counterfactual)

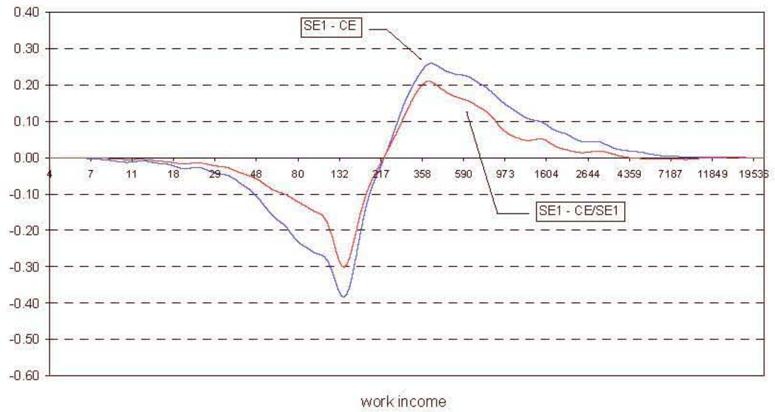


Graph A.2

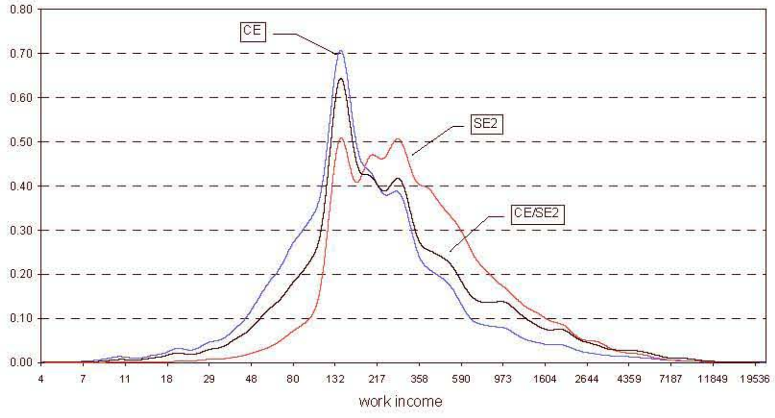
Real densities for SE1 and CE and counterfactual for CE with schooling of SE1



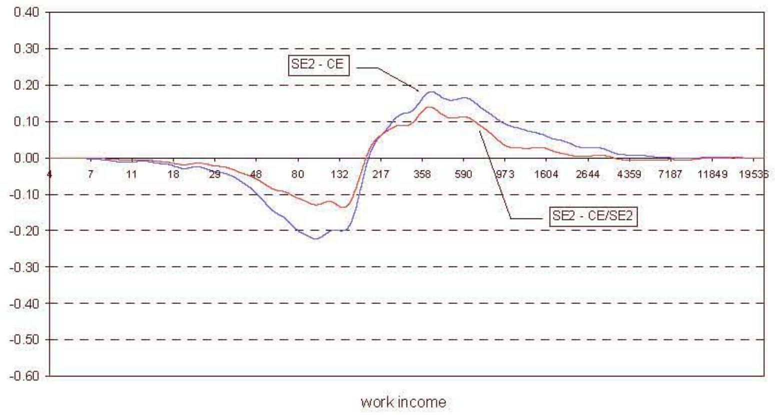
Differences between the distributions of SE1 and CE (real and counterfactual)



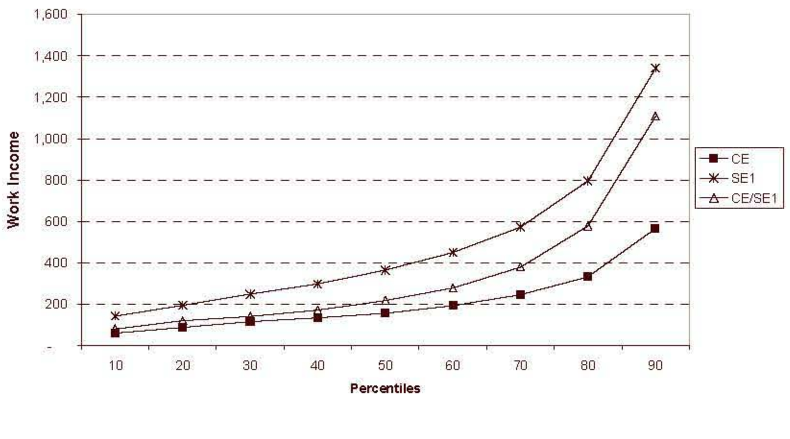
Graph A.3
Real densities for SE2 and CE and counterfactual for CE with schooling of SE2



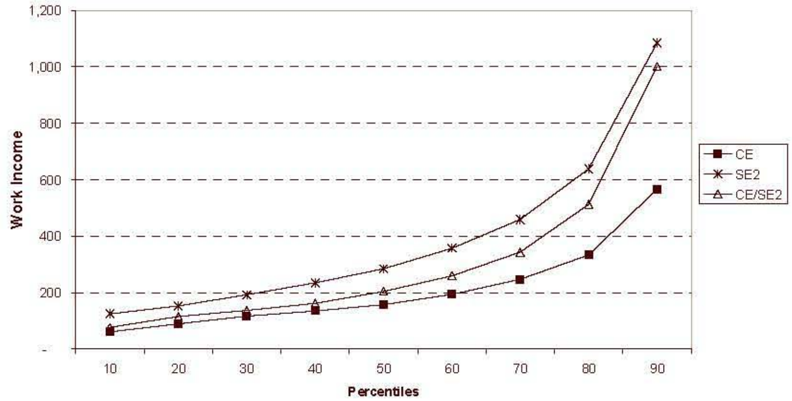
Differences between the distributions of SE2 and CE (real and counterfactual)



Graph A.4 - Evolution of Income by Percentile
for CE, SE1 and CE/SE1



Graph A.5 – Evolution of Income by Percentile
for CE, SE2 and CE/SE2



Graph A.6 - Evolution of Income by Percentile
for NE, SP and NE/SP

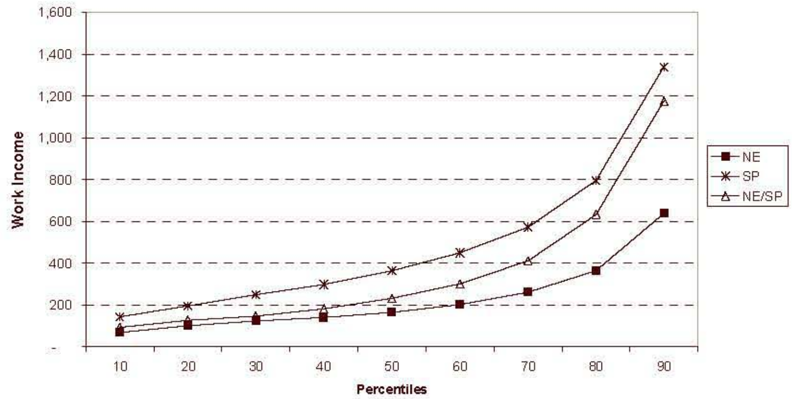


Table A.1 – Notable Points of the Estimated Income Distributions

	Estimated Distributions(*)											
	Brazil	CE	SP	NE	SE1	SE2	CE/SP	CE/SE1	CE/SE2	NE/SP	NE/SE1	NE
Percentile												
10	116	61	187	70	144	125	87	82	77	94	89	
20	147	90	250	101	196	153	125	122	116	128	124	
30	189	117	302	124	250	192	147	143	137	148	144	
40	237	136	358	141	299	235	183	172	162	183	174	
50	293	158	428	166	365	284	233	219	204	233	219	
60	365	194	523	202	450	358	296	279	260	302	281	
70	469	247	672	262	572	459	412	380	344	412	384	
80	665	334	916	365	796	639	639	578	513	633	578	
90	1,130	567	1,510	639	1,339	1,086	1,200	1,108	1,002	1,176	1,086	
Mean	572	338	775	358	676	553	584	547	503	569	534	
SD	905	697	1,056	704	977	855	1,091	1,041	978	1,033	687	
Differences of Percentiles												
10-90	1,014	506	1,323	569	1,195	961	1,113	1,025	926	1,082	997	
10-50	177	97	242	96	221	159	145	137	128	139	130	
50-90	837	409	1,082	473	974	801	967	888	798	943	867	
20-80	518	244	666	264	600	486	514	457	397	505	454	

(*) The values presented were extracted from the estimated (in log), taking the exponential of each point.
The means and standard deviations were calculated directly from the reweighted sample with the income in level.

Table A.2 - Income Ratio in the Percentiles

Percentile	CE x SP	CE/SP x SP	CE x SE1	CE x SE2	CE/SE1 x SE1	CE/SE2 x SE2
10	33%	47%	42%	49%	57%	61%
20	36%	50%	46%	59%	62%	76%
30	39%	49%	47%	61%	57%	71%
40	38%	51%	45%	58%	58%	69%
50	37%	54%	43%	55%	60%	72%
60	37%	57%	43%	54%	62%	73%
70	37%	61%	43%	54%	66%	75%
80	36%	70%	42%	52%	73%	80%
90	38%	79%	42%	52%	83%	92%

Percentile	NE x SP	NE/SP x SP	NE x SE1	NE x SE2	NE/SE1 x SE1	NE/SE2 x SE2
10	38%	50%	49%	56%	62%	67%
20	41%	51%	52%	66%	63%	78%
30	41%	49%	50%	64%	58%	73%
40	39%	51%	47%	60%	58%	70%
50	39%	54%	45%	58%	60%	72%
60	39%	58%	45%	57%	63%	73%
70	39%	61%	46%	57%	67%	76%
80	40%	69%	46%	57%	73%	82%
90	42%	78%	48%	59%	81%	90%

Table A.3 – Coefficients of Gini and Theil

			Gini Coefficient	Theil Coefficient
Brazil			0.551	0.607
Regions	Original Weighting	CE	0.582	0.768
		NE	0.574	0.731
		SE1	0.523	0.539
		SE2	0.534	0.573
		SP	0.505	0.498
	Reweighting by Education	CE/SP	0.622	0.797
		CE/SE1	0.621	0.805
		CE/SE2	0.618	0.810
		NE/SP	0.609	0.760
		NE/SE1	0.608	0.766
		NE/SE2	0.604	0.768

Remark: SE1 = MG + RJ + SP

SE2 = MG + RJ

Table A.4 – Percentage Explained by Schooling

Metric	SP x CE	SE1 x CE	SE2 x CE	SP x NE	SE1 x NE	SE2 x NE
Mean	56.3%	61.8%	76.7%	50.6%	55.3%	69.2%
Median	27.8%	29.7%	36.9%	25.5%	26.9%	32.6%
Theil Coefficient	-10.7%	-16.2%	-21.5%	-12.3%	-18.0%	-23.4%
Gini Coefficient	-51.9%	-66.8%	-75.0%	-51.4%	-66.9%	-75.0%
Kullbach-Leibler Distance	55.3%	58.7%	60.3%	51.5%	55.0%	57.3%
Chernoff Distance	56.0%	59.1%	60.4%	52.3%	55.5%	57.5%
Sibson Distance	52.5%	56.9%	59.2%	49.2%	53.6%	56.5%

ENSAIOS ECONÔMICOS DA EPGE

482. RETORNOS ANORMAIS E ESTRATÉGIAS CONTRÁRIAS - Marco Antonio Bonomo; Ivana Dall'Agnol – Junho de 2003 – 27 págs.

483. EVOLUÇÃO DA PRODUTIVIDADE TOTAL DOS FATORES NA ECONOMIA BRASILEIRA: UMA ANÁLISE COMPARATIVA - Victor Gomes; Samuel de Abreu Pessoa;Fernando A . Veloso – Junho de 2003 – 45 págs.

484. MIGRAÇÃO, SELEÇÃO E DIFERENÇAS REGIONAIS DE RENDA NO BRASIL - Enestor da Rosa dos Santos Junior; Naércio Menezes Filho; Pedro Cavalcanti Ferreira – Junho de 2003 – 23 págs.

485. THE RISK PREMIUM ON BRAZILIAN GOVERNMENT DEBT, 1996-2002 - André Soares Loureiro; Fernando de Holanda Barbosa - Junho de 2003 – 16 págs.

486. FORECASTING ELECTRICITY DEMAND USING GENERALIZED LONG MEMORY - Lacir Jorge Soares; Leonardo Rocha Souza – Junho de 2003 – 22 págs.

487. USING IRREGULARLY SPACED RETURNS TO ESTIMATE MULTI-FACTOR MODELS: APPLICATION TO BRAZILIAN EQUITY DATA - Álvaro Veiga; Leonardo Rocha Souza – Junho de 2003 – 26 págs.

488. BOUNDS FOR THE PROBABILITY DISTRIBUTION FUNCTION OF THE LINEAR ACD PROCESS – Marcelo Fernandes – Julho de 2003 – 10 págs.

489. CONVEX COMBINATIONS OF LONG MEMORY ESTIMATES FROM DIFFERENT SAMPLING RATES - Leonardo R. Souza; Jeremy Smith; Reinaldo C. Souza – Julho de 2003 – 20 págs.

490. IDADE, INCAPACIDADE E A INFLAÇÃO DO NÚMERO DE PESSOAS COM DEFICIÊNCIA - Marcelo Neri ; Wagner Soares – Julho de 2003 – 54 págs.

491. FORECASTING ELECTRICITY LOAD DEMAND: ANALYSIS OF THE 2001 RATIONING PERIOD IN BRAZIL - Leonardo Rocha Souza; Lacir Jorge Soares – Julho de 2003 – 27 págs.

492. THE MISSING LINK: USING THE NBER RECESSION INDICATOR TO CONSTRUCT COINCIDENT AND LEADING INDICES OF ECONOMIC ACTIVITY - JoãoVictor Issler; Farshid Vahid – Agosto de 2003 – 26 págs.

493. REAL EXCHANGE RATE MISALIGNMENTS - Maria Cristina T. Terra; Frederico Estrella Carneiro Valladares – Agosto de 2003 – 26 págs.

494. ELASTICITY OF SUBSTITUTION BETWEEN CAPITAL AND LABOR: A PANEL DATA APPROACH - Samuel de Abreu Pessoa ; Silvia Matos Pessoa; Rafael Rob – Agosto de 2003 – 30 págs.

495. A EXPERIÊNCIA DE CRESCIMENTO DAS ECONOMIAS DE MERCADO NOS ÚLTIMOS 40 ANOS – Samuel de Abreu Pessoa – Agosto de 2003 – 22 págs.

496. NORMALITY UNDER UNCERTAINTY – Carlos Eugênio E. da Costa – Setembro de 2003 – 08 págs.

497. RISK SHARING AND THE HOUSEHOLD COLLECTIVE MODEL - Carlos Eugênio E. da Costa – Setembro de 2003 – 15 págs.

498. REDISTRIBUTION WITH UNOBSERVED 'EX-ANTE' CHOICES - Carlos Eugênio E. da Costa – Setembro de 2003 – 30 págs.

499. OPTIMAL TAXATION WITH GRADUAL LEARNING OF TYPES - Carlos Eugênio E. da Costa – Setembro de 2003 – 26 págs.

500. AVALIANDO PESQUISADORES E DEPARTAMENTOS DE ECONOMIA NO BRASIL A PARTIR DE CITAÇÕES INTERNACIONAIS - João Victor Issler; Rachel Couto Ferreira – Setembro de 2003 – 29 págs.

501. A FAMILY OF AUTOREGRESSIVE CONDITIONAL DURATION MODELS - Marcelo Fernandes; Joachim Grammig – Setembro de 2003 – 37 págs.

502. NONPARAMETRIC SPECIFICATION TESTS FOR CONDITIONAL DURATION MODELS - Marcelo Fernandes; Joachim Grammig – Setembro de 2003 – 42 págs.

503. A NOTE ON CHAMBERS'S "LONG MEMORY AND AGGREGATION IN MACROECONOMIC TIME SERIES" – Leonardo Rocha Souza – Setembro de 2003 – 11págs.

504. ON CHOICE OF TECHNIQUE IN THE ROBINSON-SOLOW-SRINIVASAN MODEL - M. Ali Khan – Setembro de 2003 – 34 págs.

505. ENDOGENOUS TIME-DEPENDENT RULES AND THE COSTS OF DISINFLATION WITH IMPERFECT CREDIBILITY - Marco Bonomo; Carlos Viana de Carvalho – Outubro de 2003 – 27 págs.

506. CAPITAIS INTERNACIONAIS: COMPLEMENTARES OU SUBSTITUTOS? - Carlos Hamilton V. Araújo; Renato G. Flôres Jr. – Outubro de 2003 – 24 págs.

507. TESTING PRODUCTION FUNCTIONS USED IN EMPIRICAL GROWTH STUDIES - Pedro Cavalcanti Ferreira; João Victor Issler; Samuel de Abreu Pessoa – Outubro de 2003 – 8 págs.

508. SHOULD EDUCATIONAL POLICIES BE REGRESSIVE ? Daniel Gottlieb; Humberto Moreira – Outubro de 2003 – 25 págs.

509. TRADE AND CO-OPERATION IN THE EU-MERCOSUL FREE TRADE AGREEMENT - Renato G. Flôres Jr. – Outubro de 2003 – 33 págs.

510. OUTPUT CONVERGENCE IN MERCOSUR: MULTIVARIATE TIME SERIES EVIDENCE - Mariam Camarero; Renato G. Flôres Jr; Cecílio Tamarit – Outubro de 2003 – 36 págs.

511. ENDOGENOUS COLLATERAL - Aloísio Araújo; José Fajardo Barbachan; Mario R. Páscoa – Novembro de 2003 – 37 págs.

512. NON-MONOTONE INSURANCE CONTRACTS AND THEIR EMPIRICAL CONSEQUENCES - Aloísio Araujo; Humberto Moreira – Novembro de 2003 – 31 págs.

513. EQUILIBRIA IN SECURITY MARKETS WITH A CONTINUUM OF AGENTS - A. Araujo; V. F. Martins da Rocha; P. K. Monteiro – Novembro de 2003 – 17 pág's.

514. SPECULATIVE ATTACKS ON DEBTS AND OPTIMUM CURRENCY AREA: A WELFARE ANALYSIS - Aloisio Araujo; Márcia Leon – Novembro de 2003 – 50 págs.

515. O MÉTODO GENERALIZADO DOS MOMENTOS(MGM): CONCEITOS BÁSICOS - Renato G. Flôres Jr – Novembro de 2003 – 27 págs.

516. VARIÁVEIS INTRUMENTAIS E O MGM: USO DE MOMENTOS CONDICIONAIS - Renato G. Flôres Jr – Novembro de 2003 – 27 págs.

517. O VALOR DA MOEDA E A TEORIA DOS PREÇOS DOS ATIVOS - Fernando de Holanda Barbosa – Dezembro de 2003 – 17 págs.

518. EMPRESÁRIOS NANICOS, GARANTIAS E ACESSO À CRÉDITO - Marcelo Côrtes Néri; Fabiano da Silva Giovanini - Dezembro de 2003 – 23 págs.

519. DESENHO DE UM SISTEMA DE METAS SOCIAIS - Marcelo Côrtes Néri; Marcelo Xerez - Dezembro de 2003 – 24 págs.

520. A NEW INCIDENCE ANALYSIS OF BRAZILIAN SOCIAL POLICIES USING MULTIPLE DATA SOURCES - Marcelo Côrtes Néri - Dezembro de 2003 – 55 págs.

521. AN INTRA-HOUSEHOLD APPROACH TO THE WELFARE COSTS OF INFLATION - Rubens Penha Cysne – Janeiro de 2004 – 16 págs.

522. CENTRAL LIMIT THEOREM FOR ASYMMETRIC KERNEL FUNCTIONALS - Marcelo Fernandes; Paulo Klinger Monteiro – Fevereiro de 2004 – 23 págs.

523. THE TRADE-OFF BETWEEN INCENTIVES AND ENDOGENOUS RISK - Aloísio Araujo; Humberto Moreira; Marcos H. Tsuchida – Fevereiro de 2004 – 21 págs.

524. DO DIVIDENDS SIGNAL MORE EARNINGS ? - Aloísio Araujo; Humberto Moreira; Marcos H. Tsuchida – Fevereiro de 2004 – 26 págs.

525. Biased managers, organizational design, and incentive provision - Cristiano M. Costa; Daniel Ferreira; Humberto Moreira – Fevereiro de 2004 – 11 págs.

526. Land taxes in a Latin American context - Juliano J. Assunção; Humberto Moreira – Fevereiro de 2004 - 19 págs.

527. Indicadores coincidentes de atividade econômica e uma cronologia de recessões para o Brasil - Angelo J. Mont'alverne Duarte; João Victor Issler; Andrei Spacov - Fevereiro de 2004 – 41 págs.

528. TESTING UNIT ROOT BASED ON PARTIALLY ADAPTIVE ESTIMATION - Zhijie Xiao; Luiz Renato Lima – Março de 2004 – 27 págs.

529. DO SHOCKS PERMANENTLY CHANGE OUTPUT? LOCAL PERSISTENCY IN ECONOMIC TIME SERIES - Luiz Renato Lima; Zhijie Xiao – Março de 2004 – 21 págs.

530. A NEW PERSPECTIVE ON THE PPP HYPOTHESIS - Soyoung Kim; Luiz Renato Lima – Março de 2004 – 36 págs.

531. TRADE LIBERALIZATION AND INDUSTRIAL CONCENTRATION: EVIDENCE FROM BRAZIL - Pedro Cavalcanti Ferreira; Giovanni Facchini – Março de 2004 - 25 págs.

532. REGIONAL OR EDUCATIONAL DISPARITIES? A COUNTERFACTUAL EXERCISE - Angelo José Mont'Alverne; Pedro Cavalcanti Ferreira; Márcio Antônio Salvato – Março de 2004 – 25 págs.