Trends in the Relative Distribution of Wages by Gender and Cohorts in Latin America

Ana Maria Hermeto
Affiliation: CEDEPLAR/UFMG
Address: Av. Antônio Carlos, 6627 – FACE/UFMG – Belo Horizonte, MG, Brazil – 31270-901
Phone Number: 55-31-34097157
E-mail: ahermeto@cedeplar.ufmg.br

Abstract

Most of previous works regarding gender wage differentials in Latin America are based on parametric methodologies for estimation of wage differentials by gender, using traditional mincerian regressions or quantile regression estimates, with classical or bayesian statistics. The aim of our paper is to analyze and decompose changes in earnings relative distribution between men and women in different cohorts, using the relative distribution framework. This methodology considers non-parametrical tools which allow an exploratory analysis that is independent of parametric assumptions on the mathematical form of the response-variable probabilities. We use density estimates of the kernel probability for each sex and cohort and decompositions of the relative distribution to get substantive evidences for gender differentials and relative mobility in Latin American countries (Brazil, Argentina, Uruguay, Mexico and Chile), from the 1980s to 2010s. We use microdata from the National Household Sample Surveys to analyze the wage differentials between male and female workers. Our sample was restricted to those working at the survey reference week and earned a positive wage in the period. We constructed a pseudo-panel of these repeated cross-sections, which allows us to follow cohorts over time.

Introduction

A high gender wage inequality in Latin America persists over time. Literature is replete of evidences and discussions about this trend (for a review, Coelho and Corseuil, 2002; Henriques, 2000). Part of this literature is concerned with group differences and gender comparisons, disregarding the wider trend of wages stagnation and the polarization of wages. Whenever the attribute of interest is continuous, as, for example, these gender wage comparisons, often they are condensed in terms of medians or means. However, this usual parametric analysis of location and shifts provides a restricted framework for comparisons. That is, we have a very good notion of how individuals are allocated to positions at the wage distribution, but we do not know the structure of these positions. If the structure was stable, the restricted focus on allocation would be justifiable. Nevertheless, last decades in Brazil constitute a period of economic restructuring in several dimensions, and the impact on wage distributions between and within groups was huge. The structure of the wage inequality and the median trends are not the same for the sex groups.

Examining the specific Latin American wage distributions in last decades, we verify that real median wages fluctuate a lot and the wage variance persists in a high level. A set of important questions is hidden behind these summary statistics: (1) the upper and lower tails of the wage distribution increased at the same rate? (2) Are there further facts in the narrowing of the gender wage gap then the convergence of median wages between the groups? Information to answer these questions is available in data, but is inaccessible using standard statistical methods, like linear regressions, Gini and Theil indexes, etc.

Non-parametric approaches brought into consideration the fact that it was not strictly necessary to adopt assumptions about the mathematical form of the probabilities distribution of a variable. Most of parametric models (classical regressions and their decompositions) are sensitive to violations of their hypothesis, turning into misleading answers to studies questions (DINARDO e TOBIAS, 2001). Moreover, the non-parametric methodology allows the analysis of data as they are, without any prior distribution assumption.

Notions of relative distribution (Handcock e Morris, 1999) represent a non-parametric statistical framework for the comparative analysis of distributions differences and shifts, and can be used as a basis for exploratory, descriptive and analytical techniques. This framework combines graphical tools of exploratory data analysis to summary statistical measures, decomposition and inference. Relative distribution is similar to the density ratio, being defined as the random variable obtained by the transformation of a comparison group variable to the cumulative distribution function (CDF) for a reference group. This transformation yields a set of observations, the relative data, which represents the scale of the original comparison value in terms of the reference group CDF. Then, the density and the CDF of the relative data are used to fully describe and analyze distributive differences. In this sense, analysis progresses beyond means and variances, making use of detailed information inherent to distributions. Summary measures based on relative density are used to test hypothesis about differences in distributions. The analytical framework is general and flexible enough, as the relative density is decomposed in effects of differences in shape and location, and in effects that represents compositional changes in covariates and changes in the relation between the covariates and the dependent variable. Decomposition methods allow separating location, structure and compositional effects; then distinguishing the impact of changes of the population composition (a demographic process) from changes in the attribute location (a social or economic process).

Inequality is a good application of relative distribution methods because is a distribution property, rather than an individual property. Thus, it would be expected that statistical methods used to analyze inequality focused on the distributive analysis. However, in general, they do not; traditional statistical methods used in social sciences are based on linear models and their extensions. They are not designed to represent the details of data distributive patterns. On the contrary, the model the conditional mean, assuming the residual variance as homogeneous, treated as a noise parameter. Consequently, these methods do not take into account the larger share of the distributive information on data. The Lorenz curve and the Gini index, which represents distributive patterns associated to inequality, are a special case of relative distribution methods.

Methods of relative distribution aim at connecting exploratory tools and parametric restrictions to put the comparative distribution analysis in a robust statistical basis. The general non-parametric framework is based on the definition of a "relative distribution", a transformation of two distributions data in one distribution that contains all the required information for the scale independent comparison. Intuitively, the relative distribution is a transformation of data from two distributions (reference and comparison – for example, men and women, or two cohorts) into one distribution that contains all information for comparison between them. The relative distribution is the set of percentile positions that the observations of one distribution would have if they were located in another distribution. For example, the set of positions women would have if they were located in the male wage distribution. Therefore, the relative distribution method, proposed by Handcock and Morris (1999), constitutes a valuable instrument for the substantive analysis of the wage inequality and provides a consistent framework for data analysis.

Despite the development of data analysis based on the relative distribution and other non-parametric methods, there are few studies in Brazil that examine the evolution of gender wage inequality using these tools, particularly following cohorts over time. Besides, although inequality studies should focus on the distribution as a whole, most of empirical studies are stuck to methodologies based on mean values, which are not always representative of the population. The relative distribution method fills this lack and provides a framework of summary measures and figures that allows a substantive examination of data. Hence, the main objective of this article is to apply relative distribution definitions to the study of gender inequality in Latin America, contributing to the understanding of unexplored aspects and to the empirical debate.

This is a preliminary version of the paper, yet only with the application for Brazil. This application will be replicated to other countries in a comparative analysis.

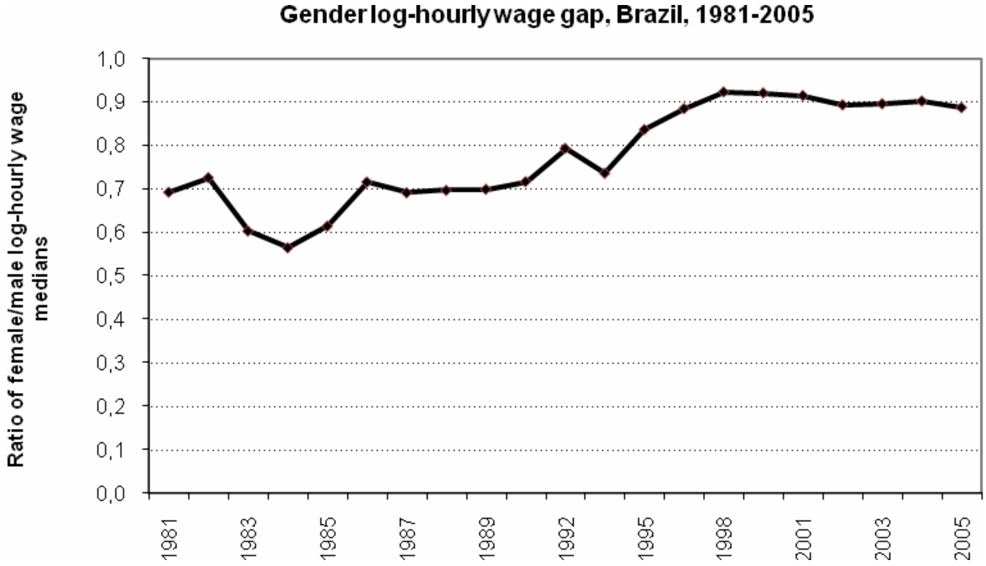
Overview

As mentioned above, the relative distribution is a statistical tool that fully represents differences between distributions, yielding a comprehensive framework that contains a graphical component, that simplifies the data exploratory analysis, a statistically basis for the development of summary measures which rely on hypothesis, and the potential for decomposition, that allows the discussion of complex hypothesis about origins of the distributive changes between and within groups. The integration of these analytical components in the context of the full distributive information reveals complex patterns and data relations, converting the relative distribution approach appropriate to the inequality issue in Brazil.

The gender wage gap in Brazil is a good example of the limitation of traditional summary measures, which typically focus on statistics which summarize the differential location of male and female wages, such as the median wages ratio (Figure 1). A different picture emerges if the complete female distribution in relation to the male is analyzed. This relative distribution is shown as relative deciles (Figure 2), which is essentially a rescaled density ratio: the probability ratio of women to men to be located in each level of the wage scale. To the log-hourly wage of each working woman is assigned the position she would have in the male distribution in the period, and these positions are plotted in a histogram. The cutting points of histogram classes are defined by the deciles of the male wage distribution, in such a way the frequency of each class represents the proportion of women located at each decile of the male wage scale over time. If the female and male wage distribution were the same, the relative deciles would assume the uniform value of 10% throughout the wage scale, provided that 10% of the female contingent would be at each male decile.

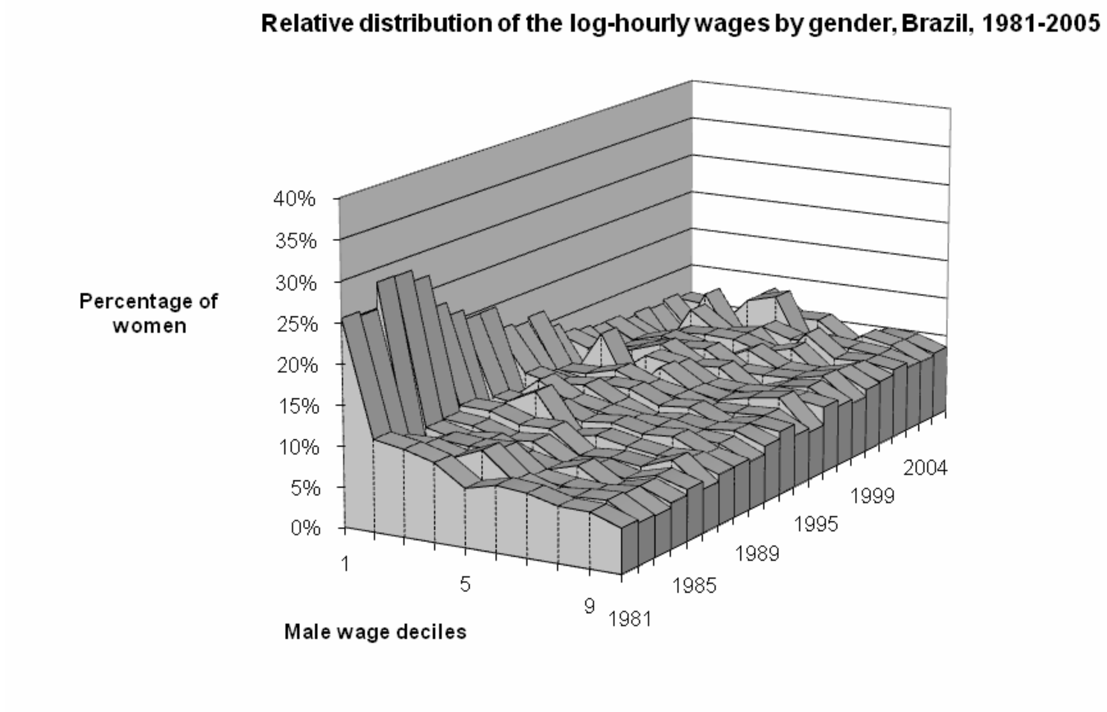
Figure 2 shows that in Brazil, the relative distribution is not uniform over time: a large share of the female distribution is concentrated on the lower tail of the male distribution, althought it changes over time. In 1981, 30% of all women were in the lowest decile of the male distribution and almost 70% (cumulative sum of 1-5 deciles) earned less than the median working man. In 2005, it has changed, though more than 60% still earned less than the median man. The absence of women in the upper tail of the male wage distribution also changed: 4% of women were there in 1981 e 9% twenty-five years later. This converging trend shows up in the middle of the 1990´s.

Figure 1



Source: PNAD microdata, 1981-2005. IBGE.

Figure 2



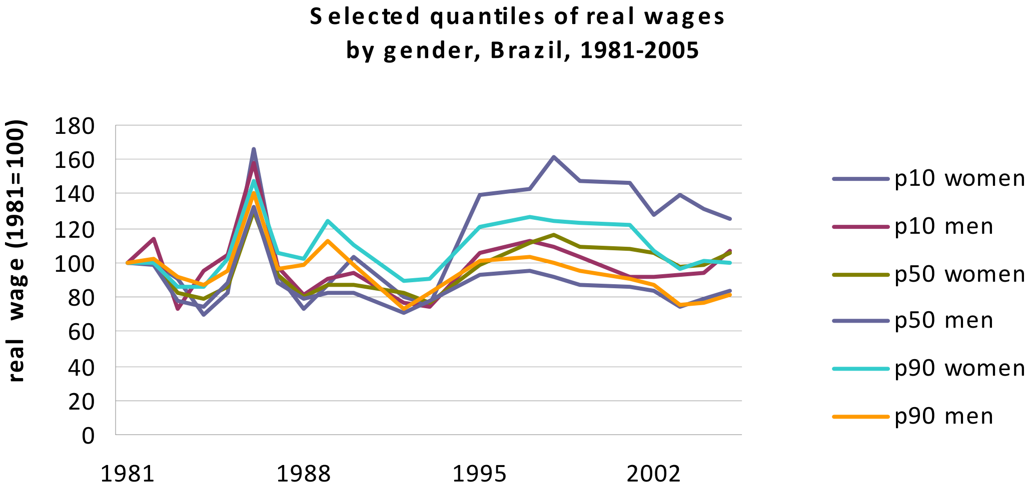
Source: PNAD microdata, 1981-2005. IBGE.

While the median ratio plotted at Figure 1 suggests that women improved over the decades, the relative distribution shows that this improvement has been more intense to women in the lower tail of the wage distribution: 70% of the total change in the relative density occurred below the median of male wages, 50% only in the lowest decile. If the skills upgrading was to explain this, it was not referring to the upper end of the scale, but to those at the bottom. Thus, median wages trends at figure 1 provide an incomplete picture of changes of male and female wages, occulting key aspects of the trend, inducing to misleading interpretations and a research agenda in the wrong extreme of the wage scale.

Patterns revealed at figure 2 are more informative about the key aspects of the changes. At the same time, this figure is harder to be interpreted because represents the combined outcomes of several factors: a median wage gap between the groups, changes in this gap over time and changes in the shape of the male and female wage distribution. The relative distribution can be decomposed in shares that account for each of these effects: the decomposition clears that the female gains in the bottom of the distribution are due to increasing female wages, but also to decreasing male wages; and that, in the other quantiles, the decline of male wages is still more relevant.

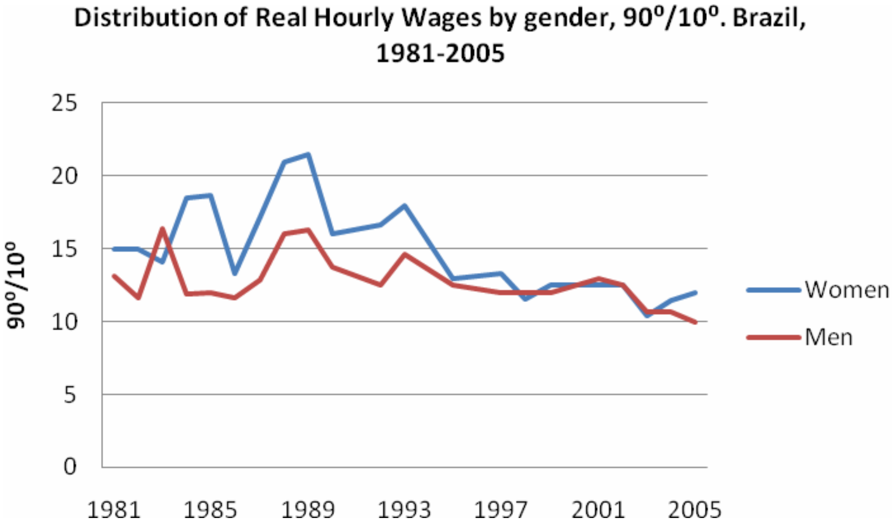
Figure 3 shows trends over time of real wages for the 10th, 50th e 90th percentiles. In this case, we divided values by the percentile value at the baseline year, such as values above 100 imply that individuals in this percentile had wage gains in this year, while values below 100 indicate a decline in the real wage for that percentile. In this sense, beyond conjuncture shifts common to all groups, we can find out declines in wages for men in all deciles and raises for women, remarkably for those in the lowest decile, and particularly in the 1990s. These trends evidence a great difference in male and female achievements. Comparing patterns for men and women, it is important to notice that these trends were constructed within groups, so female gains are not enough to equalize their distribution to the male one, and the gender gap persists, although women had relative and absolute advances in last decades. Another common aspect for men and women in this period was the growth of inequality within groups in the 1980s and the first half of the 1990s and afterwards a decline (figure 4).

Figure 3



Source: PNAD microdata, 1981-2005. IBGE.

Figure 4



Source: PNAD microdata, 1981-2005. IBGE.

A crucial issue for group comparisons is how they are ranked. Supposing that female and male wages are compared in 1981 and again in 2005, the groups are more equal in 2005 than they were in 1981? An approach is to compute an inequality measure within groups, such as the Gini index, and to compare the four resulting measures (one for each distribution by sex and year). A more concise approach is to begin with a measure that directly captures the gender comparison and then compares the changes in this measure over time. This approach is adopted by the relative distribution methods, which perform the same role for comparisons between groups as the Lorenz curves for comparisons within groups.

Data and Methods

We use microdata from the Brazilian Household Sample Survey (*Pesquisa Nacional por Amostra de Domicílios* - PNAD), from 1981 to 2005. This survey is yearly conducted by the Brazilian Census Bureau (*Instituto Brasileiro de Geografia e Estatística – IBGE*) and presents a comprehensive data source, in particular about the labor market and earnings, statistically representative of Brazil. To analyze the wage differentials between male and female workers, our sample was restricted to those working at the survey reference week and earned a positive wage in the period. We constructed a pseudo-panel of these repeated cross-sections, which allows us to follow cohorts over time.

To access the gender wage inequality incorporating cohorts in Brazil, we estimated classic mincerian regressions in three points (1981, 1993 e 2005), considering only men and we obtained the returns to each male worker attribute. Using these coefficients, we estimated a contra factual predicted wage for women: wages that women would earn if they had the estimated male returns for their own attributes structure. Formally:

$$\log hwage_{i,men,t} = \beta_0 + \beta_1 y\_schooling_{i,men} + \beta_2 age_{i,men,t} + \beta_3 age_{i,men,t}{}^2 + \sum_{j=1}^{4} \beta_j region_{i,j,men,t} \quad (1)$$
$$+ \beta_8 metrop_{i,men,t} + \beta_9 urb_{i,men,t} + \varepsilon_{i,men,t}$$

$$\log hwage\_predicted_{i,women,t} = \hat{\beta}_0 + \hat{\beta}_1 y\_schooling_{i,women,t} + \hat{\beta}_2 age_{i,women,t} + \hat{\beta}_3 age_{i,women,t}^{2}$$
$$+ \sum_{j=1}^{4} \hat{\beta}_j region_{i,j,women,t} + \hat{\beta}_8 metrop_{i,women,t} + \hat{\beta}_9 urb_{i,women,t} \qquad (2)$$

Where: *loghwage* is the logarithm of the hourly-wage, *y_schooling* is the years of schooling; *region* are indicator variables for Brazilian macro-regions (southeast, center-west, south and northeast), *metrop* is an indicator variable for metropolitan areas, *urb* is an indicator variable for urban areas and *loghwage_predicted* is the predicted wage for women using estimated male returns. Then we constructed our response variable of interest, the *gender wage gap*, which represents the gap between observed and contra factual standardized female wages, yielding a measure of the gender differential (or discriminatory component):

$$gender\ wage\ \mathrm{gap}_{i,woman,t} = \log hwage_{i,woman,t} - \log hwage\_predicted_{i,woman,t} \quad (3)$$

The analysis of the gender wage gap distribution followed two cohorts in different stages of their life cycles: individuals of 25-36 years in 1981 (cohort 1) and those in the same age group in 1993 (cohort 2). Cohort age groups were defined to split individuals in ages of different wage mobility patterns. Table 1 ilustrates our pseudo-panel.

Table 1: Scheme of analysis by cohorts of women of the gender wage differential

Age/Period	1981	1993	2005
25-36	Cohort 1	Cohort 2	
37-48		Cohort 1	Cohort 2

Our variable of interest for the computation of the relative distribution is the gender wage gap, for the defined cohorts in two points in time: members of cohort 1 of 25-36 years of age in 1981 and 37-48 years in 1993, and members of cohort 2 in the same age groups in 1993 and 2005. In each exercise we estimated the relative distribution of the gender wage gap for the two sets of reference-comparison: two cohorts when they have the same age, and the same cohort in different phases of their life cycles.

The relative distribution presents some fundamental properties: i. it is not affected by the scale choice, being invariant to any monotonic transformation of the original variable, i.e. wages vs. log-wages); ii. Its basic unit of analysis is the population and not the individual; iii. It measures the proportion of individuals and executes their ranking, and not the wage values, as traditional methodologies (HANDCOCK e MORRIS, 1999)[1]. Formalizing, consider Y_0 as the random variable of interest for the reference population. The probability density function (*pdf*) of Y_0 is written as $f_0(y)$ and their cumulative density function (*cdf*) as $F_0(y)$. Consider the same measures for a comparison population Y: the *pdf* of Y is $f(y)$ and its distribution is $F(y)$. The relative distribution of Y to Y_0 (R) is defined by the random variable distribution:

$R = F_0(Y)$ (4)

Where:

R: relative distribution of the variable of interest (gender wage gap)
F_0: cumulative distribution function of the gender wage gap in the reference population
Y: gender wage gap in the comparison population

[1]Relative distribution analysis was performed in R software (R CORE DEVELOPMENT TEAM, 2009), with *reldist* command. Figures were constructed using Handcock e Aldrich (2002) codes.

Thus, the relative distribution R is obtained from Y transformed by the cumulative distribution function of Y_0, F_0. A property of R ensures that it is continuous at the [0,1] range, and we write the sample version of R as r, the relative data. An important amount related to the cumulative distribution function is its inverse function, which derives the quantile function, written as:

$$Q(p) = F^{-1}(p) = \inf_x \{x | F(x) \geq p\} \quad (5)$$

Where:

Q(.): quantile function of the gender wage gap
$F^{-1}(.)$: inverse of the gender wage gap *cdf*
x: gender wage gap
p: quantile of interest

The quantile function value, *Q(p),* defines the value of the *p* quantile of the wage distribution. By definition, the relative distribution is equivalent to a monotonic transformation of the variable of interest (equation 4). If its distribution is known, we can demonstrate equivalence between the original distribution and its transformation:

$$F_Y(y) = P(X \leq h^{-1}(x)) = F(h^{-1}(y)) \quad (6)$$

Where Y is the variable of interest and *h(x)* is its monotonic transformation. Hence, equation 6 shows that the cumulative distribution function of Y is equivalent to the same distribution function for a monotonic transformation of Y. Using this property, we can derive the cumulative distribution function of the R random variable:

$$G(r) = F(F_0^{-1}(r)) = F(Q_0(r)), \quad 0 \leq r \leq 1 \quad (7)$$

The first derivative of G(r) yields the relative density function:

$$g(r) = \frac{f(Q_0(r))}{f_0(Q_0(r))}, \quad 0 \leq r \leq 1 \quad (8)$$

The relative density *g(r)* can be interpreted as a density ratio: the ratio of the comparison population on the reference population in a given level of the response variable Y $(Q_0(r))$. Intuitively, the relative density informs how individuals in the comparison population would be located at the distribution of individuals in the reference population. If the variable of interest is uniformly distributed in both populations, the *g(r)* function assumes the value of 1.

We use the kernel smoothing for the estimation of the probability density of the gender wage gap in the reference and in the comparison populations, $f_0(y)$ and *f(y).* This non-parametric smoothing permits that we do not adopt an unnecessary or wrong assumption about the mathematical form of the distribution function of our variable of interest. The estimation of the kernel densities requires two components, the bandwidth and the smoothing function (*kernel*). The bandwidth provides the distance from a point x_0 of the variable distribution to which the density will be smoothed and the *kernel* smoothing function estimates local means. According to Dinardo and Tobias (2001), the choice of the *kernel* does not significantly affect the estimated probability curve shape, but the bandwidth has significant implications, as it contains a trade-off between bias and variance[2]. Here we assumed the optimal criteria proposed by Silverman (1986), who suggests, for

[2] Segundo Dinardo e Tobias (2001), esse *trade-off* existe porque um aumento da largura do intervalo acarreta em uma menor variância e no aumento da precisão da estimativa (sobresuavização). Entretanto, com uma suavização excessiva dos dados, corre-se o risco de obter uma densidade viesada.

the probability density estimation ($\hat{f}_n(x)$), the utilization of an Epanechnikov kernel (*k(x)*) with a bandwidth (h^*) defined as:

$$\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n}k\left(\frac{x - X_i}{h}\right) \quad (9)$$

$$k(x) = \frac{3}{4}\left(1 - x^2\right)I(x) \quad (10)$$

$$h^* = 1{,}3643\ \delta n^{-0,2}\sigma \quad (11)$$

Where x is the function value to be smoothed, *k(x)* is the kernel function, *I(x)* is the indicator function, h^* is the optimal bandwidth, δ is a constant for the Epanechnikov kernel (1,7188), n is the sample size and σ is the sample standard error.

In this article, we estimated the probability density of the gender wage gap for each cohort of women in two stages of their life cycle: with more (25-36 years) or less (37-48 years) mobility chances. From this exercise, it was possible to investigate how the cohorts improved (or worsened) their position in terms of the gender wage gap. The relative distribution approach also enabled the decomposition of effects in changes in the median or in the structure of the distribution, leading to enlightening analysis of the effects of changes over time.

Results

In this section, we present the results of our relative distribution application and the decompositions of the gender wage gap. As mentioned above, this variable is a measure of the gender discrimination in the Brazilian labor market: positive values indicate that observed female wages were higher than the predicted female wages using male returns, and, so, women were better off. On the other hand, negative values indicate that women earned lower wages than the predicted from the male returns, and, then, they were worse off.

Female predicted contra factual wages were constructed assuming male returns to individual, regional, and productivity attributes. For such, we estimated OLS mincerian regressions to obtain estimated coefficients and predict female wages, for their covariate composition, but male returns in three years: 1981, 1993 e 2005. The estimated results are in Table 2. The signals of the coefficients were mostly significant at 1% and consonant to the labor economics literature.

Table 2: OLS Regression estimates of Mincerian equations. Men, Brazil, 1981-1993-2005

Independent variables	1981		1993		2005	
	Coefficient (Std.Error)	p-value	Coefficient (Std.Error)	p-value	Coefficient (Std.Error)	p-value
y_schooling	0,119 (0,001)	0,000	0,125 (0,000)	0,000	0,121 (0,001)	0,000
age	0,130 (0,022)	0,000	0,083 (0,006)	0,000	-0,018 (0,037)	0,628
age2	-0,002 (0,000)	0,000	-0,001 (0,000)	0,000	0,000 (0,000)	0,369
urb	0,338 (0,010)	0,000	0,272 (0,010)	0,000	0,198 (0,013)	0,000
metrop	-0,131 (0,005)	0,000	-0,103 (0,004)	0,000	-0,074 (0,005)	0,000
Southeast	0,130 (0,024)	0,000	0,099 (0,020)	0,000	0,041 (0,018)	0,027
South	0,098 (0,025)	0,000	0,173 (0,021)	0,000	0,088 (0,020)	0,000
Northeast	-0,105 (0,025)	0,000	-0,273 (0,021)	0,000	-0,325 (0,019)	0,000
Center-West	0,066 (0,028)	0,020	0,127 (0,023)	0,000	0,115 (0,023)	0,000
constant	-1,656 (0,333)	0,000	-1,444 (0,117)	0,000	0,559 (0,784)	0,476
N. Observations	34599		44015		26431	
R-squared	0,457		0,442		0,435	
F-statistic	3234,670		3872,000		2258,570	
Prob>F	0,000		0,000		0,000	

Figure 5 shows the estimated kernel probability density for the gender wage gap, comparing two cohorts of women in each stage of their life cycle (25-36 and 37-48 years of age). The median of the distribution is lower than zero for the two cohorts in both stages of their life cycles. It means that the summary measure of the distributions points to a worse position of women in the labor market. For women in the age group of potential larger mobility (25-36), the distribution of the wage gap is more flattened for the cohort 2, i.e., lower density in the median gap, and there was an increase in the density for positive values of the gender wage gap. It suggests an improvement for women, in terms of the gender wage gap for the cohort 2. In the case of the age group of potential smaller wage mobility (37-48), Figure 1 reveals the same trend. Although the density of women with median gap has increased from cohort 1 to 2, the density with a positive gap increase substantially, indicating an improvement of wage positions for women in cohort 2.

Figure 5: Kernel density estimator of gender wage differential by cohort and life cycle, Brazil.

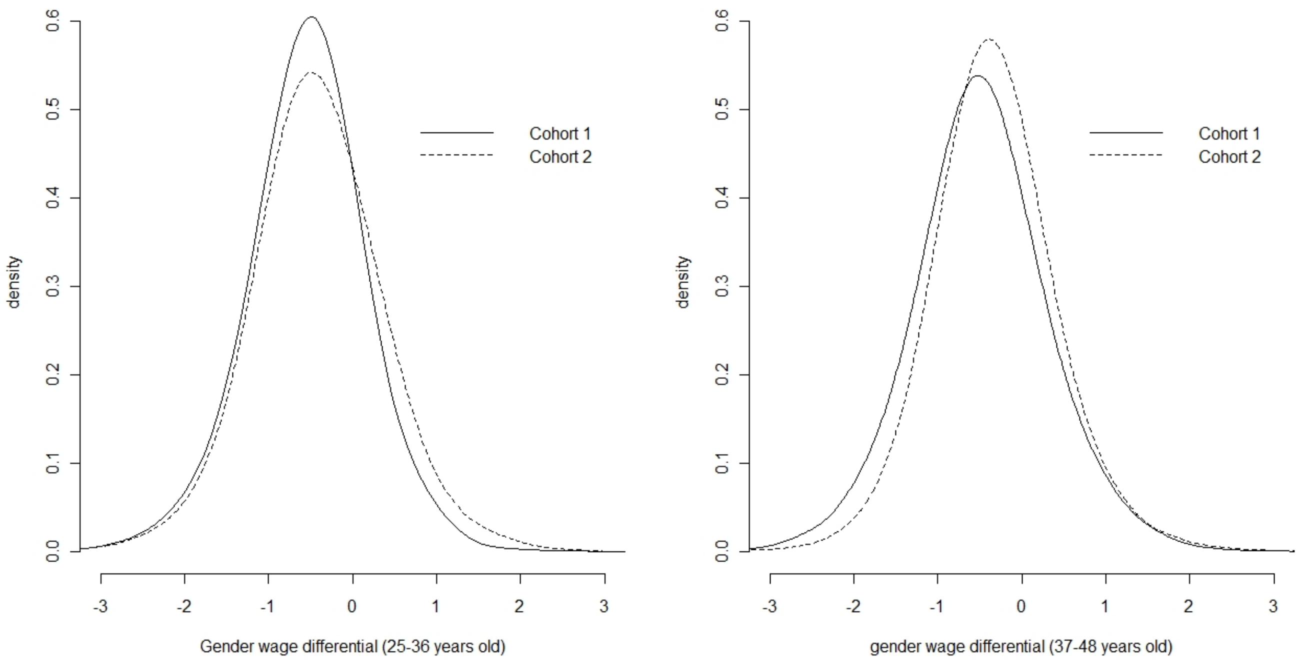
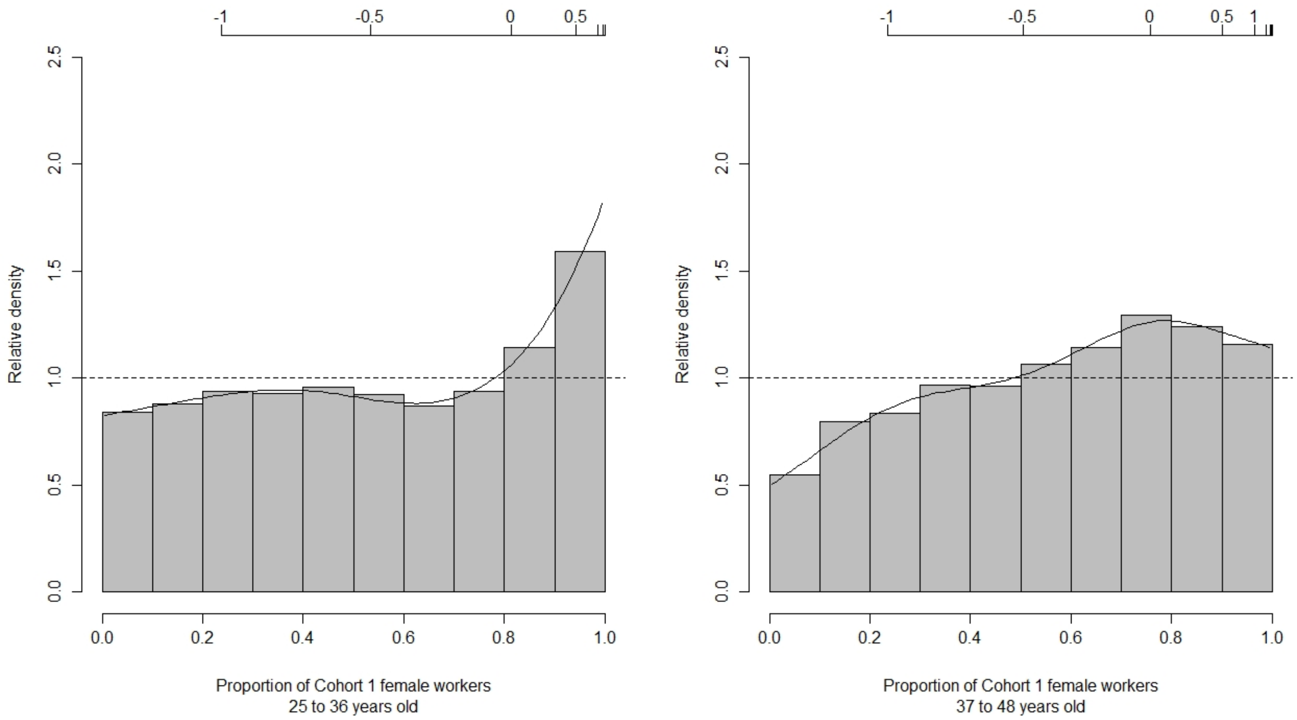


Figure 6 organizes prior information into the relative distribution framework. Again, two cohorts are compared in each stage of the working life table. The x axis denotes the deciles of the distribution of the wage gap in the reference population, the cohort 1. The y axis denotes the measure of the relative density: the proportion of individuals in the comparison group (cohort 2) who would be situated in each decile of the wage gap distribution in the reference group (cohort 1). The upper axis denotes the cutting points of the wage gap. The bars represent the ratio between individuals in the comparison group (cohort 2) and the reference group (cohort 1) in each distribution decile, and their smoothed values compose the curve. This figure evidences, in the left side, that the relative density is less than 1 till the 8th decile of the wage distribution and afterwards assumes values greater than 1. This means that cohort 2 was constituted by a smaller contingent of female workers with a negative gender gap than cohort 1 (values of the gap less than zero up to the 8th decile). From this decile of the gender gap distribution, the relative density is greater than 1, suggesting that there was more individuals in cohort 2 situated in the deciles of the positive gap of cohort 1 (upper axis). Hence, in the youngest age group, the gender gap tended to be more favorable to women in the subsequent cohorts.

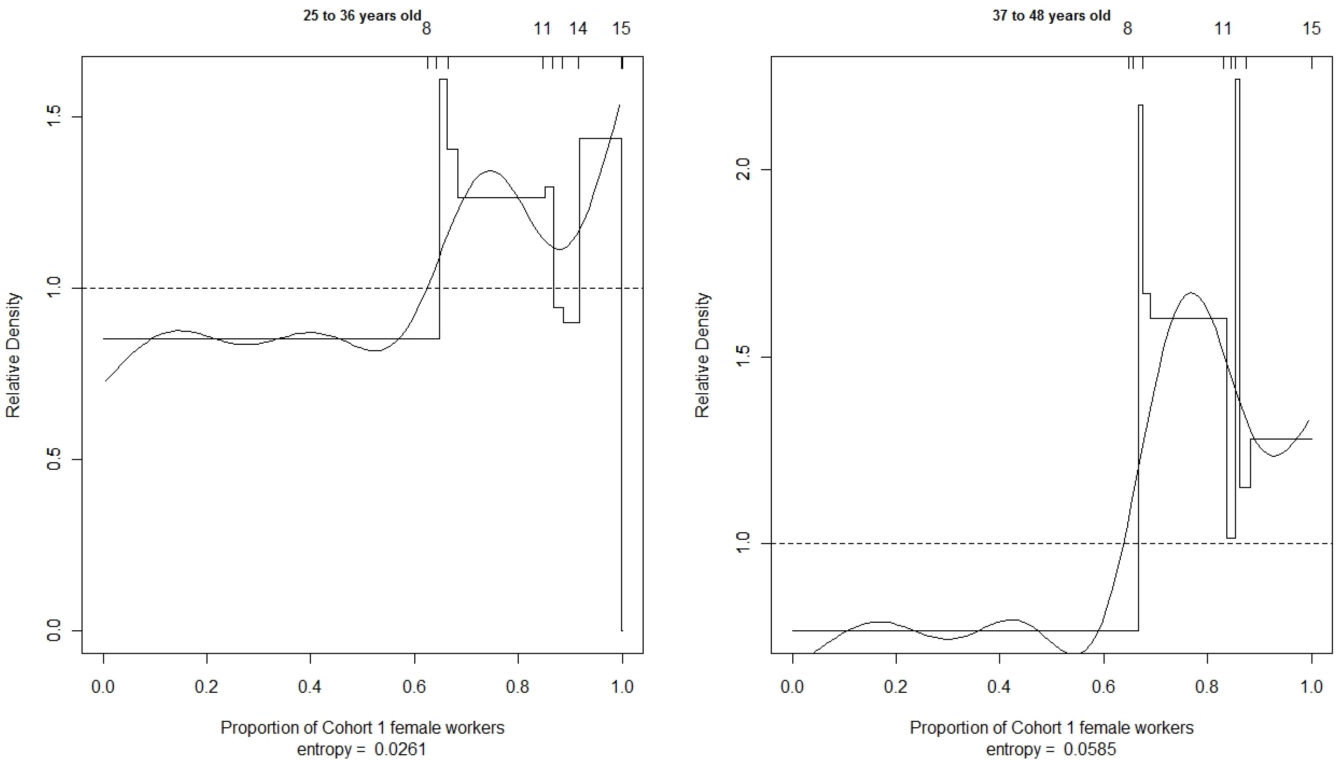
The same trend of reduction of the gender wage gap between the cohorts was evidenced when we considered individuals in an older age group (37-48), as we can see at the right frame at Figure 6. Relative density is less than 1 up to the 5th decile of the distribution, and then turns to be greater than 1. Despite the improvement of position of women of cohort 2 in relation to cohort 1, the gender wage inequality does not evolved as favorable to women as we could observe for youngest age group (25-36) considered. There is yet a considerable proportion of women with a negative wage gap (up to the 8th decile, the value at the upper axis is negative and the relative density is greater than 1).

Figure 6: Relative Density of the gender wage differential. Cohort 1 versus Cohort 2, by age group, Brazil



We are also interested in access how the educational composition of the female labor force varied between cohorts and age groups. The relative distribution approach allows comparisons between two populations when the response variable is categorical. So, we computed the relative distribution of the years of schooling for women in each cohort. Results are shown in Figure 7. It signalizes that the cohort 2 is more educated than cohort 1. The relative density is lower than 1 up to 8 years of schooling (upper axis), which means that there was a smaller proportion of women at cohort 2 situated at the lowest deciles of educational distribution of cohort 1. At the same time, the relative density is greater than 1 for those with more than 8 years of schooling. Analogously, it implies that there were a greater proportion of women in cohort 2 with more than 8 years of schooling than in cohort 1.

Figure 7: Relative Density of the years of schooling. Cohort 1 versus Cohort 2, by age group, Brazil



A comparison of the evolution of the gender wage gap within each cohort over time was accomplished, decomposing the global relative distribution for each cohort in effects of changes in the median and changes in the structure. Chart (a) at Figures 8 and 9 displays the global relative density, or how women of 37-48 years of age would be placed in the gap scale of women of 25-36 years of age in cohort 1. Chart (b) displays the relative density adjusted for changes in the median. If the adjusted curve is increasing, there was in the period a shift to the right in the distribution of the response variable, i.e., an increase in the median. Oppositely, there was a reduction in the median of the wage gap distribution. Chart (c) displays the path of the relative distribution adjusted for changes in the wage structure. If this curve has a U shape, this suggests a polarization of the distribution, i.e., an increase of density in the first and last deciles. On the other hand, if the curve has an inverse U shape, the trend was to equalizing the distribution around its median. Above each chart of the relative distribution decomposition in effects of changes in the median and in the structure, it is shown the estimated entropy, which indicates the dissimilarity measure between the reference and comparison populations, and is equally decomposable in effects of changes in median and structure.

Figure 8 shows the evolution of the gender wage gap in cohort 1 between 1981 and 1993, i.e., when they were 25-36 years of age (reference group) until there were 37-48 (comparison group). Chart (a) displays the global relative density, with a dissimilarity of 0.0135. We can observe that women in cohort 1 improved their position in terms of the gender wage gap over time, as the relative density is increasing. This indicates that, in 1993, there were fewer women of cohort 1 in the first deciles of the gap distribution. Chart (b), which shows the relative density, adjusted to effects of changes in the median, displays an increasing curve. It suggests an increasing median wage gap over their life cycle. However, the effect of the change in the median explains just one

Figure 8: Location/Shape relative decomposition of the gender wage differential. Cohort 1, 25-36 versus 37-48 years, Brazil.

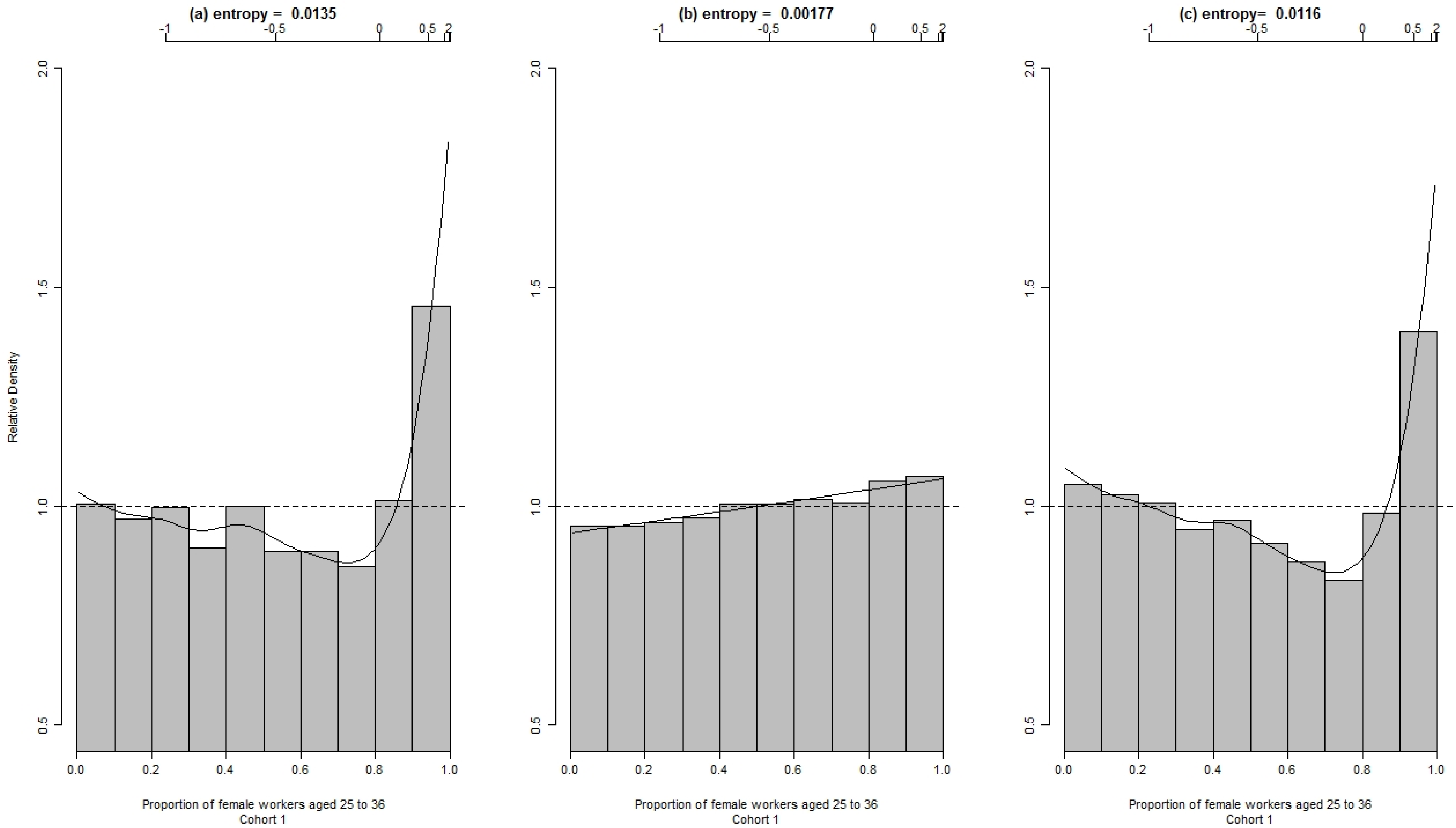
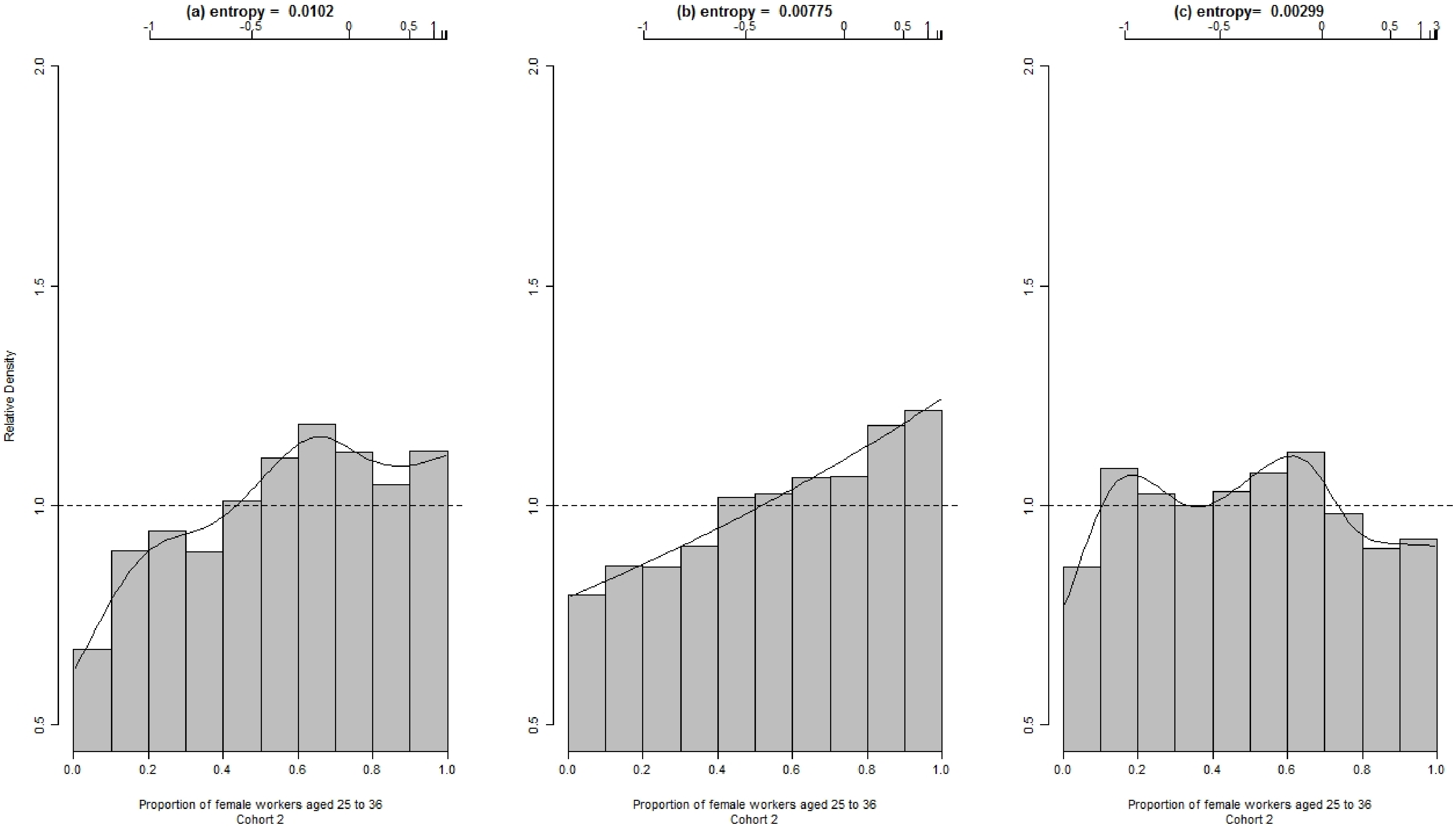


Figure 9: Location/Shape relative decomposition of the gender wage differential. Cohort 2, 25-36 versus 37-48 years, Brazil.



part of the observed global relative density, i.e., only 13% of the global entropy. Chart (c) expresses that the effect of changes in the structure was the most important determinant of the global relative density, as it explains 87% of the global entropy and points to an interesting outcome: a polarization of the gender distribution. Within cohort 1, over time, more women in the oldest age group studied presented high negative and positive gaps.

The evolution of the gender wage gap of women in cohort 2, nevertheless, is distinct. Figure 9 shows that, in chart (a), the global relative density of those women increased over their life cycle, indicating an improvement in terms of the gender wage gap. The dissimilarity between the distributions of the younger women is 0.0102. The adjustment of the relative distribution for the effects of changes in the median (chart (b)) was the preponderant effect, explaining 75% of the dissimilarity between the distributions of younger and older women of cohort 2. Similarly, it points to a more favorable trend for women towards a positive gender gap. The adjusted relative density for changes in the structure was less decisive (25% of global entropy), and also evidences a slight polarization of women in cohort 2.

The final application of this article was the adjustment of the gender gap relative distribution for the effects of changes in the composition in terms of the educational level of women. The question was how the higher education of cohort 2 contributed to explain the observed wage gap relative distribution between cohorts 1 and 2, which is more advantageous for cohort 2. Figures 10 and 11 plots the wage gap relative distribution in relation to the education covariate, decomposing the relative distribution in effects of changes in the median and in the structure. Chart (a) displays the global relative density; chart (b) displays the relative density adjusted for changes in the educational composition and chart (c) displays the residual relative distribution. Entropy measures were also computed, and are decomposable in these effects. In figures10 and 11, we can observe a global relative density favorable to cohort 2. Estimated entropy measures in the both stages of the cohorts working life cycles are around 0.02. Charts (b) in both figures reveal uniform densities, which mean that there was not a significant change in the marginal density of the wage gap, adjusted by the education level. The global entropy is not significantly explained by this component. The residual effects (charts (c)) are the main effect explaining the relative density between cohorts 1 and 2, for both age groups studied.

<u>Final remarks</u>

The reduction of the gender inequality in the Brazilian labor market is evidenced by the recent literature. However, few authors evaluate the distribution of the wage gap. If the mean wage gap is not representative of the Brazilian population, then policies based in these measures would be translated into undesired outcomes. In our article, we seek to contribute to the debate incorporating the analysis of the wage gap over time between and within cohorts.

Using a non-parametric methodology of relative distribution (Handcock e Morris, 1999), we could construct simple figures and an exploratory and contra factual investigation of the gender wage inequality in Brazil, over last decades. We found out that, for each age group, women in more recent cohorts are better off in terms of the gender gap. Besides, we verified that each cohort experienced an improvement in the gender gap over the life cycle. For the oldest cohort, changes in the structure of the wage gap were preponderating, and for the youngest cohort, the change in the median gap was the main determinant of the upward mobility over its working life cycle. Finally, results point that changes in the marginal density of education, related to the wage gap (adjusting the relative distribution), were not crucial for the observed shifts in the wage gap between the cohorts, by age group, even when we consider that the more recent cohort is more educated than the earliest studied here.

Figure 10: Schooling compositional adjustment of the relative density of the gender wage differential. Cohort 1 versus Cohort 2, 25-36 age group, Brazil.

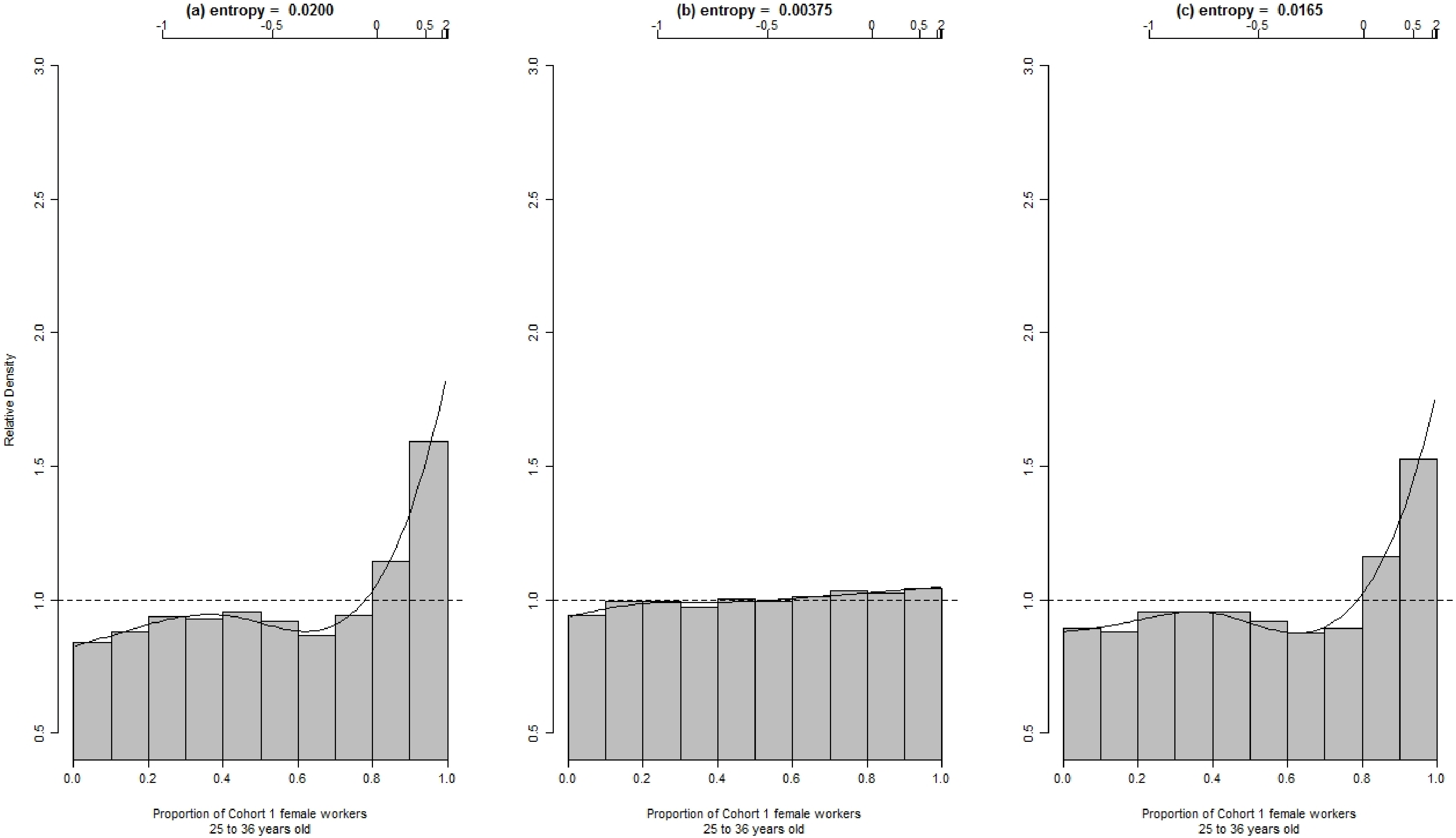
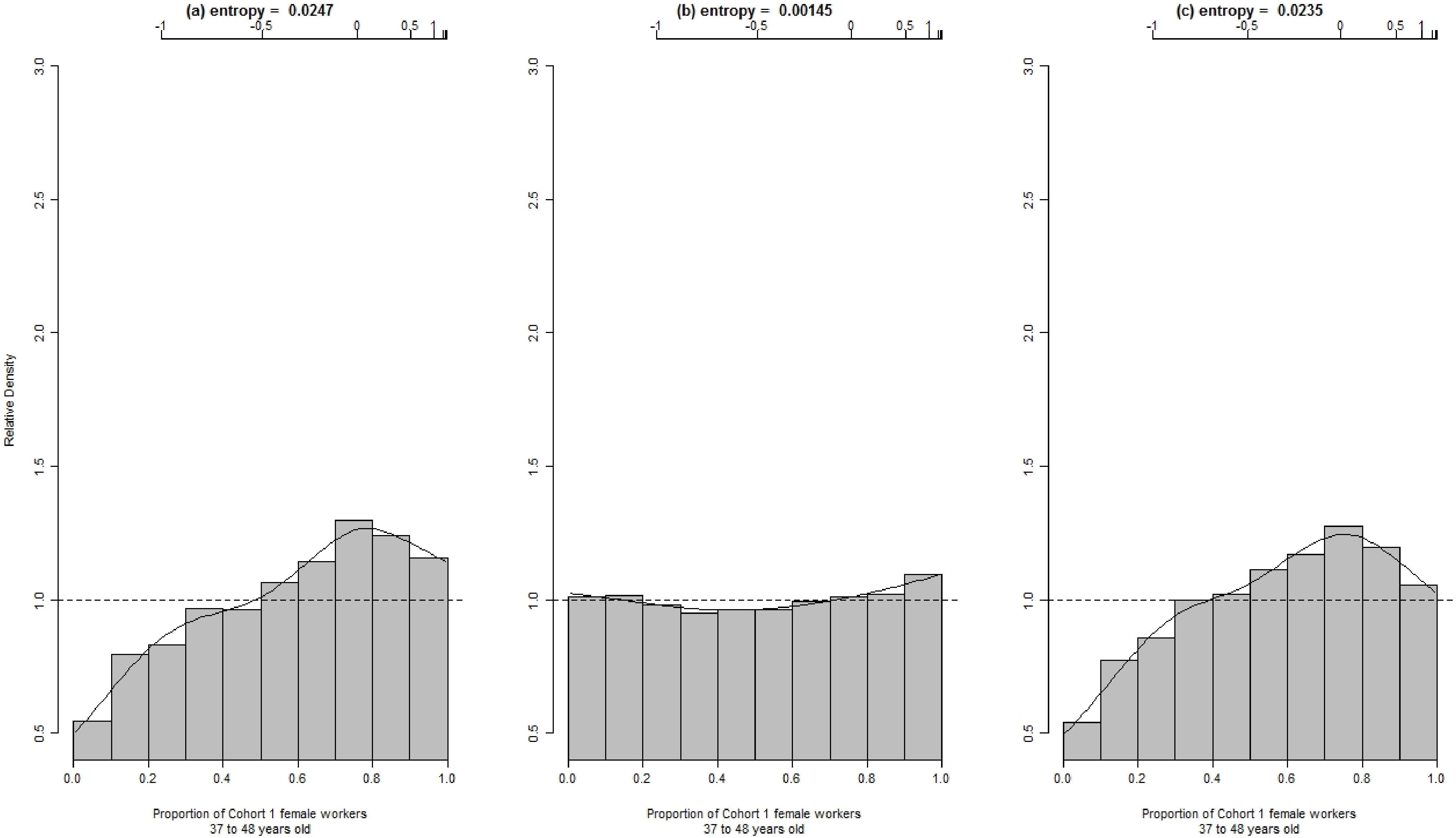


Figure 11: Schooling compositional adjustment of the relative density of the gender wage differential. Cohort 1 versus Cohort 2, 37-48 age group, Brazil.



References

Altonji, J.G., Blank, R.M. 1999. Race and gender in the labor market. In: Ashenfelter, O., Card, D. (eds.) *Handbook of Labor Economics, V.3*. Elsevier Science.

Barros, R.P., Mendonça, R.S.P. 1995. Os determinantes da desigualdade no Brasil. *Texto para Discussão 377, IPEA.*

Barros, R.P., Henriques, R., Mendonça, R.S.P. 2001. A estabilidade inaceitável: desigualdade e pobreza no Brasil*. Texto para Discussão 800, IPEA*.

Bernard, A.B., Jensen, J.B. 1998. *Understanding Increasing and Decreasing Wage Inequality*. Report 6571: National Bureau of Economic Research.

Bernhardt, A., Morris, M., Handcock, M. 1995. Women's gains or men's losses? A closer look at the shrinking gender gap in earnings. *American Journal of Sociology*, 101: 302-28.

Bernhardt, A., Morris, M., Handcock, M., Scott, M. 1998. Inequality and mobility: trends in wage growth for young adults. *American Journal of Sociology*.

Blackburn, M.L., Bloom, D.E., Freeman, R. 1990. The declining economic position of less skilled American men. In: Burtless, G. (ed.) *A Future of Lousy Jobs?* Washington, DC: Brookings Institution.

Blau, F.D. 1996. Where are we in the economics of gender? The gender pay gap. *NBER Working Paper 5664*.

Blau, F.D. 1998. Trends in the well-being of American women, 1970-1995. *Journal of Economic Literature*: 36:112.

Blau, F.D., Kahn, L.M. 2000. Gender differences in pay. *NBER Working Paper* 7732.

Buchinsky, M. 1994. Changes in the U.S. wage structure 1963-1987: application of quantile regression. *Econometrica*, 62(2): 405-58.

Buchinsky, M. 1999. Recent advances in quantile regression models. *Journal of Human Resources*, 33(1): 88-126.

Coelho, A.M., Corseuil, C.H. 2002. Diferenciais salariais no Brasil: um breve panorama. *Texto para Discussão 898, IPEA.*

Danziger, S., Gottschalk P. (eds.) 1997. *Uneven Tides: Rising Inequality in America*. New York: Russell Sage Foundation.

Dinardo, J.; Tobias, J.L. 2001. Nonparametric density and regression estimation. *Journal of Economic Perspectives*, 15 (4): 11-28.

Dooley, M., Gottschalk, P. 1982. Does a younger male labor force mean growing earnings inequality? *Monthly Labor Review*, 105: 42-45.

Dooley, M., Gottschalk, P. 1984. Earnings inequality among males in the United States: trends and the effect of labor force growth. *Journal of Political Economy*, 92: 59-89.

Dooley, M., Gottschalk, P. 1985. The increasing proportion of men with low earnings in the United States. *Demography*, 22: 25-34.

Fernandes, R.; Menezes Filho, N.A. 1999. A evolução da distribuição dos rendimentos do trabalho no Brasil Metropolitano: uma análise para o período de 1983 a 1997. *Anais do 27º Encontro Nacional de Economia*. Belém: ANPEC.

Ferreira, F.H.G.; Barros, R.P. 1999. The Slippery Slope: Explaining the Increase in Extreme Poverty in Urban Brazil: 1976-1996. *Texto para Discussão 404, PUC-RJ*

Fortin, N.M., Lemieux, T. 1997. Rank regressions, wage distributions, and the gender gap. *Journal of Human Resources*, 33(3): 610-43.

Freeman, R. 1976. *The Overeducated American*. New York: Academic Press.

Freeman, R. 1994. *Working Under Different Rules*. New York: Russell Sage Foundation

Handcock, M.S.; Morris, M. 1999. *Relative Distribution Methods in the Social Sciences*. New York: Springer-Verlag.

Handock, M.; Aldrich, E. M. 2002. *Applying relative distribution methods in R*. Center for Statistics and Social Sciences. University of Washington (*Working Paper* n. 27).

Henriques, R. (org.) 2000. *Desigualdade e Pobreza no Brasil*. Rio de Janeiro: IPEA.

Hoffman, R. 1998. *Distribuição de Renda: Medidas de Desigualdade e Pobreza*. São Paulo, EDUSP.

Howell, D., Duncan, M., Harrison. B. 1998. *Low Wages in the U.S. and High Unemployment in Europe: A Critical Assesment of the Conventional Wisdom*. CEPA Working Paper No. 5. New York: New School for Social Research.

Juhn, C., Murphy, K.M., Pierce, B. 1991. Accounting for the slowdown in black-white wage convergence. In: Kosters, M.H. (ed.) *Workers and Their Wages.* Washington, D.C.: American Institute Enterprise Press.

Juhn, C., Murphy, K.M., Pierce, B. 1993. Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3): 410-42.

Katz, L.F., Autor, D.H. 1999. Changes in the wage structure and earnings inequality. In: Ashenfelter, O., Card, D. (eds.) *Handbook of Labor Economics.* Amsterdan: North-Holland.

Katz, L.F., Murphy, K.M. 1992. Changes in relative wages, 1963-1987: supply and demand factors. *Quarterly Journal of Economics*, 107: 35-78.

Lam, D. 1999. Generating Extreme Inequality: Schooling, Earnings and Intergenerational Transmission of Human Capital in South Africa and Brazil. (mimeo.)

Lerman, R.I. 1997. Reassessing trends in US earnings inequality. *Monthly Labor Review* :17-25.

Massey, D.S. 1996. The age of extremes: Concentrated affluence and poverty in the twenty-first century. *Demography*, 33: 395-412.

Menezes-Filho, N.A.; Fernandes, R.; Pichetti, P. 2000. The distribution of male wages in Brazil: some stylised facts. In: Henriques, R. (org.) *Desigualdade e Pobreza no Brasil*. Rio de Janeiro: IPEA.

Morris, M. 1996. Vive la Difference: persistence and change in gender inequality. In: Baron, J., Grusky, D., Treiman, D. (eds.) *Social Differentiation and Social Inequality: Theoretical and Empirical Inquiries*, pp.211-240. Boulder: Westview Press.

Morris, M., Bernhardt, A. Handcock, M. 1994. Economic inequality: new methods for new trends. *American Sociological Review*, 59: 205-19.

Morris, M., Bernhardt, A. Handcock, M. 2000. Trends in earnings dispersion in the National Longitudinal Survey of Youth and the Current Population Survey. *Monthly Labor Review,* 123: 48-61.

Morris, M., Bernhardt, A., Handcock, M., Scott, M. 2001. *Divergent Paths: Economic Mobility in the New American Labor Market*. New York: Russell Sage Foundation.

Morris, M., Handcock, M.S. 1998. Relative distribution methods. *Sociological Methodology*, 28: 53-89.

Morris, M., Western, B. 1999. Inequality in earnings at the close of the twentieth century. *Annual Review of Sociology*, 25: 623-57.

Murphy, K., Welch, F. 1993. Inequality and relative wages. *The American Economic Review*, 83(2): 104-109.

Mwabu, G., Schultz, T. P. 1996. Education returns across quantiles of the wage function: alternative explanations for returns to education by race in South Africa. *American Economic Review*, 86(2): 335-9.

Neal, D., Rosen, S. 1998. Theories of the distribution of labor earnings. *NBER Working Paper 6378*.

Oliveira, A.M.H.C. 2001. *Occupational Gender Segregation and Effects on Wages in Brazil*. Trabalho apresentado no Encontro da IUSSP, Salvador, Agosto de 2001.

Oliveira, A.M.H.C., Rios-Neto, E.L.G. 2003. *Tendências da Desigualdade Salarial para Coortes de Mulheres Brancas e Negras No Brasil*. Trabalho aceito para apresentação no XXXI Encontro Nacional de Economia, em Porto Seguro, 10 a 12 de Dezembro de 2003.

R Development Core Team. 2009. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Ramos, C.A.; Santana, R. 2000. Desemprego, pobreza e desigualdade. In: Henriques, R. (org.) *Desigualdade e Pobreza no Brasil*. Rio de Janeiro: IPEA.

Ramos, L. 1993. *A Distribuição de Rendimentos no Brasil*: 1976/85.

Ramos, L.; Vieira, M.L. 2000. Determinantes da desigualdade de rendimentos no Brasil nos anos 90. In: Henriques, R. (org.) *Desigualdade e Pobreza no Brasil*. Rio de Janeiro: IPEA.

Silverman, B. W. 1986. *Density Estimation.* Londres: Chapman and Hall.

Soares, S.S.D. 2000. O Perfil da Discriminação no Mercado de Trabalho – Homens Negros, Mulheres Brancas e Mulheres Negras. *Texto para Discussão 769, IPEA*.

Suen, W. 1997. Decomposing wage residuals: unmeasured skill or statistical artefact? *Journal of Labor Economics*, 15(3): 555-66.