Human Capital and the Recent Decline of Earnings Inequality in Brazil[*]

Priscilla Albuquerque Tavares[**]
Naercio Aquino Menezes-Filho[***]

Abstract

Earnings inequality has started to decrease in Brazil in recent years, after remaining very high for decades. We describe this decline using a flexible decomposition technique and assess the contributions of education and experience. We conclude that the main factors leading to the reduction of inequality in Brazil are the drop in education earnings differentials and the decline in the dispersion within demographic groups. The paper demonstrates the powerful impact that education can have on reducing inequality.

Keywords: Human Capital, Income Inequality, Wages, Education.

JEL Codes: J31, J45.

1. Introduction

Brazil has the world's eighth largest economy (IMF, 2008). Nevertheless, 21.4% of the country's people live in poverty, and 7.3% in extreme poverty (IPEADATA, 2009). This contradiction is the result of the country's glaring income inequality (UNDP, 2010).[1] However, after decades remaining at a very high and stable level, inequality has recently started to decline in Brazil and in several other Latin-American countries (Lopez-Calva and Lustig, 2010). The aim of this paper is to understand the reasons behind the decline in Brazilian inequality, using a flexible econometric approach and focusing on the role that education and age have played on this phenomenon.

The focus of this paper is on observable skills because human capital is one of the main determinants of earnings and; therefore, affects earnings inequality. Moreover, education has improved substantially in recent years in several Latin American countries. Therefore, it is of interest to investigate whether the decline of inequality is related to this education upgrade in Brazil, a major Latin American country that has always been seen as very unequal.

The relationship between education and inequality depends on two factors: the education inequality among workers in the job market (composition effect) and the monetary value the market attributes to each additional year of schooling (price effect), as described by Ram (1990) and Knight and Salbot (1983). In Brazil, both the education wage differentials and the great educational disparity among workers have been traditionally important to explaining wage inequality (Lam and Levinson, 1992). In this paper we assess what has been happening in recent years in regards to education inequality in Brazil, and we thoroughly examine the impact of this factor on earnings inequality.

There is mounting evidence in the literature that the behavior of income inequality is better explained by models that allow for wage changes that are different for workers located at different points of the wage distribution. Autor et al. (2005), for example, argue that the wage differentials in the upper part of the distribution (90th/50th percentile) have increased continuously in the United States since the 1970s. However, in the bottom part (50th/10th percentile) of the distribution, inequality increased in the 1980s and has remained virtually unchanged since that time. Lemieux (2006a,b) corroborates these results by arguing that changes in the returns to measured skills have played a significant role in the growth of inequality since the early 1970s. He also asserts that the long-run increase in American income inequality is concentrated in the upper part of the distribution and is basically due to the rising returns to postsecondary education.

In Brazil, it is also very instructive to observe how earnings have changed in the different parts of a distribution. Figure 1 describes the evolution of real wages in Brazil since 1995 at the 10th, 50th and 90th percentiles. Wages in 1995 are

---

[1]In a comparison with 126 countries, Brazil is the 10th most unequal.

set to zero, so that the points in the figure are cumulative values with respect to 1995. The figure shows, quite interestingly, that wages in the bottom part of the distribution increased much more than in the median and top. While wages at the 10th percentile grew by about 57% , median wages increased by 13% and wages at the 90th percentile actually fell in real terms.

Figure 1
Real wage changes for selected percentiles
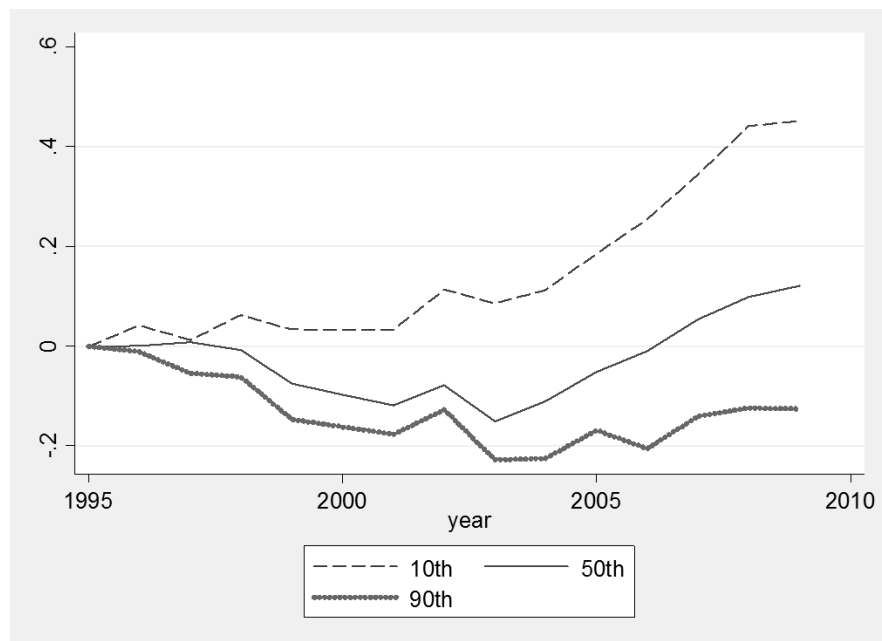
Source: 1995 to 2009 PNAD.
Cumulative changes in the logarithm of real hourly
wages (2005 prices) with respect to 1995.

In light of this scenario, this article examines the effects of changes in the "composition" of workers' attributes and their "prices" on income inequality in Brazil between 1995 and 2009 using a quantile regression approach. This permits an evaluation of wage changes at different points of the earnings distribution. This contrasts with recent literature that has examined the issue of wage inequality in Latin America.

Ferreira et al. (2008), for example, undertake a preliminary investigation into the behavior of inequality in Brazil between 1981 and 2004. The authors focus on the role of inflation, but also examine the behavior of the returns to educa-

tion, rural-urban convergence and social assistance to the poor. Manacorda et al. (2010) scrutinize the behavior of the returns to education in five Latin-American countries. Using a model of demand for skills, the authors determine that the rise in the supply of workers with intermediate education has depressed wage differentials. Neither of these papers; however, decompose the role human capital played on inequality into components between and within demographic groups at different points along the earnings distribution.[2]

This paper is organized into five sections, including this introduction. The next section describes the data and presents some descriptive evidence. Section 3 establishes the econometric methodology, while Section 4 highlights the econometric results. Section 5 serves to conclude the paper.

2. Data

We use data from the National Household Survey (*Pesquisa Nacional por Amostra de Domicílios* – PNAD/IBGE) from 1995 to 2009, conducted by the Brazilian Institute of Geography and Statistics.[3] The worker sample consists of men from 25 to 60 years of age, with a strictly positive principle job income and workweek. We split the sample into 1,980 cells, defined by the survey year, worker age (in years) and education, grouped into four categories: zero to three; four to seven; eight to 11; and 12 or more years of study. To measure labor income we use the logarithm of real hourly wages, at 2005 prices (lw),[4] and our measures of inequality will be the variance of (log) wages, which is perfectly decomposable. Table 1 presents the basic descriptive statistics of this variable.

Figure 2 describes the evolution of the Gini coefficient calculated on the basis of two different income measures: household per capita income and labor market earnings. The figure shows that both measures of inequality fell substantially in recent years when compared to their 1995 levels. Earnings inequality fell 23.4% from an initial value of 0.394, while income inequality fell 10.6% from a value of 0.600 in 1995. Therefore, it seems that in order to understand the reasons behind the drop in overall inequality, it is important to grasp the determinants of inequality in the labor market.

Between 1995 and 2009, the education quality of the Brazilian labor force increased significantly, with the average level of schooling rising from 6.1 to eight years. Figure 3 describes the education composition of the workforce in Brazil over our sample period. The share of the least educated (less than three years of education) fell from 30.2% in 1995 to around 17.2% in 2009, whereas the share of individuals with high school education rose from 18% to 33%. The share of

---

[2]Menezes-Filho et al. (2006) used a similar methodology to decompose the evolution of inequality in Brazil between 1981 and 1997, before inequality started to fall.

[3]For 1991, 1994 and 2000, when the PNAD was not conducted, we interpolated the variables using the simple average of the two adjacent years.

[4]We used the income deflator of Corseuil and Foguel (2002).

Table 1
Data description

| Education group	Number of observations	Population represented	Mean number of cell observations	Minimum number of cell observations	Maximum number of cell observations	Median of the real log hourly wage
Primary	228,829	106,205,082	462	172	643	0.76
Secondary	282,775	131,180,480	571	183	825	1.25
High School	351,875	160,504,275	711	79	1623	1.76
College	114,263	53,195,202	231	39	526	2.80

Source: 1995 to 2009 PNAD.

Note: Primary: zero to three years of schooling;

Secondary: four to seven years of schooling;

High School: eight to 11 years of schooling;
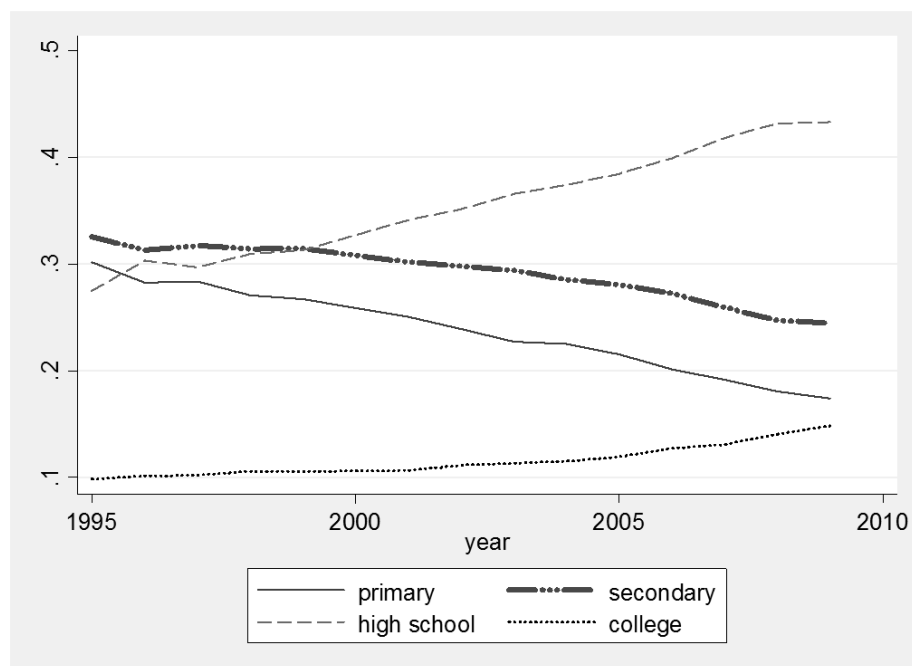
College: 12 or more years of schooling.

Figure 2
Gini Index – Per capita family income and earnings
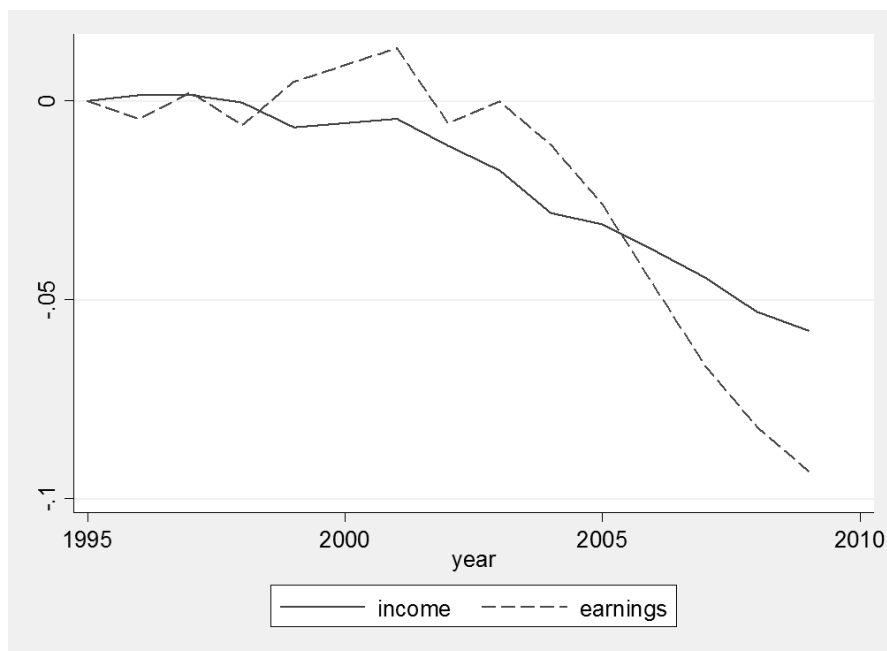
Sources: 1995 to 2009 PNAD and IPEADATA.

individuals with college education rose from 9.8% to 14.2%. These are rapid changes for such a small period of time and are likely to have an impact on the labor market and inequality.

Figure 3
Education composition of labor force (1995-2009)

Sources: 1995 to 2009 PNAD.

Notes: (a) Male workers aged 24 to 56 years; (b) Education groups: primary – zero to three years of schooling; secondary – four to seven years of schooling; high school – eight to 11 years of schooling; college – 12 or more years of schooling.

The impact of the changes in the education composition on the labor market can be seen in Figure 4, which depicts the behavior of the education wage differentials over time. Returns to secondary education (with respect to illiteracy) and to high school education fell quite substantially between 1995 and 2009. Returns to college education increased quite rapidly between 1995 and 2003, but fell afterwards. Although our aim in this paper is not to explain the behavior of the education wage differentials, related research shows that they reflect the evolution of the relative supply of the four education groups depicted above (Binelli et al., 2008).

Figure 5 shows the behavior of the average returns to education in our sample

Figure 4
Education wage differentials (1995-2009)
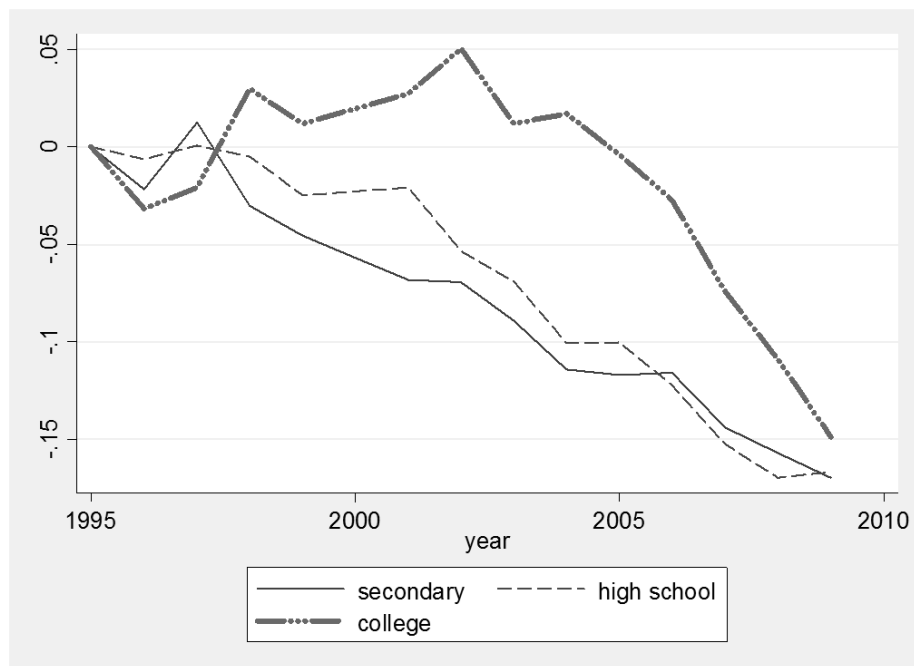
Source: 1995 to 2009 PNAD.

Note: Cumulative change in log wage differentials with respect to the previous education groups.
Education groups: primary – zero to three years of schooling; secondary – four to seven years of
schooling; high school – eight to 11 years of schooling; college – 12 or more years of schooling.

period. Returns fell almost continuously between 1995 and 2009, despite the rise, documented above, in the college education wage differentials in the 1990s. This reflects the fact that the majority of the Brazilian population has far less than college education, thus returns to more basic education levels dominate the behavior of average returns.

Figure 5
Mean returns to education (1995-2009)
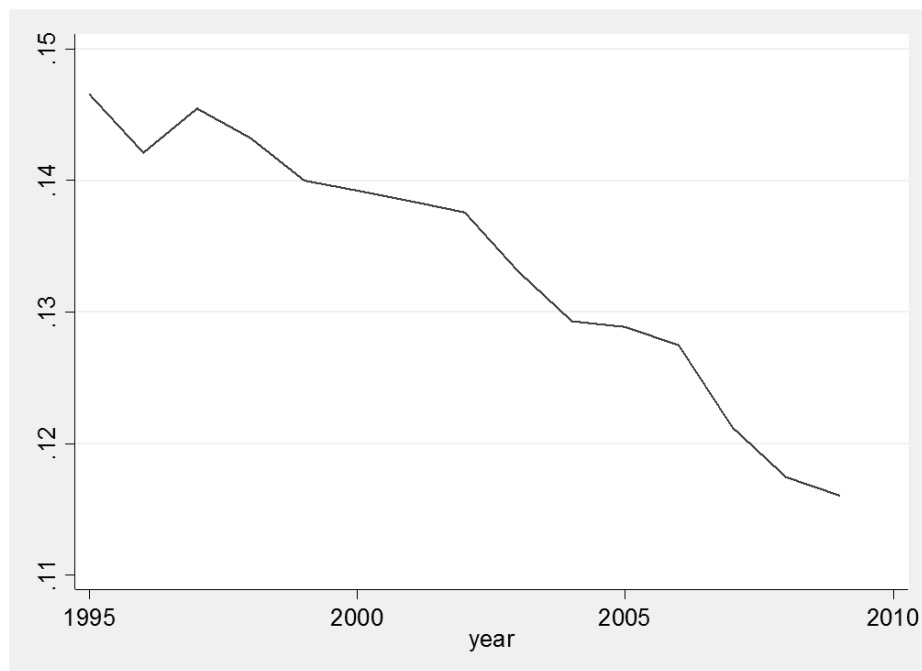
Source: 1995 to 2009 PNAD.

Note: Regression coefficients of the logarithm of real hourly wages (2005 prices) against years of schooling, age and age squared.

The behavior of wage inequality within the education groups is also of substantial interest, as it allows us to infer the evolution of the demand for other (unobserved) measures of skills. We can see from Figure 6 that both the 90th-50th and the 50th-10th quantile differentials fell for most education groups. The exceptions are the 50th-10th quantile differential in the primary education group, and the 90th-50th quantile differential in the college-educated group, which have both increased in recent years. Therefore, it seems that inequality is increasing in the top of the distribution, similar to the trend occurring in the United States

(Lemieux, 2006a), as well as in the very bottom of the distribution. In the following section, we attempt to describe these patterns using a flexible decomposition approach.

Figure 6
Within-group variance by education group
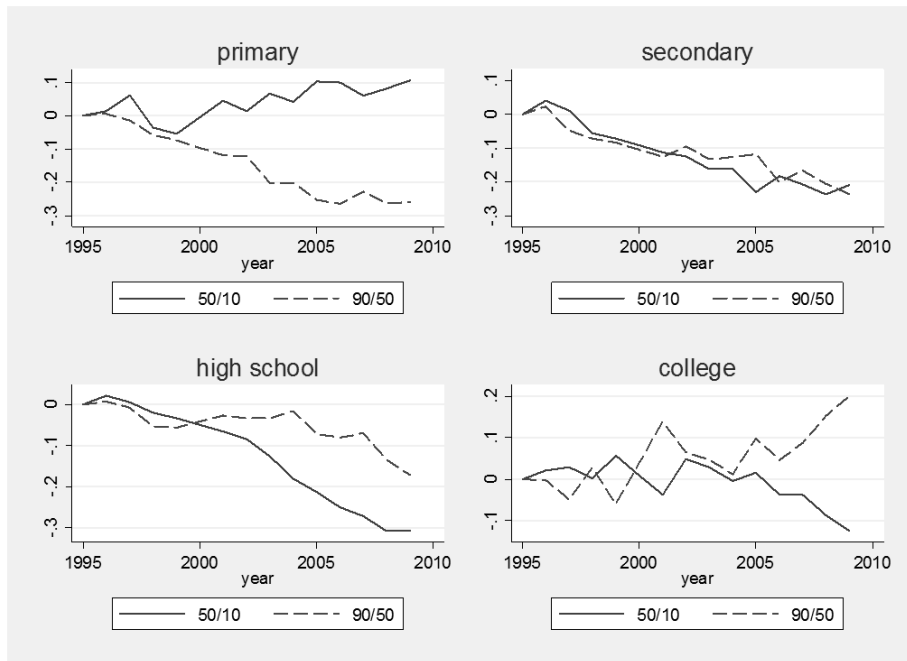
Source: 1995 to 2009 PNAD.

Note: Cumulative changes in the logarithm of real hourly wages differentials with respect to 1995. Education groups: primary – zero to three years of schooling; secondary – four to seven years of schooling; high school – eight to 11 years of schooling; college – 12 or more years of schooling.

3.   Econometric Methodology

Our estimation model is based on the work of MaCurdy and Mroz (1995) and Gosling et al. (2000), where log wage is described by polynomials of time trends $T(t)$, age $A(a_{it})$ and cohort $C(c_i)$ effects and their interactions $R(a_{it}, c_i, t)$:

$$lw_{it} = \alpha + A(a_{it}) + T(t) + C(c_i) + R(a_{it}, c_i, t) + u_{it} \tag{1}$$

where α is a constant and u is an error term. The time trends capture the effects of interactions between changes in the supply and demand of the different demographic groups, which may reflect skill biased technological change, trade effects, etc. This term captures common shocks in the wage distribution within all education groups, except by age factor. The inclusion of the time trends term is the only strategy for taking into account the life cycle of differences in wage fluctuations across generations.

The age and cohort effects capture the wage changes related to the life cycle (age and experience) of workers and specific generational characteristics (different productive patterns and conditions that exist when entering the job market). The age functions measure wage distribution changes for a specific education group in a given generation, and reflect life cycle wage changes unrelated to labor market experience (one of the most important determinants of worker productivity). The cohort functions measure wage variations between generations related to different education-specific cohort attributes, in terms of unobserved ability. This factor is important once education policy or institutional labor market changes affect wage distribution. These two influences are difficult to take into account. Given the existence of an exact linear relationship between age, time and cohort effects,[5] for identification purposes we apply exclusion restrictions on the coefficients of the cohort terms. Thus, the model now includes functions of age, time trends and interactions between only these components:

$$lw_{it} = \alpha + A(a_{it}) + T(t) + R\left(a_{it}, t\right) + u_{it} \tag{2}$$

We estimate this model for 21 log wage quantiles (q), separately for the four schooling groups:[6]

$$lw_{it}^q = \alpha^q + A\left(a_{it}, ed_i\right)^q + T\left(t, ed_i\right)^q + R\left(a_{it}, t, ed_i\right)^q + u_{it}^q \tag{2'}$$

The interpretation of the components of the regression is simple: for a given quantile of the distribution, differences in the coefficients of the functions: $T(t, ed_i)^q$, $R(a_{it}, t, ed_i)^q$ and $A(a_{it}, ed_i)^q$ among education groups capture changes in the return to education and experience, and the interaction between these two attributes. For a given education group, differences among the coefficients of the functions: $T(t, ed_i)^q$, $A(a_{it}, ed_i)^q$ and $R(a_{it}, t, ed_i)^q$ across quantiles reflect changes in the intra-group wage dispersions. The estimated quantile models give us the conditional distribution of log wages. From this distribution, it is possible to recover the unconditional distribution and decompose the log wage variance by considering counterfactual exercises that explain the different effects of education on wage inequality.

---

[5]The worker's age is determined by the survey year minus the birth cohort ($i = t - c$).
[6]1st, 5th, 10th, 15th, 20th, 25th, 30th, 35th, 40th, 45th, 50th, 55th, 60th, 65th, 70th, 75th, 80th, 85th, 90th, 95th and 99th.

Thus, the decomposition of the variance consists of measuring the portions of the wage dispersion attributed to the differences of workers' productive attributes (between-group inequality) and the alterations in unobserved productive characteristics in the same group (within-group inequality):

$$Var(lw_t) = \sum_z f_{zt} Var(lw_{zt}) + \sum_z f_{zt} \left[E\left(lw_{zt}\right) - E\left(lw_t\right) \right]^2 \qquad (3)$$

where f_{zt} is the relative weight of cell z in year t; $E(lw_{zt})$ and $Var(lw_{zt})$ are the mean and variance of the log wages in cell z in year t; and $E(lw_t)$ and $Var(lw_t)$ are the mean and variance of log wages in the labor market in year t. In equation 3, the first term and the second term on the right side refer to the within-group and between-group dispersion, respectively.

The within-group variance is affected by changes in labor force composition and wage dispersion within each group of workers that have the same level of schooling and age. In turn, the composition and price effects of education and age affect inter-group or between-group variance. The composition effect of education, evaluates how changes in the educational makeup of the workforce affect wage inequality over time. To estimate this effect, we calculate the variance between groups by maintaining steady values for the wage returns to education and experience, and the age composition of the workers.

The price effect of education evaluates how changes in the variations in wages paid to workers with different qualification levels affect wage inequality over time. To estimate this effect, we calculate the inter-group component of inequality, keeping the education and age composition of the workers and the wage returns to experience fixed. To maintain a fixed value for the returns to education and experience, we attribute the trend and interaction terms of the regression to zero before predicting the log wages. In order to fix the workforce composition, we maintain the relative weights of the education and age cells at their base-year levels (1995).

Therefore, the estimation procedure is done in two steps. In the first step, we estimate the log wage equations and obtain the conditional distributions of log wages. In the second phase, we recover the unconditional distributions for each counterfactual exercise (price effect and composition effect). The procedures are described below:

First step – estimation: the models for the quantiles (2') are estimated by the means of third-order polynomials in the functions for age, time and interactions:

$$lw_i^q = \alpha_q + A_1 a_i + A_2 a_i^2 + A_3 a_i^3 + T_1 t + T_2 t^2 + T_3 t^3 + R_1 a_i t + R_2 a_i^2 t + R_3 a_i t^2 + u \quad (2")$$

The error term includes macroeconomic cyclical effects:

$$u = u_{it} + \bar{u}_t$$

These effects refer to the macroeconomic changes that occurred in a determined period (including changes in inflation, joblessness and economic activity) and are orthogonal to the age and trend effects. In other words, they do not include any trends.[7]

The models are estimated by the smoothed least absolute deviations method. This technique consists of applying a weighted least-squares estimator to the context of quantile regressions, with desirable properties in small samples (Horowitz, 1998). The coefficients are simple order statistics of each age, year and education cell. The weights are based on the variance of each estimated order statistic $(\hat{\beta}_q)$, given by $Var(\hat{\beta}_q) = \frac{q(1-q)}{N_z f(q)^2}$, where N_z is the number of observations in cell z and $f(q)$ is the density of log wages in each cell at the qth quantile. This density is estimated nonparametrically from a Gaussian kernel distribution: $f(q) = \frac{1}{N_z h} \sum_{i=1}^{N_z} \phi(\frac{lw_{iz}-\hat{\beta}_q}{h})$, in which lw_{iz} is the logarithm of the wage of each individual i in the same cell z, h is the fixed window (bandwidth) of half a standard deviation of the log wages in each cell z and $\phi(.)$ is the standard normal density function (Koenker and Portnoy, 1998).

This procedure is equivalent to choosing the vector of the parameters that minimize the quadratic form:

$$\left(\hat{\beta}_q - Z\beta_q\right) Var\left(a\hat{beta}_q\right)^{-1} \left(\hat{\beta}_q - Z\beta_q\right) \tag{4}$$

where Z is a set of linear restrictions that transform the unrestricted model (1) into a restricted model (2).[8] In our case, the restriction implies that the age, trend and (orthogonal) time dummies sufficiently explain the behavior of each estimated order statistic across cells and over time. Imposing the restrictions means estimating weighted least squares regressions of the grouped data, for each quantile and education group separately. This procedure will give us consistent estimates of β. Under the null hypothesis that the restrictions are valid, the minimized value follows a chi-square distribution with degrees of freedom equal to the number of restrictions. In order to construct the test statistics, we only need to sum up the weighted squared residuals. This involves the estimated percentiles, minus the predicted values, minus the orthogonal time dummies.

Second step – recovery of unconditional distribution: If q and q_z correspond to a determined quantile of the unconditional and conditional wage distributions, respectively, then $q = Pr(lw < lw^q)$ and $q_z = Pr(lw < lw_z^q)$, where q and q_z. The relationship between q and q_z is given by $q = \int_{z=1}^{Z} q_z \frac{N_z}{N}$, where $\frac{N_z}{N}$ is the relative size of cell z and Z is the total number of cells, which can be substituted by $q = \sum_{z=1}^{Z} q_z \frac{N_z}{N}$ if the variables defining the cells are discrete. Given a set of predicted conditional quantiles $z\beta^q$, it can be estimated to which conditional

---

[7]See MaCurdy and Mroz (1995).
[8]See Rothenberg (1971) and Chamberlain (1993).

quantile of cell z a given log wage level (lw) would correspond:

$$q_z^{lw} = \frac{1}{2}\left[\max_q\left(z\beta^q \leq lw\right) + \min_q\left(z\beta^q \geq lw\right)\right]$$

This procedure is carried out for the range of log wages observed in our data $[-4, 8]$, utilizing an equally spaced difference of 0.05 between the log wages. From this, 21 results for w^q are found corresponding to the considered quantiles that characterize the unconditional distribution, as well as this distribution's mean and variance. If $q = Pr(lw < lw^q) = G(lw^q)$ is an accumulated density function, then the empirical probability density function referring to quantile q can be written as $g(lw^q) = G(lw^q) - G(lw^{q-\varepsilon})$, where $q - \varepsilon$ is the neighboring quantile. Thus, the mean and variance of the unconditional log wage distribution are given by:

$$E(lw_t) = \sum_q lw^q g(lw^q)$$

and

$$Var(lw_t) = \sum_q \left[lw^q - E(lw_t)\right]^2 g(lw^q)$$

The unconditional distributions are then obtained separately for each year, considering each desired counterfactual exercise. The counterfactual exercise is obtained by fixing the wage returns or the workforce composition.

This methodology can be seen as an extension of the traditional approach of variance decomposition, wherein all log wage conditional distributions (beyond conditional variance) can be recovered by innumerous quantiles of non-linear functions. The main advantage of this approach is that it provides a natural way to decompose wage structure: composition effect is clearly interpreted as changes in workers' observable attributes; and differences in quantiles can be seen as an estimate of variations in the non-observed wage component. However, this technique and other decomposition methods do not allow the establishment of behavioral relationships or the estimation of structural parameters. However, this descriptive methodology is useful for quantifying the contribution of different factors – including workers' productive attributes, institutional or conjectural labor market factors and shocks – on wage distribution changes. Moreover, one convenience of using variance as an inequality measure is the potential for the decomposition of between and within-group components. This decomposition allows descriptive economic mechanisms to serve as inducers that change the course of wage inequality. These changes include a cross-generational rise in both the qualifications of workers and the labor market entry-age.

4. Results

Tables 2a through 2c present the estimation results of the 25th, median and 75th quantile regressions for the different education groups, serving as examples of

the patterns found in the other percentiles. The trend, age and interaction terms are statistically significant in most education groups. The differences in magnitude of the trend coefficients among the quantiles in an education group are evidence of changes in the wage distribution for workers with the same level of schooling and experience. The interactions of trend and age reflect changes in the returns of experience over time, which may also impact dispersion within groups. However, it is easier to grasp the information contained in the estimated models by utilizing graphs.

Figure 7 shows that changes in the variance predicted by the model closely follow changes in the actual variance, which was computed using individual data. This demonstrates the model's good fit and that the predicted variance can be used to carry out the counterfactual exercises for the different effects of education on wage inequality.

Figure 7
Fit of model (1995-2009)
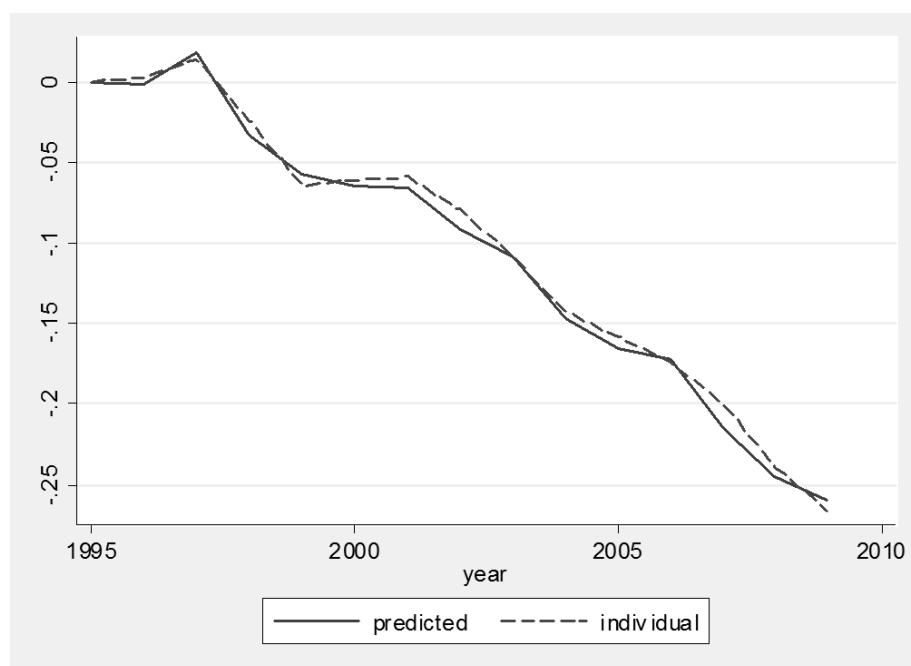
Source: 1995 to 2009 PNAD.

Note: Cumulative change in the variance of the logarithm of real hourly wages with respect to 1995.

The differences in the magnitude of age coefficients across percentiles and

Table 2a
Regression coefficients – Quantile 25

	Primary	Secondary	High School	College
Trend	-0.19*	-0.28*	-0.27*	-0.50*
	(0.06)	(0.05)	(0.05)	(0.12)
Trend^2	0.13	0.12	-0.10	0.29***
	(0.09)	(0.08)	(0.08)	(0.17)
Trend^3	0.13*	0.15*	0.21*	-0.04
	(0.04)	(0.03)	(0.03)	(0.07)
Age	0.07**	0.29*	0.45*	0.93*
	(0.03)	(0.02)	(0.02)	(0.06)
Age^2	0.00	-0.04*	-0.11*	-0.36*
	(0.02)	(0.02)	(0.02)	(0.03)
Age^3	0.00	0.00	0.01**	0.05*
	(0.00)	(0.00)	(0.00)	(0.01)
Trend*Age	0.07**	-0.11*	-0.10*	-0.10***
	(0.03)	(0.03)	(0.03)	(0.06)
Trend^2*Age	-0.01	0.02***	0.01***	0.04*
	(0.01)	(0.01)	(0.01)	(0.01)
Trend*Age^2	-0.02	0.00	0.01	-0.02
	(0.02)	(0.01)	(0.02)	(0.03)
Constant	0.20*	0.61*	1.03*	1.80*
	(0.02)	(0.01)	(0.01)	(0.03)
Chi-square test	583.57	498.56	541.82	739.16
P-value	0.00	0.45	0.07	0.00

Source: 1995 to 2009 PNAD.

Notes: *$p > 0.01$; **$p > 0.05$; ***$p > 0.10$.

Number of observations: 495 cells.

Primary: zero to three years of schooling;

Secondary: four to seven years of schooling;

High School: eight to 11 years of schooling;

College: 12 or more years of schooling.

Table 2b
Regression coefficients – Median

	Primary	Secondary	High School	College
Trend	-0.20*	-0.34*	-0.46*	-0.30*
	(0.05)	(0.05)	(0.05)	(0.10)
$Trend^2$	0.18**	-0.01	-0.07	-0.02
	(0.07)	(0.07)	(0.07)	(0.15)
$Trend^3$	0.12*	0.19*	0.23*	0.06
	(0.03)	(0.03)	(0.03)	(0.06)
Age	0.25*	0.36*	0.50*	0.98*
	(0.02)	(0.02)	(0.02)	(0.05)
Age^2	-0.05*	-0.07*	-0.12*	-0.36*
	(0.01)	(0.02)	(0.02)	(0.03)
Age^3	0.00	0.00	0.01*	0.05*
	(0.00)	(0.00)	(0.00)	(0.01)
$Trend^*Age$	-0.11*	-0.08*	-0.02	-0.08
	(0.03)	(0.03)	(0.03)	(0.06)
$Trend^{2*}Age$	0.03*	0.02*	0.01***	0.02
	(0.01)	(0.01)	(0.01)	(0.01)
$Trend^*Age^2$	-0.01	0.00	-0.04*	0.01
	(0.01)	(0.01)	(0.01)	(0.03)
Constant	0.54*	1.06*	1.52*	2.26*
	(0.01)	(0.01)	(0.01)	(0.03)
Chi-square test	402.86	556.92	500.13	621.37
P-value	1.00	0.03	0.43	0.00

Source: 1995 to 2009 PNAD.

Notes: *$p > 0.01$; **$p > 0.05$; ***$p > 0.10$.

Number of observations: 495 cells.

Primary: zero to three years of schooling;

Secondary: four to seven years of schooling;

High School: eight to 11 years of schooling;

College: 12 or more years of schooling.

Table 2c
Regression coefficients – Quantile 75

	Primary	Secondary	High School	College
Trend	-0.20*	-0.48*	-0.53*	-0.13
	(0.06)	(0.06)	(0.06)	(0.11)
Trend^2	-0.14	0.08	-0.08	-0.21
	(0.08)	(0.08)	(0.09)	(0.15)
Trend^3	0.26*	0.16*	0.21*	0.11
	(0.04)	(0.04)	(0.04)	(0.07)
Age	0.32*	0.44*	0.56*	1.02*
	(0.03)	(0.03)	(0.03)	(0.05)
Age^2	-0.06*	-0.09*	-0.11*	-0.39*
	(0.02)	(0.02)	(0.02)	(0.03)
Age^3	0.00	0.00	0.01	0.06*
	(0.00)	(0.00)	(0.00)	(0.01)
Trend*Age	-0.06	-0.10*	0.01	-0.04
	(0.03)	(0.03)	(0.03)	(0.06)
Trend^{2*}Age	0.02*	0.02*	0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Trend*Age^2	-0.01	-0.01	-0.04	0.05
	(0.02)	(0.02)	(0.02)	(0.03)
Constant	1.02*	1.47*	2.00*	2.72*
	(0.02)	(0.01)	(0.02)	(0.03)
Chi-square test	476.59	638.61	570.14	638.79
P-value	0.72	0.00	0.01	0.00

Source: 1995 to 2009 PNAD.

Notes: *$p > 0.01$; **$p > 0.05$; ***$p > 0.10$.

Number of observations: 495 cells.

Primary: zero to three years of schooling;

Secondary: four to seven years of schooling;

High School: eight to 11 years of schooling;

College: 12 or more years of schooling.

schooling groups reveal that returns to experience vary a great deal with human capital and ability. Figure 8 shows that wage inequality increases with age for all education groups, but this effect is much stronger for the less educated. This indicates that there are unobserved productivity differences amongst workers that are revealed on the job and that this heterogeneity is higher among the less educated.

Figure 8
Wage inequality over life-cycle by education
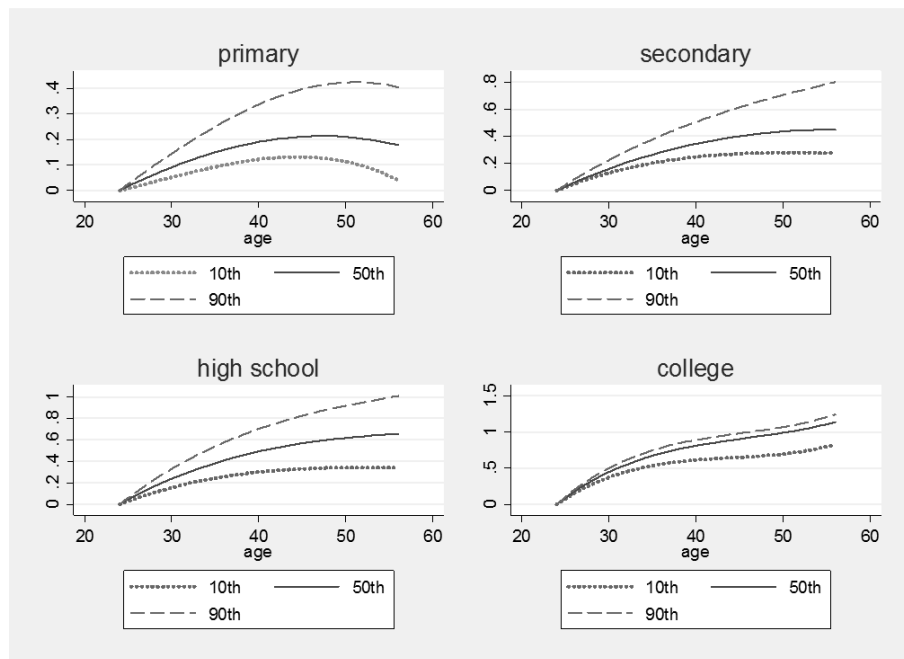
Source: 1995 to 2009 PNAD.

Note: Cumulative changes in the logarithm of real hourly wages (2005 prices) by education group with respect to 25 years of age.

Figure 9 illustrates the behavior of wages over the life cycle for the different education groups, over time. It seems that returns to experience are higher for the more educated workers, indicating that returns to specific human capital (on the job training) depend on general human capital. Over time, returns to experience have flattened out for the primary and high school educated workers, but have remained stable for the other education groups. Therefore, the decline in the returns to experience for less skilled and semi-skilled workers may also have contributed to the fall in earnings dispersion, as is documented.
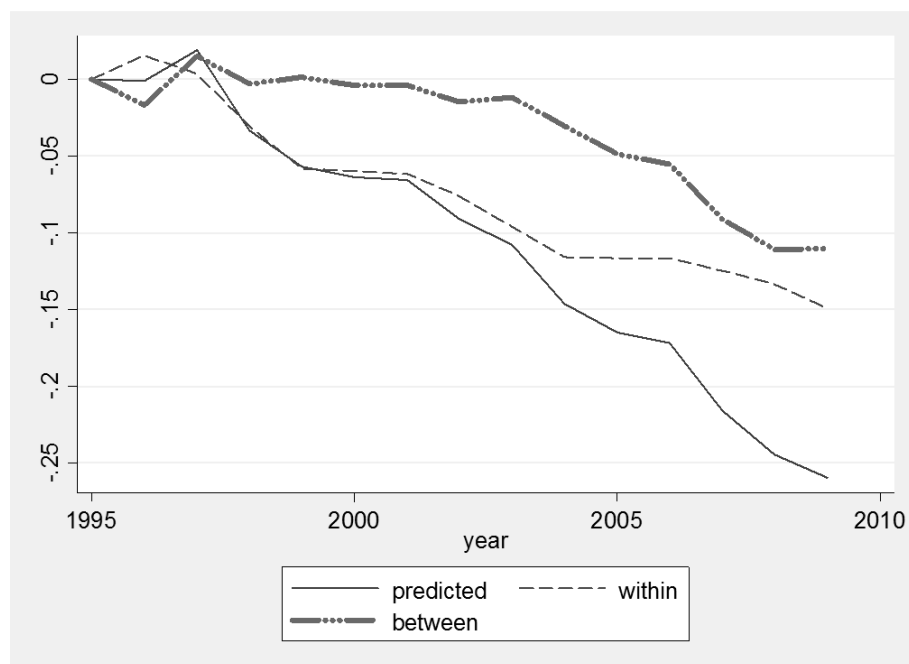
Figure 9
Age returns by education group
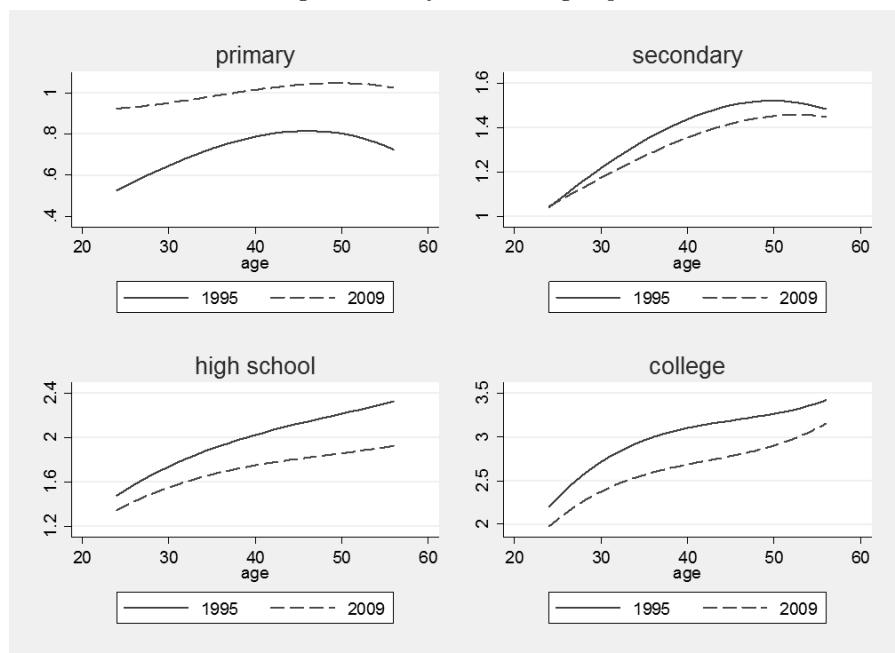
Source: 1995 to 2009 PNAD.

Note: Predicted logarithm of real hourly wages (2005 prices) by education group.

4.1 Variance decomposition

Figure 10 documents the role played by the within and between-group components of wage inequality in Brazil over the past 15 years. The behavior of overall inequality closely follows the pattern of within-group inequality until 2001, with the between-group component remaining basically constant until that same year. This means that most of the changes occur within education and age cells. However, after 2001 inequality starts to fall more rapidly than the within-group component, due to the fall in the between-group component. As we proceed, we try to disentangle the underlying effects of the behavior of this last component.

Figure 10
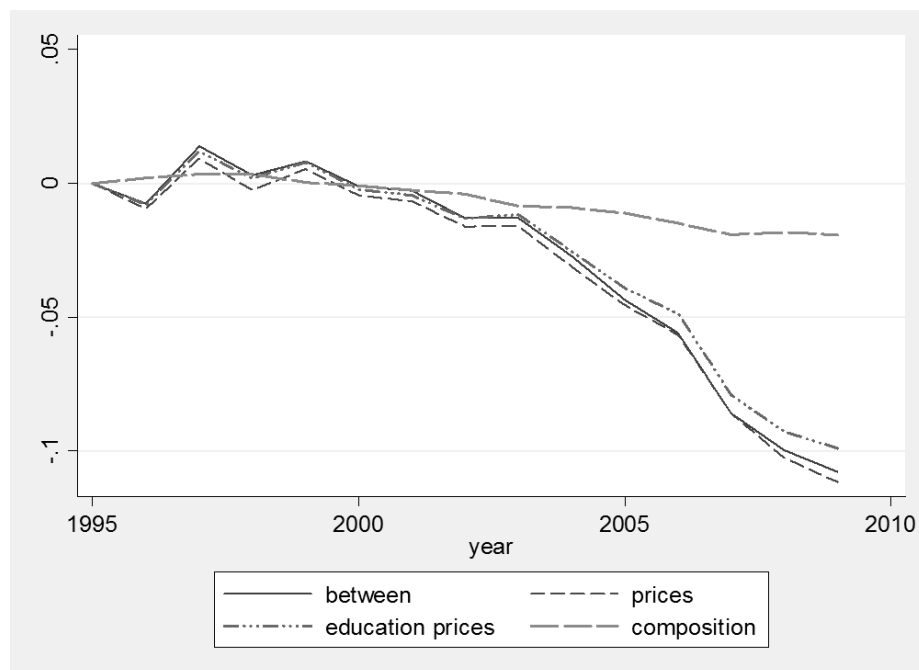Variance Decomposition (1995-2009)

Source: 1995 to 2009 PNAD.

Note: Cumulative change in the within-group and between-group components of the variance of the logarithm of real hourly wages (2005 prices) with respect to 1995.

Figure 11 decomposes the between-group component into a price and a composition effect. To calculate the composition effect we hold the wages for each percentile in each cell constant at their 1995 level and allow the cell weights to vary. To compute the price effect we hold the weights constant and allow the wage differentials to move, as described above in Section 3. While the composition effect

contributes to the decrease in inequality after 1998, most of the drop is caused by the decline in the wage differentials across groups, especially after 2001. Moreover, most of this effect reflects the decline in the education wage differentials, since holding the age differentials constant has very little impact on the behavior of the price effects. Therefore, it seems that the bulk of the decrease in inequality between groups reflects the decline in the education wage differentials.

Figure 11
Between-group variance component (1995-2009)
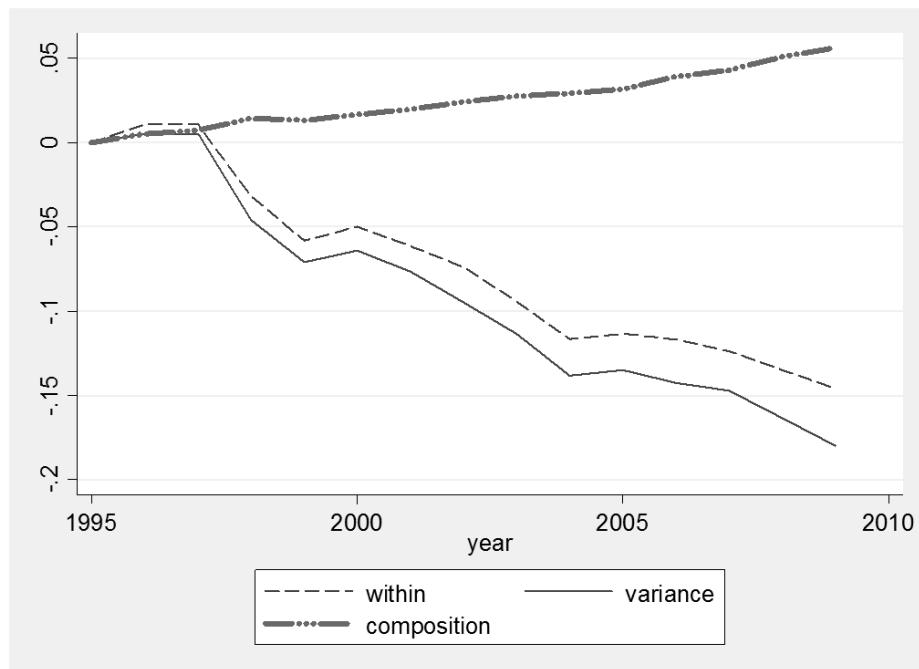

Source: 1995 to 2009 PNAD.

Note: Cumulative change in the price and composition effects of the between-groups variance of the logarithm of real hourly wages (2005 prices) with respect to 1995. Composition effect: Wage inequality due to the changes in education and labor force age compositions (wage returns fixed at 1995 level). Price effect: Wage inequality due to changes in wage returns (education and age compositions fixed at 1995 levels).

4.2 Within-group variance

Figure 12 plots the behavior of the within-group component over time. It is clear that the behavior of this component reflects two forces acting in opposite

directions. The "pure" within-group effect has shaped the overall declining trend of inequality over time, but the composition effect (also called "mechanical" by Lemieux, 2006a) contributed to a continuous rise in inequality, since more educated and older groups are more unequal. As a result, within-group inequality fell less rapidly then it would have otherwise.

Figure 12
Within-group variance component (1995-2009)

Source: 1995 to 2009 PNAD.

Note: Cumulative change in the within-group components of the variance of the logarithm of real hourly wages (2005 prices) in relation to 1995. Composition effect: Intra-group wage inequality due to changes in the education and age compositions of the labor force (intra-cell variances fixed at 1995 levels). Intra-cell variance effect: Intra-group wage inequality due to changes in the intra-cell variances of education (education and age compositions fixed at 1995 levels).

What other factors could explain the behavior of the within-group inequality? Aside from human capital, our regressions do not allow inferences about other forces that can affect within-group variance. Nevertheless, we can make some speculations about economic factors that might have affected wage inequality within groups of workers with the same level of schooling in this period. One

possible explanation is the increase in the real value of the minimum monthly wage that took place between 1995 and 2009. Figure 13 shows that the minimum wage almost doubled in our sample period, at the same time that inequality was falling substantially within groups. Future research assessing the effects of minimum wage policy on earnings inequality is needed to investigate this possibility in more detail.

Figure 13
Minimum wage and wage inequality (1995-2009)

Source: 1995 to 2009 PNAD.

Note: Real minimum wage expressed in Brazilian currency (Real, R$).

Table 3 summarizes our main results by presenting the contribution of each component to the reduction of inequality, year by year. The table shows that the variance of wages declined by 0.25 between 1995 and 2009, from an initial value of 0.93. This represents a reduction of 26%. Moreover, the reduction of the education and age wage differentials accounts for about 44% of the total drop. Changes in the education and age composition of the workforce explain about eight percent of the change in inequality, as the new generations of workers become increasingly more educated. The highest contributing factor within groups is the price effect, which is in the range of 70%. Finally, had all of the other forces remained constant,

the higher human capital of the workforce would have contributed to an increase in the overall variance of earnings by about 22%. This results from the fact that inequality is higher among the more educated and experienced workers, as is noted above (the mechanical effect).

Table 3
Regression Contribution of each component to the reduction of inequality

Year	Variance changes	Between component (prices) %	Between component (composition) %	Within component (prices) %	Within component (composition) %
1996	0,01	-3,00	0,60	1,65	1,63
1997	0,02	0,38	0,14	0,20	0,29
1998	-0,03	0,09	-0,12	1,60	-0,50
1999	-0,05	-0,11	-0,01	1,42	-0,26
2000	-0,05	0,09	0,02	1,25	-0,32
2001	-0,06	0,10	0,04	1,19	-0,31
2002	-0,09	0,19	0,05	1,09	-0,28
2003	-0,11	0,15	0,08	1,06	-0,26
2004	-0,14	0,22	0,06	0,96	-0,20
2005	-0,16	0,29	0,07	0,86	-0,20
2006	-0,17	0,33	0,09	0,83	-0,23
2007	-0,21	0,41	0,09	0,70	-0,20
2008	-0,23	0,44	0,08	0,70	-0,22
2009	-0,25	0,44	0,08	0,71	-0,22

Source: 1995 to 2009 PNAD.

5. Conclusions

This paper evaluates the factors that have contributed to the decline in earnings inequality in Brazil, an occurrence happening for the first time in decades, by means of a flexible decomposition technique and counterfactual exercises. The variance in (log) earnings declined by about a quarter between 1995 and 2009. We find that, until the end of the 1990s, most of the decrease happened within education and age groups, with our observable measures of skill serving a very small role. However, in the new century, the between-group component also contributed significantly to the drop in inequality, mostly as a result of the fall in the education wage differentials. Returns to experience have also declined, especially among the less skilled workers.

We find that the education composition of the workforce also contributed to the decline in inequality between groups, but increased the within-group dispersion. Overall, the results indicate the powerful impact that education can have on reducing earnings inequality.

References

Autor, D. H., Katz, L. F., & Kearney, M. S. (2005). Rising wage inequality: The role of composition and prices. Working Paper 11628, NBER.

Binelli, C., Meghir, C., & Menezes-Filho, N. (2008). Education and wages in Brazil. University College London (mimeograph).

Chamberlain, G. (1993). Quantile regressions, censoring and the structure of wages. *Advances in Econometrics*, 1:171–209.

Corseuil, C. H. & Foguel, M. N. (2002). Uma sugestão de deflatores para rendas obtidas a partir de algumas pesquisas domiciliares do IBGE. Discussion Paper 897, IPEA.

Ferreira, F., Leite, P., & Litchfield, J. (2008). The rise and fall of Brazilian inequality: 1981 – 2004. *Macroeconomic Dynamics*, 12:199–230.

Gosling, A., Machin, S., & Meghir, C. (2000). The changing distribution of male wages in the UK. *The Review of Economic Studies*, 67:635–66.

Horowitz, J. L. (1998). Bootstrap methods for median regression models. *Econometrica*, 66:1327–1351.

IBGE (1995). Pesquisa Nacional por Amostra de Domicílios. Brazilian Geography and Statistics Institute.

IMF (2008). World Economic Outlook Database. International Monetary Fund.

IPEADATA (2009). Brazilian Applied Economics Research Institute.

Knight, J. & Salbot, R. (1983). Education expansion and the Kuznets effect. *American Economic Review*, 73:1132–36.

Koenker, R. & Portnoy, S. (1998). Quantile regressions. Technical Report 97-0100, University of Illinois.

Lam, D. & Levinson, D. (1992). Declining inequality in schooling in Brazil and its effects on inequality in earnings. *Journal of Development Economics*, 37:199–225.

Lemieux, T. (2006a). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, 96:461–98.

Lemieux, T. (2006b). Post-secondary education and increasing wage inequality. *American Economic Review*, 96:1–23.

Lopez-Calva, L. & Lustig, N. (2010). *Declining Inequality in Latin America: A Decade of Progress.* UNDP, New York.

MaCurdy, T. & Mroz, T. (1995). Estimating macro effects from repeated cross-sections. Discussion paper, Stanford University.

Manacorda, M., Sanchez-Paramo, C., & Schady, N. (2010). Changes in returns to education in Latin America: The role of demand and supply of skills. *Industrial and Labor Relations Review*, 63:307–326.

Menezes-Filho, N., Fernandes, R., & Picchetti, P. (2006). Rising human capital but constant inequality: The education composition effect in Brazil. *Revista Brasileira de Economia*, 60:407–24.

Ram, R. (1990). Education expansion and schooling inequality: International evidence and some implications. *Review of Economics and Statistics*, 72:266–73.

Rothenberg, T. (1971). *Efficient Estimation with a Priori Information.* Yale University Press, New Haven.

UNDP (2010). Human Development Report. United Nations Development Programme, New York.