The World Bank Economic Review, 0(0), 2021, 1–31
doi: 10.1093/wber/lhab005
Article

OXFORD

Labor Market Experience and Falling Earnings Inequality in Brazil: 1995–2012

Francisco H. G. Ferreira, Sergio P. Firpo, and Julián Messina

Abstract

The Gini coefficient of labor earnings in Brazil fell by nearly a fifth between 1995 and 2012, from 0.50 to 0.41. The decline in other measures of earnings inequality was even larger, with the 90-10 percentile ratio falling by almost 40 percent. Applying micro-econometric decomposition techniques, this study parses out the proximate determinants of this substantial reduction in earnings inequality. Although a falling education premium did play a role, in line with received wisdom, this study finds that a reduction in the returns to labor market experience was a much more important factor driving lower wage disparities. It accounted for 53 percent of the observed decline in the Gini index during the period. Reductions in horizontal inequalities – the gender, race, regional and urban-rural wage gaps, conditional on human capital and institutional variables – also contributed. Two main factors operated against the decline: a greater disparity in wage premia to different sectors of economic activity, and the "paradox of progress": the mechanical inequality-increasing effect of a more educated labor force when returns to education are convex.

JEL classification: D31, J31

Keywords: earnings inequality, Brazil, returns to experience

1. Introduction

Rising income inequality has attracted a great deal of recent attention in both academic and policy circles.[1] The sustained rise in wage and income inequality in the United States since 1980 has garnered particular interest, with a large literature debating the relative importance of various contributing factors, such as changes in technology; the role of international trade; changes in the relative supply of skills; and changes in labor market institutions, such as falling unionization rates and real minimum wages. Inequality has also been rising in many other countries, both rich and poor (Alvaredo and Gasparini 2015; Morelli et al. 2015), often leading to a perception that it is rising everywhere. The World Inequality Report 2018, for

Franciso H. G. Ferreira (f.d.ferreira@lse.ac.uk) is at the International Inequalities Institute, London School of Economics, UK, and IZA, Bonn, Germany. Sergio P. Firpo (firpo@insper.edu.br) is at the Insper Institute of Education and Research, São Paulo, Brazil, and IZA, Bonn, Germany. Julián Messina (julianm@iadb.org) is at the Inter-American Development Bank, Washington, DC, and IZA, Bonn, Germany. The authors are grateful to Camila Galindo, Elizabeth Jaramillo, Nicolás Guida Johnson, and Juan Pablo Uribe for excellent research assistance. They also thank the editor, Eric Edmonds; two anonymous referees; Edmund Amann; Armando Barrientos; Dean Jolliffe; Nora Lustig; Ben Ross Schneider; and seminar participants at MIT, Brown University, George Washington University, Tulane University, the World Bank, the University of Cape Town, King's College London, and USP-Ribeirão Preto, as well as at various conferences, for comments on earlier versions. All errors are our own. A supplementary online appendix is available with this article at *The World Bank Economic Review* website.

1 See, e.g., Stiglitz (2012) and Piketty (2014).

example, claims that "income inequality has increased in nearly all world regions in recent decades, but at different speeds" (Alvaredo et al. 2018, 6).

But inequality can also fall. A recent World Bank study found that income inequality declined (by at least one Gini point) in 39 of 91 countries for which data were available in the 1993–2008 period (World Bank 2016). Most of these reductions took place in emerging and developing countries: inequality fell in 10 African and 11 Latin American countries, for example. Understanding the nature and drivers of such inequality reductions would seem to be as important as studying the increases, yet much less scholarly attention has been paid to them, even where rich data constellations are available.

This paper investigates the proximate determinants of a substantial decline in the inequality of labor incomes in Brazil – the largest economy in Latin America, and one of the 10 largest in the world – during the 1995–2012 period.[2] Starting from very high levels, the Gini coefficient of the country's distribution of household per capita income fell by 12 percent, from 0.59 in 1995 to 0.52 in 2012: a seven-point reduction. At least half of that decline has been attributed to changes in the distribution of labor earnings (Barros et al. 2010; Azevedo, Inchauste, and Sanfelice 2013), and the Gini coefficient for that distribution fell by 18 percent in the same period, from 0.50 to 0.41.

By any recorded historical standards, this is a substantial fall in wage inequality.[3] What explains it? The dominant view in the literature on income disparities in Latin America attributes the bulk of the equalization during the 2000s to falling returns to schooling (e.g., López-Calva and Lustig 2010; Gasparini et al. 2011). The same is true for studies focusing specifically on Brazil (e.g., Barros et al. 2010). This view casts the Brazilian inequality trajectory as a "reverse Goldin and Katz (2008) story," in which the main roles are played by education and its labor market returns.[4]

Yet, during this period the Brazilian labor market underwent other rapid transformations as well, with potential implications for inequality. Formal employment grew by 9 percentage points (Maurizio 2014). Gender, racial, and experience wage gaps declined (Ñopo 2012; Messina and Silva 2018). And relative wages changed across sectors, urban-rural locations, and regions, either triggered by long-lasting effects of trade liberalization (Dix-Carneiro and Kovak 2017), the commodity boom (Adão 2015), or the emergence of China (Costa, Garred, and Pessoa 2016). Little is known about the contribution of these each of these factors to the reduction in wage inequality.

This study applies a decomposition method based on recentered influence function (RIF) regressions (Firpo, Fortin, and Lemieux 2009) to decompose the observed decline in inequality into shares due to each of these (and other) factors. RIF regressions are designed to enable detailed decompositions that parse out the structure and composition effects of observed covariates, identified under the standard assumptions discussed by Fortin, Lemieux, and Firpo (2011).[5] The analysis suggests that, although Brazil did experience a large increase in the supply of education, which contributed to a decline in the returns to schooling (López-Calva and Lustig 2010), these falling returns were not the quantitatively most important driver of falling inequality.

Instead, the bulk of the decline in wage inequality over the 1995–2012 period can be accounted for by two other factors. The first, and most important, is a marked reduction in the returns to labor market

2 During this period, average real earnings declined by 12 percent between 1995 and 2003, then rose by 38 percent between 2003 and 2012. See section 6 for details.

3 For comparison, the average decline in income inequality across 15 Latin American countries between 2000 and 2010 was 3.5 Gini points (Ferreira et al. 2013). In the United States, the total *increase* in the Gini coefficient of household income in the four and a half decades between 1967 and 2011 was 8 points, from 0.40 to 0.48 (Jacobson and Occhino 2012).

4 More recently, others have attributed part of the decline to the direct and indirect effects of a rising real minimum wage policy (Engbom and Moser 2017; Alvarez et al. 2018).

5 The terms "structure and composition effects" are now more commonly used, but they are closely related to the older terminology of "price and endowment effects" used in the literature that followed Oaxaca (1973) and Blinder (1973).

experience, an effect consistent with the age-biased technical change hypothesis of Behaghel and Greenan (2010).[6] In our preferred specification, falling returns to experience account for 4.7 of the 9.0 Gini points decline in the full period. Surprisingly, this quantitatively dominant effect has been missing entirely from the discussion in the existing literature. The second factor is a reduction in wage gaps associated with the race, gender, and spatial location of workers, conditional on workers' observed human capital.[7] The schooling premium did also decline but, when the other factors are properly accounted for, its quantitative importance during 1995–2012 is small and statistically insignificant.

Another two factors acted to partly offset the decline in wage inequality: first, in the presence of convex returns to schooling (which were observed in this study), a broad increase in the educational attainment of the labor force has a mechanical inequality-increasing effect, sometimes referred to as the "paradox of progress" (Bourguignon, Ferreira, and Lustig 2005). On its own this effect contributed a 3.2-point increase in the Gini coefficient for wages over the period. Second, there was an increase in wage gaps among different sectors of economic activity, conditional on other worker characteristics – particularly in the upper half of the wage distribution.

The paper is organized as follows. The next section contains a very brief review of the literature on the decline in Brazil's inequality since the mid-1990s. Section 3 describes the data sets and provides some descriptive statistics for the main variables used in the subsequent analysis. The empirical approach is described in section 4, and section 5 presents the results, including a robustness section. Section 6 concludes.

2. A Brief Review of the Literature

Ferreira, Leite, and Litchfield (2008) and Barros et al. (2010) were two of the first papers to document a decline in household income inequality in Brazil in the late 1990s and 2000s. Though they used completely different decomposition techniques, they broadly agreed on the main mechanisms driving the decline. Ferreira et al. (2008) performed a set of static and dynamic decompositions of scalar inequality indices for the 1981–2004 period, in the spirit of Cowell and Jenkins (1995) and Mookherjee and Shorrocks (1982). They found that increases in the quantity of education, given convex returns, accounted for much of the rise in inequality between 1981 and 1993, whereas the decline between 1994 and 2004 was accounted for by three main factors: (i) declining returns to education, (ii) income convergence between rural and urban areas, and (iii) increases in social assistance transfers targeted to the poor.

Barros et al. (2010) developed a microsimulation–based decomposition of an identity relating household per capita income to its labor and non-labor components, demographic characteristics, participation rates, and (linearized) returns to skill. They used this technique to study inequality trends between 2001 and 2007 and concluded that "... the decline resulted from three main factors: an increase in contributory and noncontributory government transfers; a decline in wage differentials by educational level ...; and an improvement in spatial and sectoral integration of labor markets, in particular among metropolitan and non-metropolitan areas." Although the authors considered the quantity and returns to schooling explicitly within their decomposition framework, they do not examine the role of labor market experience at all: it does not feature in their decomposition.

Barros et al. (2010) also estimate that 40–50 percent of the reduction in inequality in household per capita incomes arose from changes in the distribution of non-labor incomes, mostly public transfers. Using similar techniques, Azevedo et al. (2013) estimate a broadly similar share, arising primarily from increases in pension incomes, and both higher and better-targeted social assistance transfers. Both papers attribute another 10 percent or so to demographic factors, chiefly the rapid decline in family sizes (and thus in dependency ratios), which was most pronounced among poorer households. The remaining 50 percent or

6 See also Friedberg (2003) and De Koning and Gelderblom (2004).

7 For lack of a better term, this study refers to these closing gaps as reductions in "horizontal inequality".

so of the decline in inequality in household incomes is attributed to changes in the distribution of labor earnings, which are the focus of this paper. And in terms of that distribution, the early studies pointed unambiguously to falling returns to education as the key driver of inequality reduction.

This view that falling returns to education were central to the decline in earnings inequality was not specific to Brazil. Similar assessments were made for various other Latin American countries over the same period. A comprehensive discussion of the trends in Argentina, Brazil, Mexico, and Peru can be found in López-Calva and Lustig (2010), who note that the returns to tertiary relative to primary education declined in all four countries. They go on to argue that "... the data suggest that while during the 1990s the demand for skills dominated ..., in the last ten years the growth in the supply of skills outpaced demand and the college premium consequently shrank. Or, to use Tinbergen's language, in the race between skill-biased technological change and educational upgrading, the latter took the lead" (p. 13). Acosta et al. (2019) and Lustig, López-Calva, and Ortiz-Juárez (2013) reach broadly similar conclusions.

This became the "conventional view" on the determinants of the decline in earnings inequality in Brazil (and much of Latin America) over the last decade, which is still echoed by some more recent papers. Looking at wages for male workers in Brazil between 1995 and 2012, for example, Wang, Lustig, and Bartalotti (2016, 49) also find that the wage structure effect with respect to schooling was the key driver of lower wage inequality. Their abstract states: "Education expansion had an equalizing impact on wage distribution, primarily through the decline in return to education." [8]

Jaume (2017) sheds light on a possible mechanism: he applies a variant of the Acemoglu and Autor (2011) model of production using tasks of varying complexity to study the effect of an educational expansion in the context of the Brazilian labor market. He finds that the increases in education can help explain the decline in wage inequality, through a process of educational downgrading, where workers with certain educational levels become more likely to perform simpler tasks, as time goes by and average education levels rise.

Overall, the bulk of the existing literature suggests that declining returns to education were the main driver of the recent reduction in Brazilian wage inequality. Falling spatial gaps and a rising minimum wage also feature as likely contributors. The role of declining returns to labor market experience is notable for its absence from the discussion.

3. Data and Descriptive Statistics

The present study relies on two complementary microdata sources on labor market earnings during the 1995–2012 period. The primary data set used for most of the analysis is the *Pesquisa Nacional por Amostra de Domicílios* (PNAD), an annual, nationally representative household survey, covering both rural and urban areas.[9] It was fielded by the Brazilian Census Bureau (*Instituto Brasileiro de Geografia e Estatística*, IBGE) every year, except for census years. To construct inequality trends, all available years in the period 1995–2012 are used, but for the decompositions the present study uses the PNAD datasets for six years: 1995, 1996, 2002, 2003, 2011, and 2012.

8 In their regression-based analysis, Wang, Lustig, and Bartalotti (2016) control for a quadratic on experience, and their table 7 shows large negative effects of returns to experience on various measures of inequality. But these effects are not discussed in the text at all.

9 Except for the rural areas of Acre, Amapá, Amazonas, Pará, Rondônia, and Roraima states, which correspond to the Amazon rain forest. These areas, which according to census data account for 2.3 percent of the Brazilian population, were excluded from the survey before 2003. To preserve sample comparability, the study excludes these areas from the sample in all years.

The PNAD is a long-standing household survey, which was fielded in comparable form between 1976 and 2015.[10] The survey was always fielded in the last quarter of the year, in 1100 municipalities representative of the entire country. The survey's question on labor earnings has remained identical through the period of analysis. For employees, this refers to the gross monthly monetary salaries the worker would normally earn in September from his or her main occupation. There is a separate question for income in kind from the main occupation, and, if this is non-zero, it is added to monetary salaries.

The corresponding question for self-employed workers asks for "net monthly withdrawals": that is, gross income from self-employment minus expenses incurred in the job or business. When the remuneration is variable, workers are asked to report their average monthly earnings (or withdrawal). Thus, monthly earnings exclude one-off bonuses, the "thirteenth payment" (Christmas bonus), and irregular in-kind payments. This study's main measure of earnings is total monthly earnings in the main job, and it is expressed in real values using the CPI deflator with base-year 2005 – this is our dependent variable Y.[11] If the worker has more than one job, the main job is defined in the following order: (i) the job held for the longest period in the preceding year; (ii) paid work has priority over unpaid work; (iii) the job with the greatest number of weekly hours normally worked; (iv) the higher yield normally received.

The following filters were applied to the data to generate the working sample. The sample includes all workers aged 18–65 who reported positive earnings during the survey's reference month.[12] Monthly earnings are trimmed at the 1st and 99th percentiles to eliminate outliers. Sample sizes vary somewhat by year, from a minimum of 105,566 workers in 1996 to a maximum of 136,963 in 2012.

The data also include information on various earning covariates. All such variables used in the analysis are categorical, except for schooling (Ed) and experience (Exp) in the labor market, which are measured in years.[13] An important caveat is that the PNAD data used are not a panel, and the survey does not ask the kind of retrospective questions on job transitions that would make it possible to construct reliable measures of actual labor market experience. Therefore, the study uses potential experience, defined as the worker's age minus years of schooling minus 6, in the main analysis. The analysis is complemented using actual experience data from the RAIS dataset, which is described below.

Demographic worker characteristics include a gender dummy (G) and a three-way categorical variable for race, R (white, black, and other).[14] The analysis also distinguishes between rural and urban workers, through a dummy variable (U). Rural and urban areas are classified in accordance with the Brazilian census definitions. Spatial variables also include dummies for the five main geographic regions of Brazil: North, Northeast, Center-West, Southeast and South, grouped together as the variable set *Region*. Finally, regarding sectoral distribution, workers are divided into 17 different sectors of economic activity (see table 2). These sector dummies are grouped together in the set *Sector*.

A three-way categorical variable for formality (F) is defined as follows: Workers are classified as formal employees if they have a job that is properly registered in their work-card or "*carteira de trabalho*," which provides workers with various social security benefits, including rights to pensions, unemployment insurance, and severance payments. Workers whose employers have not registered their job in the "*carteira*

10 From 2016 on, the PNAD was replaced by the PNAD Contínua, which is fielded on a rolling basis across all 12 months of the year. IBGE does not recommend comparisons across the two series (which overlap between 2012 and 2015), because of methodological differences. This is, of course, outside the study period of this article.

11 Results for hourly wages are qualitatively similar and are reported in the supplementary online appendix.

12 The composition of the sample of workers with positive earnings naturally changes over time. The effect of such changes, at least insofar as they reflect changes in the joint distribution of observable characteristics, is captured by the composition effects of the decomposition method described in section 4.

13 As discussed below, RIF regression–based decompositions perform better when not all variables are categorical.

14 The "black" category combines the "*pretos*" and "*pardos*" response items from the PNAD, as per increasingly common practice. The "other" category combines indigenous people and those reporting "yellow" as their skin-color: a group largely consisting of Asian immigrants and their descendants.

de trabalho" are considered informal employees. The third category consists of own-account or self-employed workers ("*conta própria*"). The vector of observed covariates used in our analysis is thus given by eq. (1):

$$X = \{Ed,\ Exp,\ F,\ G,\ R,\ U,\ Region,\ Sector\} \tag{1}$$

The PNAD is our primary dataset because it is a representative sample of all Brazilian workers, including formal and informal employees, as well as the self-employed. The informal sector (informal employees and the self-employed) comprised 52 percent of Brazilian workers at the beginning of the study period, and 43 percent at the end. Any study seeking to analyze changes in the earnings distribution across the whole country must therefore include these workers. The description of trends in earnings and other variables in the next section is therefore based primarily on the PNAD dataset.

Nevertheless, because of the aforementioned lack of information on actual experience in the PNAD, the study also uses data from the *Relação Anual de Informações Sociais* (RAIS), a comprehensive administrative dataset maintained by Brazil's *Ministério do Trabalho e Emprego* (Labor and Employment Ministry). The RAIS data have the drawback that they only cover workers with a signed labor card: that is, formal sector employees. On the other hand, their advantages include the fact that it is a long-term panel that follows the universe of workers over their entire formal-sector careers. Because wage information is directly provided by employers to the ministry of labor and is linked to a number of employee benefits, it is generally considered to be more accurate than wage data in PNAD. With the exception of race, which is observed in RAIS with considerable noise and hence is excluded from the analysis, variable definitions are identical to those from the PNAD, with an additional variable for actual experience (*Exp_act*), which corresponds to the sum of years each worker was matched to a firm in the RAIS data.[15]

The RAIS data are used to analyze the evolution of the returns to actual work experience during the period of analysis, 1995–2013. To construct actual work experience, the study builds the entire work history of each worker in the formal sector. RAIS does not contain retrospective information of previous work spells. Thus, if a worker is first observed at, say, age 40 in the first available wave of RAIS, it is not possible to construct her work history prior to that year to compute her actual years of experience. To maximize the coverage of actual experience, the sample is extended as far back as possible, to 1986. This makes it possible to build the entire work histories of workers who were at most 18 years old in 1986. Thus, the maximum actual work experience that is observed is 28 years in 2013.

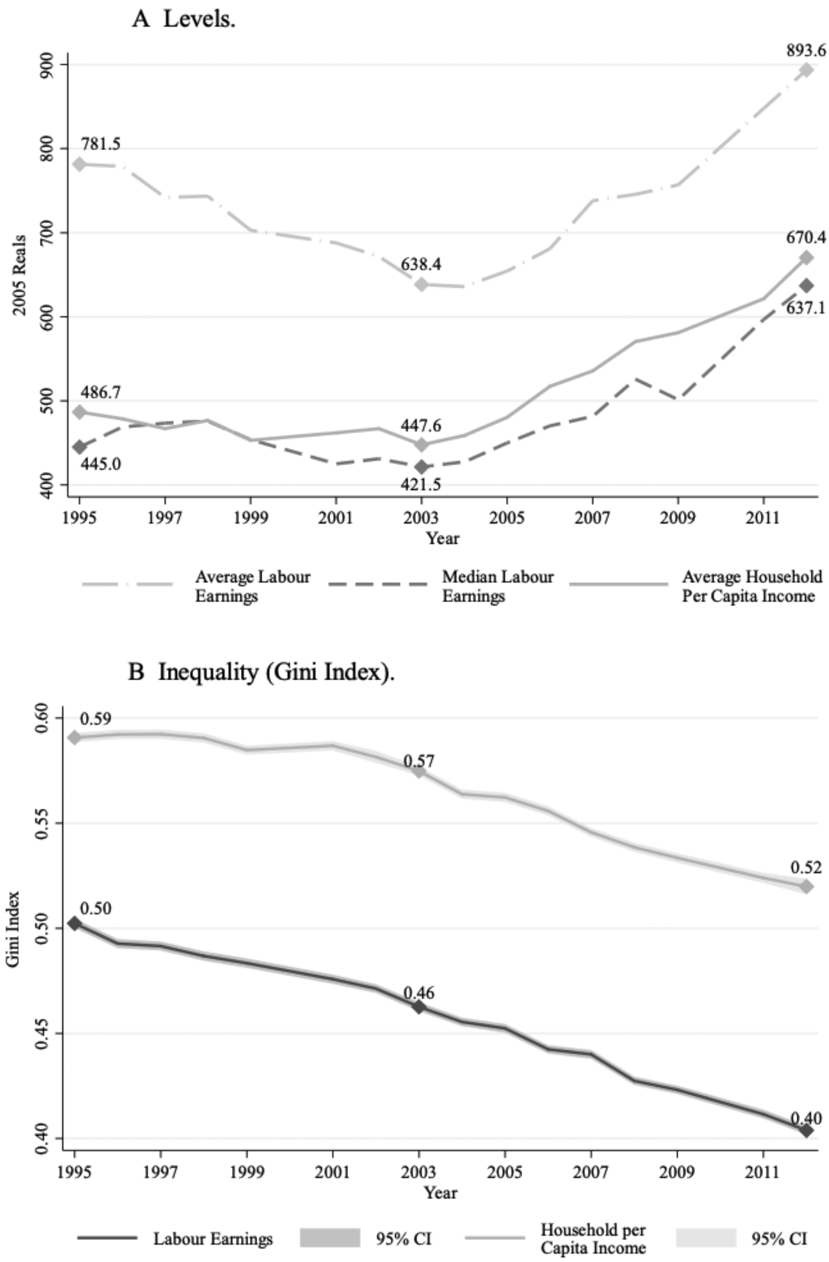
The study works with a 3 percent random sample of all worker identifiers observed in RAIS during 1986–2013. The main measure of earnings is the monthly earnings in December, and it is expressed in real values using the CPI deflator with base-year 2005. When a worker has more than one job at a point in time, the study keeps the highest-paying job. To preserve the symmetry with the treatment that is given to monthly earnings in PNAD, the 1st and 99th percentiles are trimmed each year. Overall, entire work histories are computed for a total of 3,546,849 workers that are observed, on average, for 9 years.

Descriptive Statistics and Basic Trends

This subsection provides a preliminary description of the changes in earnings (both levels and inequality) and in the vector of covariates *X* during the study period: 1995 to 2012. Summary statistics refer to the study sample constructed from the PNAD data, as described in this subsection. Panel A in fig. 1 shows mean and median labor earnings in levels, as well as mean household per capita income (HPCY) for comparison. From 1995 to 2003, both earnings and household incomes were either stable or declining: median labor earnings and average HPCY were roughly constant, while mean labor earnings fell by 18 percent. This changed around 2002–2003, when all three series began to trend sharply upwards. Average

15 By its nature, the RAIS data contains no information on any time workers might have spent in informal sector work or self-employment.

Figure 1. Inequality and Level Trends of Household Incomes and Labor Earnings in Brazil, 1995–2012

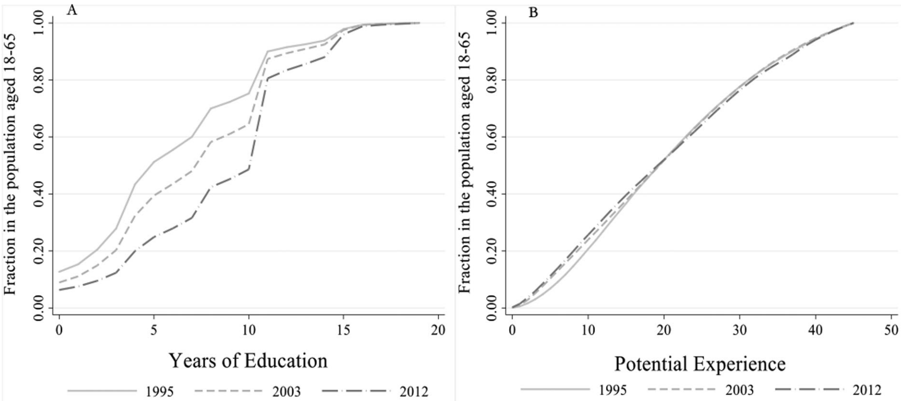


Source: *Pesquisa Nacional por Amostra de Domicílios* (PNAD).
Note: All measures are calculated over the estimating sample (formal, informal, and self-employed of ages 18–65). Negative incomes and the 99th and 1st percentiles of earnings are trimmed. Labor earnings refer to monthly earnings reported in the main occupation. The household per capita income includes all the incomes perceived by the household members.

labor earnings, for example, increased by about 40 percent in real terms from 2003 to 2012. Median earnings and HPCY also grew rapidly in this second subperiod.[16] Because of this important inflection

16 The causes of this inflection and of the boom decade of 2002–2013 go beyond the scope of this paper: 2002 was the year in which President Luiz Inácio Lula da Silva took office. It is also now commonly viewed as the beginning of the

Figure 2. Schooling and Experience Cumulative Distribution Functions by Year. (a) Panel A. Education. (b) Panel B. Potential Experience



Source: *Pesquisa Nacional por Amostra de Domicílios* (PNAD).

Note: The figures show the cumulative distribution functions (CDF) of years of education (panel a) and potential experience (panel b), which is defined as age-years of education-6. The CDFs are calculated for all the individuals over the estimating sample (formal, informal, and self-employed of ages 18–65). The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 includes 2011 and 2012 PNAD samples.

in income and earnings levels around 2002–2003, all of the analysis are conducted separately for the 1995–2003 and 2003–2012 subperiods, as well as for the entire period.

Panel B of fig. 1 shows the point estimates and bootstrapped 95 percent confidence intervals for the Gini coefficients of HPCY and of labor earnings. Over the 17-year study period, HPCY inequality fell by about 12 percent and earnings inequality by as much as 18 percent, when both are measured by the Gini coefficient. The decline is not specific to the Gini, though: in fact, the reductions are proportionally larger when measured by the Theil index and by the 90-10 percentile ratio, at 34 percent and 37 percent, respectively. Upper-tail inequality (measured by the 90-50 percentile ratio) declined steadily during 1995–2012. Instead, lower-tail inequality (measured by the 50-10 percentile ratio) increased during 1995–2003, to sharply decline thereafter.[17]

Turning to the covariates, the study looks first at trends in schooling and experience in the labor force. Figure 2 (panel A) shows the empirical cumulative distribution functions (CDFs) for years of schooling in the working-age population (18–65), at three points in time: 1995, 2003, and 2012.[18] These distribution functions illustrate an impressive expansion in the supply of years of schooling in the labor force. The proportion of the working-age population with at least 10 years of schooling, for example, doubled from 25 percent to 50 percent between 1995 and 2012.[19]

Figure 2 (panel B) plots the empirical CDFs for potential experience, defined as age minus years of schooling minus six, for each worker. In contrast to the education CDFs, these are remarkably stable.

commodity price super-cycle, which benefited all commodity-exporting countries in Latin America, including Brazil. Messina and Silva (2019) discuss possible mechanisms through which the super-cycle may have affected the proximate drivers of earnings inequality.

17 These series are not shown but are available from the authors on request.

18 Throughout the rest of the paper, two pairs of surveys are pooled for each of these points in time to increase precision in the decompositions and descriptive statistics. Thus, the study pools 1995–1996, 2002–2003, and 2011–2012, although they are referred to as 1995, 2003, and 2012 to simplify the exposition.

19 This increase in the supply of educated workers reflects educational policy changes dating back to the late 1980s, but also the subsequent decentralization of basic education funding from the state level to the municipal level, as well as changes in the funding system with the creation of FUNDEB (*Fundo Nacional para o Desenvolvimento da Educação Básica*) to reallocate funding according to demand (see Cruz and Rocha 2018).

Table 1. Summary Statistics: PNAD Samples

	Mean		
	1995	2003	2012
Earnings (Y)			
Earnings in 2005 reals	780.12	655.01	870.97
log(Earnings)	6.22	6.10	6.48
Years of education (Ed)	6.31	7.41	8.92
Years of potential experience (Exp)	23.07	22.51	22.32
Formality (F)			
Self-employment	0.27	0.25	0.21
Informal	0.25	0.27	0.21
Formal	0.48	0.48	0.57
Race (R)			
White	0.56	0.54	0.48
Black	0.43	0.45	0.51
Other	0.01	0.01	0.01
Gender (G): Female	0.38	0.40	0.42
Urban-Rural (U): Rural	0.16	0.12	0.10
Region			
Northeast	0.25	0.24	0.24
North	0.04	0.06	0.06
Southeast	0.47	0.46	0.46
South	0.16	0.16	0.16
Center-West	0.07	0.08	0.08
Sector			
Agriculture, Fishing, and Mining	0.15	0.12	0.09
Industry	0.15	0.16	0.14
Construction	0.07	0.08	0.09
Services	0.62	0.64	0.67

Source: Pesquisa Nacional por Amostra de Domicílios (PNAD).
Note: All the statistics are calculated over the estimating sample (formal, informal, and self–employed of ages 18–65). The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 includes 2011 and 2012 PNAD samples.

Although the labor force aged during the study period – the proportion of the working-age population aged 30 or over increased from 64 percent to 69 percent, and those aged 45 or over increased from 22 percent to 29 percent – this did not translate into an increase in the stock of potential experience. Table 1 confirms that, on average, potential experience in the labor force experienced a slight decline from 23.1 years in 1995 to 22.3 years in 2012. This is because the increase in years of schooling documented in panel A proved sufficient to offset the effect of aging on potential experience.

The formal-informal composition of employment also changed during this period. The proportion of workers with formal contracts ("*carteira de trabalho assinada*") increased by almost a fifth, from 48 percent to 57 percent (table 1). This came at the expense of both informal employees and self-employment. Table 1 also documents a substantial increase in the female share of the labor force between 1995 and 2012, from 38 percent to 42 percent. This trend may be associated with increases in women's educational attainment and, more recently, with large increases in the provision of public childcare.[20] There were also changes in the racial composition of the labor force, with the proportion of non-white workers (mostly Afro-Brazilians and people of mixed race) in the working-age population rising by 8 percentage points, to just over 51 percent of the total.

20 From 1991 to 2007, the proportion of 0–6 year-old children attending childcare increased from 27 percent to 44.5 percent, according to the IBGE (see http://seriesestatisticas.ibge.gov.br/series.aspx?vcodigo=CAJ318).

In terms of the spatial variables (U and $Region$), table 1 points to the continued urbanization of the labor force, with the rural share of the working-age population decreasing by 38 percent, from 16 percent to 10 percent of the total. Changes in the regional composition of the labor force were not particularly pronounced. Finally, some changes in the sectoral composition are apparent.[21] Despite the commodity boom, the period was still characterized by a reduction in the proportion of workers in agriculture. The expansion of the construction sector (which gained 2 percentage points) and construction-related services (2 percentage points in the real estate sector) were also noteworthy.

As the joint distribution of earning covariates, which are denoted F_X, changed over the period, so did the conditional distribution of earnings on those covariates, $F_{Y|X}$. The study uses an extended Mincerian equation as a simple descriptive tool to document changes in the partial associations between wages and its covariates, and loosely refers to the coefficients as returns or "premiums" to various observed worker characteristics. The study thus estimates:

$$\log Y = \beta_0 + \sum_{j=1}^{4} \beta_{1,j} Ed^j + \sum_{j=1}^{4} \beta_{2,j} Exp^j + \beta_3 F + \beta_4 G + \beta_5 R + \beta_6 U + \beta_7 Region + \beta_8 Sector + \varepsilon \quad (2)$$

Table 2 reports the coefficients of eq. (2), estimated as simple OLS regressions for each of the three key years in the analysis: 1995, 2003, and 2012.[22] Years of schooling and years of potential experience are both entered as quartic polynomials.[23] For the categorical variables, the omitted categories are formal employment (for F), male (for G), indigenous and other races (for R), urban (for U), Center-West (for $Region$) and financial services (for $Sector$). For ease of visualization, fig. 3 plots the predicted earnings-education (panel A) and earnings-experience (panel B) profiles implied by the coefficients in table 2. Both curves shift down markedly over time. The convexity of the education premium and the concavity of the experience premium are preserved, but the average returns to both education and experience fall over the period.

Between 1995 and 2003, returns to secondary education (complete or incomplete) and of incomplete tertiary fall, relative to no schooling. But the returns to completing tertiary education rise relative to secondary or incomplete tertiary. In other words, even as the earnings-education profile shifts downwards in this first subperiod, it also becomes more convex. In contrast, the 2003–2012 period is characterized by a reduction of all schooling premiums, as the log-earnings curve continues to move downwards but also becomes less convex, a feature that is shared with other Latin American countries (Gasparini et al. 2011). Returns to potential experience also fall markedly, as shown in panel B of fig. 3. Most of the observed decline takes place in the second subperiod.

Table 2 also documents changes in some key wage premia, conditional on education and experience. The gap between black and white workers narrowed, as did that between both of those groups and the omitted category of (primarily) Asian-descendants.[24] The gender wage gap fell from 45 percent to 34 percent, again a trend shared with other Latin American countries (Ñopo 2012). Disparities across the five main geographical regions of the country, conditional on other observables, were generally stable over the study period, but the gap between rural and urban workers narrowed by 3 percentage points.

21 For a simpler description, here the study groups the 17 economic sectors used in the subsequent regression analysis (see table 2) into four broad groupings: agriculture, fishing, and mining; industry; construction; and services.

22 In eq. 2, β_3, β_5, β_7, and β_8 represent appropriately dimensioned vectors of coefficients.

23 The quartic polynomial specification aims to retain substantial functional form flexibility, while avoiding the use of sets of dummies, for which RIF-regressions are typically less stable.

24 The omitted group is the "indigenous and other" category that basically consists of Asian descendants, as the remaining indigenous population is very small in the Brazilian labor market, particularly since the rural areas of the Northern region are excluded from the sample. The negative coefficient on the white dummy comes from workers of Asian (mostly Japanese) descent, who have typically commanded a premium over observationally comparable white workers.

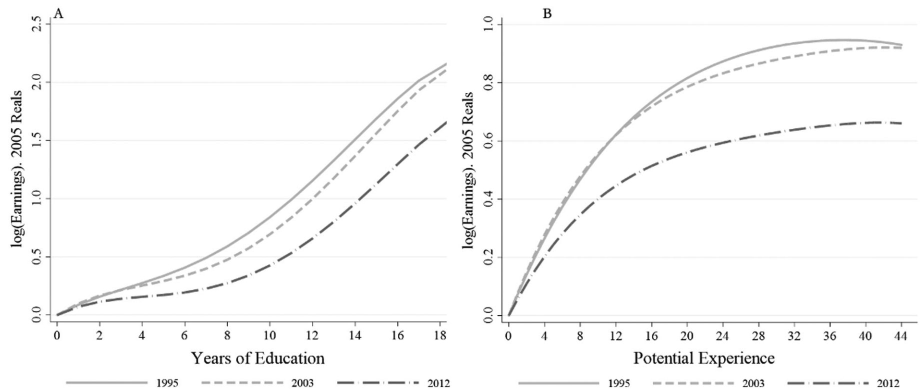
Table 2. Mincerian Equation: Coefficients of Equation (2)

	(1) 1995		(2) 2003		(3) 2012	
	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error
Years of education	0.101**	[0.004]	0.120**	[0.004]	0.089**	[0.004]
Years of education 2/100	–1.431**	[0.108]	–2.281**	[0.093]	–1.983**	[0.089]
Years of education 3/1000	1.779**	[0.105]	2.427**	[0.086]	2.032**	[0.079]
Years of education 4/10000	–0.515**	[0.033]	–0.651**	[0.026]	–0.515**	[0.023]
Potential experience	0.075**	[0.002]	0.081**	[0.002]	0.059**	[0.001]
Potential experience 2/100	–0.235**	[0.016]	–0.305**	[0.013]	–0.238**	[0.011]
Potential experience 3/1000	0.037**	[0.004]	0.058**	[0.003]	0.048**	[0.003]
Potential experience 4/10000	–0.003**	[0.000]	–0.004**	[0.000]	–0.004**	[0.000]
Self-employment	–0.061**	[0.005]	–0.258**	[0.004]	–0.158**	[0.004]
Informal	–0.257**	[0.004]	–0.332**	[0.003]	–0.281**	[0.003]
White	–0.150**	[0.026]	–0.049**	[0.019]	–0.025+	[0.015]
Black	–0.299**	[0.026]	–0.175**	[0.019]	–0.123**	[0.015]
Female	–0.447**	[0.004]	–0.404**	[0.003]	–0.340**	[0.003]
Northeast	–0.374**	[0.005]	–0.404**	[0.004]	–0.383**	[0.004]
North	–0.093**	[0.007]	–0.152**	[0.005]	–0.189**	[0.005]
Southeast	0.048**	[0.005]	–0.003	[0.004]	–0.042**	[0.004]
South	–0.047**	[0.006]	–0.049**	[0.005]	–0.034**	[0.004]
Rural	–0.183**	[0.005]	–0.105**	[0.005]	–0.149**	[0.005]
Agriculture	–0.687**	[0.014]	–0.537**	[0.012]	–0.490**	[0.012]
Fishing	–0.655**	[0.025]	–0.592**	[0.024]	–0.703**	[0.023]
Mining and quarrying	–0.319**	[0.026]	–0.214**	[0.024]	0.002	[0.019]
Manufacturing industries	–0.321**	[0.013]	–0.320**	[0.012]	–0.272**	[0.011]
Electricity, gas, and water	–0.282**	[0.019]	–0.053**	[0.019]	–0.100**	[0.019]
Construction	–0.337**	[0.013]	–0.291**	[0.012]	–0.213**	[0.011]
Trade	–0.378**	[0.013]	–0.322**	[0.011]	–0.300**	[0.010]
Hotels and restaurants	–0.382**	[0.015]	–0.339**	[0.013]	–0.325**	[0.011]
Transport and storage	–0.179**	[0.014]	–0.106**	[0.013]	–0.157**	[0.011]
Real estate	–0.349**	[0.016]	–0.266**	[0.012]	–0.243**	[0.011]
Public administration	–0.406**	[0.013]	–0.167**	[0.012]	–0.098**	[0.011]
Teaching	–0.687**	[0.013]	–0.409**	[0.012]	–0.334**	[0.011]
Social and health	–0.403**	[0.014]	–0.260**	[0.012]	–0.209**	[0.011]
Community services	–0.554**	[0.014]	–0.372**	[0.013]	–0.286**	[0.012]
Domestic service	–0.591**	[0.013]	–0.513**	[0.012]	–0.498**	[0.011]
Extra-territorial org.	0.178	[0.162]	0.247+	[0.126]	0.312*	[0.123]
Constant	5.959**	[0.032]	5.710**	[0.024]	6.248**	[0.020]
Observations	214,001		263,774		270,053	
R-squared	0.52		0.51		0.47	

Source: Pesquisa Nacional por Amostra de Domicílios (PNAD).
Note: Robust standard errors in brackets. **, * and + denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. The omitted categories are: formal employees, center-west, other race, male, urban, and financial services. The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 includes 2011 and 2012 PNAD samples.

The formality premium moved the other way, with both informal employees and the self-employed falling further behind formal sector workers. Finally, most sectoral wage gaps with respect to the omitted financial services sector also declined, but, with 17 sectors in the analysis, it is harder to obtain a sense of the overall net effect of changing industry wage premia on total wage inequality merely from inspecting changes in these coefficients.

Figure 3. Education and Experience Premium by Year. (a) Panel A. Education. (b) Panel B. Potential Experience



Source: *Pesquisa Nacional por Amostra de Domicílios* PNAD.
Note: All regressions include a quartic polynomial in education, a quartic polynomial in potential experience, two dummies for formality status, two race dummies, a gender dummy, four region dummies, sector dummies, and an indicator of work in rural areas (see table 2). The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 includes 2011 and 2012 PNAD samples.

4. Distributional Decomposition Analysis

The trends just described in section 3 reflect changes in both the joint distribution of earning covariates, F_X, and in the conditional distribution, $F_{Y|X}$. Together, they help shape the evolution of the earnings distribution, F_Y. The study now discusses how to formally connect changes in the distribution of covariates (the "composition" of the labor force) and in the sensitivity of earnings to covariates (the "structure" of the labor market), on the one hand, with changes in the earnings distribution on the other.

The question of interest is how covariates have impacted not only average earnings but also other features of the distribution, including inequality measures. The study therefore uses a decomposition approach that was designed explicitly to relate changes in the distribution of covariates to statistics defined on the distribution of earnings (F_Y), such as quantiles, the Gini coefficient, percentile ratios, or other inequality measures. This decomposition method makes it possible to understand, for example, by how much the changes in the distribution of education have contributed to the decline in the Gini coefficient over time, isolating these effects from those arising from other factors.

The approach is based on Re-centered Influence Function (RIF) regressions, which were introduced by Firpo, Fortin, and Lemieux (2009, henceforth FFL). Influence Functions (IF) were developed within the robust statistics literature (Hampel 1974) to provide a measure of how a specific feature, or summary statistic, of the distribution (such as a specific quantile, like the median, or an inequality index, like the Gini) is affected by each data point in the distribution.[25]

Let υ be a given feature, or statistic, of the earnings (Y) distribution, such as the mean μ_Y, the τ^{th} quantile q_τ, or the Gini coefficient G_Y. In general, the influence function of υ is written as $IF(Y; \upsilon)$: for each υ, it is a function of Y. The IF is constructed to have mean zero. Re-centering it, that is, defining the RIF as $RIF(Y; \upsilon) = \upsilon + IF(Y; \upsilon)$, guarantees that the mean of the RIF equals the distributional feature of interest.

Because the RIF measures how a given distributional statistic is affected by each data point in the sample over which it is calculated, it can be used to measure how that statistic changes if it were possible to arbitrarily change the data distribution. Let $F_Y(y)$ be the cumulative distribution function of earnings at a support point y. It can be written, using the law of iterated expectations, as $F_Y(y) = E[F_{Y|X}(y|X)]$.

25 It is known, for example, that the median, unlike the mean, is robust to the presence of outliers, and this difference is captured by differences in their IFs: Whereas the median has a bounded IF, the IF of the mean is unbounded.

The change in the distribution of Y that is considered by FFL keeps the conditional distribution $F_{Y|X}(y|X)$ fixed and allows a small change in the distribution of covariates.

FFL show that the effect of that marginal change in the distribution of covariates (F_X) on any given distributional feature υ of the earnings distribution (F_Y) is captured by the average derivative of the conditional expectation of RIF$(Y;\upsilon)$ on X.[26] When that conditional expectation, E[RIF$(Y;\upsilon)|X$], can be approximated by a linear regression, its average derivative will be approximated by the coefficients of a linear regression. In mathematical terms, given the linear approximation, $E[\frac{dE[RIF(Y;\upsilon)|X]}{dX}] \cong \beta^{\upsilon}$, where β^{υ} are the coefficients of a linear regression of RIF$(Y;\upsilon)$ – instead of Y – on X. The vector of coefficients β^{υ} therefore has a clear interpretation as the effect of a small location shift in the distribution of covariates on υ. Differences in υ, such as difference in means, quantiles, and Gini coefficients, can therefore be written as the sum of expected differences in conditional expectations (RIF-regressions) and differences in the distribution of the covariates.[27]

RIF-regressions thus generalize the well-known methods by Oaxaca (1973) and Blinder (1973), which were used to decompose changes in the means of earnings distributions, to a much broader class of functionals of that distribution, such as quantiles, quantile ratios, and inequality measures (Fortin, Lemieux, and Firpo 2011). For any distribution statistic, one can separate changes over time into a *structure* effect, which holds the distribution of the covariates F_X constant; and a *composition effect*, which holds the conditional distribution $F_{Y|X}$ constant. In fact, RIF-regressions, when applied to the mean, yield exactly the same decomposition proposed by Blinder (1973) and Oaxaca (1973). This happens because RIF$(Y;\mu_Y) = Y$. Therefore, the traditional Oaxaca-Blinder method is a special case of the decomposition method using RIF.

Implementation of the FFL decomposition proceeds in two stages: (i) an estimation stage, and (ii) a computation stage.

The Estimation Stage

The estimation stage consists of estimating β^{υ}, the vector of coefficients of a linear regression of RIF$(Y;\upsilon)$ on X. This is attained simply by replacing the dependent variable *log Y* in an OLS earnings regression – such as eq. (2) – with the corresponding sample version of RIF$(Y;\upsilon)$. In this paper, the choices for υ are the mean μ_Y, the Gini coefficient G_Y and the τ^{th} percentile, q_τ. All one needs to know is the exact formula for the RIF for each statistic υ, and how to compute it. Fortunately, a long list of RIFs has been presented in previous work. Essama-Nssah and Lambert (2012) offer one of these lists containing RIFs for many inequality measures, including the Gini coefficient. Formulae for the mean and quantiles are also presented there.[28]

As previously discussed, for the mean μ_Y, RIF$(Y;\mu) = Y$. This is the only case in which the RIF is identical to the original dependent variable. The influence function of the τ^{th} quantile, q_τ, is $IF(y;q_\tau) = \frac{\tau - 1\{y \le q_\tau\}}{f_Y(q_\tau)}$, where f_Y is the density function of Y.[29] Therefore, $RIF(y;q_\tau) = q_\tau + \frac{\tau - 1\{y \le q_\tau\}}{f_Y(q_\tau)}$. For the Gini coefficient, G_Y, $RIF(y;G_Y) = \frac{1}{\mu}(\mu - y(1+G_Y) + 2F_Y(y)y - 2H_Y(y))$, where $H_Y(y) = E[1\{Y \le y\}Y]$.

Naturally, implementation of RIF regressions requires using the sample counterparts of the described population objects. Thus, we replace $RIF(y;q_\tau)$ by $\widehat{RIF}(y;q_\tau)$ and $RIF(y;G_Y)$ by $\widehat{RIF}(y;G_Y)$. In practice,

26 Consider the conditional expectation of RIF$(Y;\upsilon)$ on X evaluated at x, $E[RIF(Y;\upsilon)|X = x]$. Its derivative is $dE[RIF(Y;\upsilon)|X = x]/dx$. As this derivative is a function of X, its average over the distribution of covariates X is the so-called average derivative and is commonly represented by $E[dE[RIF(Y;\upsilon)|X]/dX]$.

27 The decomposition is not exact for discrete changes, as those distributional features can only locally be approximated by expectations and the average derivative of the RIF is not necessarily constant. These limitations of the method are discussed in Fortin, Lemieux, and Firpo (2011).

28 The estimation was carried out with Stata commands rifreg and oaxaca.

29 The indicator function 1{A} equals 1 when A is true and 0 otherwise. Thus, another way of writing the influence function for the quantile would be $IF(y;q_\tau) = \tau/f_Y(q_\tau)$ if $q_\tau < y$ and $IF(y;q_\tau) = -(1-\tau)/f_Y(q_\tau)$ if $y \le q_\tau$.

for the i-th observation in the data two new variables are constructed, $\widehat{\mathrm{RIF}}(y_i; q_\tau)$ and $\widehat{\mathrm{RIF}}(y_i; \mathrm{G}_Y)$, where $\widehat{\mathrm{RIF}}(y_i; q_\tau) = \hat{q}_\tau + \frac{\tau - 1\{y_i \le \hat{q}_\tau\}}{\hat{f}_Y(\hat{q}_\tau)}$ and $\widehat{\mathrm{RIF}}(y_i; \mathrm{G}_Y) = \frac{1}{\bar{Y}}(\bar{Y} - y_i(1 + \hat{\mathrm{G}}_Y - 2\hat{F}_Y(y_i)) - 2\hat{H}_Y(y_i))$.[30] Then the following regressions are run:[31]

$$\widehat{\mathrm{RIF}}(y; q_\tau) = \hat{\beta}_0^{q_\tau} + \sum_{j=1}^{4} \hat{\beta}_{1,j}^{q_\tau} Ed^j + \sum_{j=1}^{4} \hat{\beta}_{2,j}^{q_\tau} Exp^j + \hat{\beta}_3^{q_\tau} F + \hat{\beta}_4^{q_\tau} G + \hat{\beta}_5^{q_\tau} R + \hat{\beta}_6^{q_\tau} U + \hat{\beta}_7^{q_\tau} Region + \hat{\beta}_8^{q_\tau} Sector + \varepsilon \tag{3}$$

$$\widehat{\mathrm{RIF}}(y; \mathrm{G}_y) = \hat{\beta}_0^{G_Y} + \sum_{j=1}^{4} \hat{\beta}_{1,j}^{G_Y} Ed^j + \sum_{j=1}^{4} \hat{\beta}_{2,j}^{G_Y} Exp^j + \hat{\beta}_3^{G_Y} F + \hat{\beta}_4^{G_Y} G + \hat{\beta}_5^{G_Y} R + \hat{\beta}_6^{G_Y} U + \hat{\beta}_7^{G_Y} Region + \hat{\beta}_8^{G_Y} Sector + \varepsilon \tag{4}$$

Equation (3) yields the estimates $\hat{\beta}^{\upsilon}$ when $\upsilon = q_\tau$ and eq. (4) does so for $\upsilon = \mathrm{G}_Y$.

The Computation Stage

FFL show that, under the assumption of a linear conditional expectation function as discussed above, differences in statistic υ can be written as the sum of two components: namely, the structure effect and the composition effect. For example, the difference in υ between two periods, t = 1 and t = 2, is given by:

$$\hat{\upsilon}_{t=2} - \hat{\upsilon}_{t=1} = \widehat{\Delta\upsilon} = \hat{\Delta}_S^{\upsilon} + \hat{\Delta}_X^{\upsilon} \tag{5}$$

In eq. (5), $\hat{\Delta}_S^{\upsilon}$ denotes the structure effect, which, as discussed, holds the distribution of covariates constant and captures changes in statistic υ arising from changes in the conditional distribution, using the estimates of the average derivatives of the conditional expectations (RIF-regressions), namely $\hat{\beta}^{\upsilon}$, from the estimation stage. Much as in Oaxaca (1973), the decomposition arithmetically requires appropriately weighting the differences in $\hat{\beta}^{\upsilon}$, using covariates as weights. One formulation is given by eq. (6) below:

$$\hat{\Delta}_S^{\upsilon} = \sum_{j=1}^{k} \hat{\Delta}_{S,j}^{\upsilon} = \sum_{j=1}^{k} \left[\bar{X}_{2,j}' \left(\hat{\beta}_{2,j}^{\upsilon} - \hat{\beta}_j^{\upsilon} \right) + \bar{X}_{1,j}' \left(\hat{\beta}_j^{\upsilon} - \hat{\beta}_{1,j}^{\upsilon} \right) \right] \tag{6}$$

If the structure effect is written as in eq. (6), then the composition effect – which holds the conditional distribution constant and captures changes in the joint distribution of covariates – is written as in eq. (7) below:

$$\hat{\Delta}_X^{\upsilon} = \sum_{j=1}^{k} \hat{\Delta}_{X,j}^{\upsilon} = \sum_{j=1}^{k} \left(\bar{X}_{2,j}' - \bar{X}_{1,j}' \right) \hat{\beta}_j^{\upsilon} \tag{7}$$

In eqs. (6) and (7), an upper bar denotes averages across individuals, and $\hat{\beta}_{t,j}^{\upsilon}$ and $\hat{\beta}_j^{\upsilon}$ correspond respectively to the coefficients associated with covariate j in a regression of the recentered influence function of υ on X for period t, and for the pooled sample of the two periods. Note also that the additive separability of these estimates implies that detailed decompositions, where the structure and composition effects can be evaluated separately for each element of X are straightforwardly obtained for each covariate $j = 1,\ldots,k$. The computation stage of the decomposition consists simply of using the $\hat{\beta}^{\upsilon}$ estimates obtained in

30 Here, $\hat{q}_\tau$ is the sample τ-th quantile; $\hat{f}_Y$ is a kernel density estimator; $\bar{Y}$ is the sample mean; $\hat{\mathrm{G}}$ is the Gini for the sample; $\hat{F}_Y$ is the empirical distribution function ($\hat{F}_Y(y) = \frac{1}{N}\sum_{i=1}^{N} 1\{y_i \le y\}$) and $\hat{H}_Y$ is the sample analog of H_Y, that is, $\hat{H}_Y(y) = \frac{1}{N}\sum_{i=1}^{N} 1\{y_i \le y\} y_i$.

31 For the cases of the mean and the quantiles, log-earnings are used as the dependent variable Y. For the Gini, the dependent variable in levels was used.

the estimation stage and computing the structure and composition effects, either overall or for a particular covariate j, using eqs. (6) and (7).

A relevant detail for the implementation of this study's method is the fact, noted in section 3, that many of the regressors are categorical variables. It is well-known in the literature on Oaxaca-Blinder decompositions that results are not invariant to the choice of the excluded category (Oaxaca and Ransom 1999; Yun 2005; Gardeazabal and Ugidos 2004; and Fortin, Lemieux, and Firpo 2011). Several attempts to solve the invariance problem have been proposed in the literature. This paper does not impose restrictions on the parameters beyond the usual practice of dropping one from each set of dummies and including the constant. The study opted for choosing the best performers (i.e., the categories displaying the highest wages in the pooled sample) to be the omitted categories (Asian descendant males, urban center-west, and being a formal employee). There is a practical rationale for this choice: 1995–2012 is a period of rapid inequality reduction in Brazil. Hence, the earnings of the most disadvantaged groups tended to grow faster than the earnings of the most advantaged. By selecting the most advantaged as the reference category, the role of the unobserved component in the decomposition is minimized.

5. Results

Results are presented in four parts. In the first three subsections below the study decomposes the changes between 1995 and 2012 – and for the subperiods 1995–2003 and 2003–2012 – in the three main statistics of interest (mean incomes, the Gini coefficient, and a set of quantiles and quantile ratios) as per eq. (5), into the structure ($\hat{\Delta}_S^{\upsilon}$) and composition ($\hat{\Delta}_X^{\upsilon}$) effects defined in eqs. (6) and (7). In each case, detailed decompositions of each effect by individual covariates ($\hat{\Delta}_{S,j}^{\upsilon}$, $\hat{\Delta}_{X,j}^{\upsilon}$) are also presented.[32] The first subsection looks at average earnings; the second subsection decomposes changes in the Gini coefficient; and the third subsection considers individual percentiles and three specific percentile ratios (90-10, 90-50, and 50-10). Finally, the last subsection examines the sensitivity of the study's main findings to various alternative samples, variable definitions, and regression specifications, in a robustness section.

Average Earnings

Table 3 presents the results of the FFL decomposition of differences in average earnings over time into the composition and pay structure effects, both overall and for each (set of) covariate(s). These are the decompositions from eqs. (5) to (7), when the $\hat{\beta}_j^{\upsilon}$ are obtained from the estimates in eq. (2): $\hat{\beta}_j^{\upsilon} = \hat{\beta}_j$.[33] The first column presents results for the full period, 1995–2012, whereas the next two columns refer to the first (1995–2003) and second (2003–2012) subperiods, respectively. The top panel gives log earnings in each relevant year, the difference between them, and the decomposition of this difference into the overall composition and structure effects. The second panel further decomposes composition effects into those of individual (or groups of) variables, whereas the third panel does the same for structure effects. The bottom panel combines structure and composition effects for each (set of) variable(s) and reports their total contribution.

Average earnings grew by 26 percent over the entire period, decreasing by 12 percent in the first subperiod and then increasing by 38 percent in the second.[34] For the full period, the composition effect accounts for the bulk of the change (22 percent), with the largest effect coming in 2003–2012. The most important component of the composition effect was the increase in educational attainment, which contributed 0.21 log points to the increase in earnings. The structure effects are muted overall (0.04 log points), but that

32 To save space, results for race and gender are presented together (as demographic characteristics of workers), as are those for rural-urban status and region of residence (spatial characteristics).

33 Recalling once again that $\mathrm{RIF}(Y; \mu) = Y$, so that the RIF-regression decomposition for average earnings is simply the standard Oaxaca-Blinder decomposition.

34 Approximating growth rates by log changes.

Table 3. Decomposition Results: Changes in Average Earnings

	2012–1995	2003–1995	2012–2003
Overall			
Log earnings, end of period	6.48**	6.10**	6.48**
	[0.002]	[0.002]	[0.002]
Log earnings, beginning of period	6.23**	6.23**	6.10**
	[0.002]	[0.002]	[0.002]
Difference	0.26**	–0.12**	0.38**
	[0.003]	[0.003]	[0.002]
Composition effect	0.22**	0.09**	0.14**
	[0.002]	[0.002]	[0.002]
Structure effect	0.04**	–0.21**	0.24**
	[0.002]	[0.002]	[0.002]
Composition effects			
Education	0.21**	0.10**	0.12**
	[0.001]	[0.001]	[0.001]
Potential experience	–0.02**	–0.02**	–0.01**
	[0.001]	[0.001]	[0.000]
Formality status	0.02**	–0.00**	0.02**
	[0.000]	[0.000]	[0.000]
Race and gender	–0.03**	–0.01**	–0.01**
	[0.001]	[0.001]	[0.001]
Region and urban	0.01**	0.01**	0.01**
	[0.001]	[0.001]	[0.000]
Economic sector	0.03**	0.01**	0.01**
	[0.001]	[0.000]	[0.000]
Structure effects			
Education	–0.29**	–0.08**	–0.22**
	[0.007]	[0.007]	[0.007]
Potential experience	–0.22**	–0.02	–0.20**
	[0.012]	[0.013]	[0.009]
Formality status	–0.03**	–0.07**	0.04**
	[0.002]	[0.002]	[0.002]
Race and gender	0.07**	0.03**	0.04**
	[0.003]	[0.003]	[0.003]
Region and urban	–0.05**	–0.03**	–0.01**
	[0.005]	[0.006]	[0.005]
Economic sector	0.13**	0.10**	0.04*
	[0.015]	[0.016]	[0.015]
Constant	0.43**	–0.13**	0.56**
	[0.022]	[0.024]	[0.020]
Combined effects			
Education	–0.08**	0.02*	–0.10**
	[0.007]	[0.007]	[0.007]
Potential experience	–0.24**	–0.03*	–0.21**
	[0.012]	[0.013]	[0.009]
Formality status	–0.01**	–0.07**	0.06**
	[0.002]	[0.002]	[0.002]
Race and gender	0.04**	0.01**	0.02**
	[0.003]	[0.003]	[0.003]
Region and urban	–0.03**	–0.02**	–0.01*
	[0.005]	[0.006]	[0.005]
Economic sector	0.16**	0.11**	0.05**
	[0.015]	[0.016]	[0.015]
N	484,054	477,775	533,827

Source: Pesquisa Nacional por Amostra de Domicílios(PNAD).

Note: Robust standard errors in brackets. **, * and $^{+}$ denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. Standard errors are calculated using the delta method. Each category summarizes the contribution of the sum of individual effects. Education and experience refer to quartic polynomials in education and experience, respectively, in the underlying regressions. Formality status summarizes the contribution of two dummies: informal employee and self-employed. Race and gender captures the interactions between gender and race dummies ('mestiço' and Afro-Brazilian, white, and others). Region and urban includes the interactions between an indicator variable of living in an urban area and region (northeast, north, southeast, south, and center-west). Economic sector adds 17 dummies for economic sector. Omitted categories in the underlying regressions are: formal employees, center-west, white, male, urban, and financial services. The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 includes 2011 and 2012 PNAD samples.

again reflects a negative effect during 1995–2003, offset by a positive one in 2003–2012. In disaggregated terms, these net outcomes reflect positive earnings growth effects from sectoral wage premia and, most importantly, from rising average returns to unobserved skills, captured by the constant term.

Changes in the Gini Coefficient

Turn now to changes in inequality, beginning with the Gini coefficient ($\upsilon = G_Y$), whose decompositions are described in table 4: for the full 1995–2012 period in column (1), and for 1995–2003 and 2003–2012 in columns 2 and 3, respectively.[35] These are the decompositions in eqs. (5) to (7), when the $\hat{\beta}_j^{\upsilon}$ are obtained from the estimates in eq. (4): $\hat{\beta}_j^{\upsilon} = \hat{\beta}_j^{G_Y}$. Overall, the Gini coefficient decreased by 9 percentage points, from 0.498 in 1995 to 0.408 in 2012. In 2003, it was 0.467, so most of the reduction occurred in the second subperiod. In contrast to the decomposition for average earnings, structure effects explain all of the inequality reduction between 1995 and 2012 and were indeed partly offset by a small positive composition effect: Had only changes in the distribution of covariates taken place during this period, inequality would have increased by 1 Gini point (table 4). The same pattern holds for both subperiods, as changes in the distribution of observable characteristics of the workforce were inequality-enhancing between 1995 and 2003 (0.4 Gini points) and 2003 and 2012 (0.9 points).

To understand the inequality-increasing composition effect, note that it is driven entirely by the education component (+3.2 Gini points). All other variables make negative (inequality-reducing) contributions that tend to incompletely offset the education effect. This inequality-increasing effect of an educational expansion (at given returns) is consistent with the "paradox of progress" (Bourguignon, Ferreira, and Lustig 2005): increases in educational attainment – even when its dispersion declines – may be inequality-increasing because of the marked convexity of the earnings-education profile.[36] As the distribution of education shifts to the right, density mass moves toward ranges of years of schooling with steeper returns. That contributes to an increase in average earnings (as seen in the previous subsection), but also to rising earnings inequality. This effect is illustrated in fig. 4, which superimposes the kernel density estimates of the distributions of years of schooling in 1995 and 2012 with the log-earnings – education profile, estimated as in eq. (2), but for the pooled 1995 and 2012 samples. The rightward shift in the distribution of years of schooling implies that in 2012 there is a much larger mass of workers than in 1995 in a range where schooling premia are higher.

If compositional changes were on balance inequality-enhancing, then the bulk of inequality reduction during 1995–2012 must be explained by changes in the structure effect. The main drivers of the negative structure effect were (i) a large decline in the returns to potential experience; (ii) falling conditional earnings gaps associated with the race, gender, and spatial location of workers; and (iii) an equalizing change in the structure of returns to unobserved skills, captured by the constant term.

The most important observed structure effect was the reduction in returns to experience. On its own, it contributed 4.7 points to the reduction in the Gini coefficient during 1995–2012. But although it was the most important factor, the structure effect of potential experience did not act alone. Column (1) of table 4 reports reductions in the gender and racial wage gaps (–1.52) and in regional and urban-rural disparities in wages (–0.82) for the overall period.[37] These effects are both statistically and economically significant

35 The working paper version of this paper (Ferreira, Firpo, and Messina 2017) presents parsimonious decompositions with an incrementally richer set of covariates in each decomposition. The contributions of each single factor to the changes in the Gini are relatively stable across specifications.

36 See also Knight and Sabot (1983) and Lam (1999).

37 A possible causal mechanism behind this decline in regional and urban-rural wage gaps is suggested by the work of Dix-Carneiro (2014) and Dix-Carneiro and Kovak (2017) These papers build and estimate structural dynamic equilibrium models of the Brazilian labor market, with the principal objective of understanding how the effects of trade liberalization on wages and employment evolve over time. Although they seldom discuss wage inequality explicitly, the persistent employment and wage dynamics they estimate imply a long-term decline in the wage gaps between metropolitan and

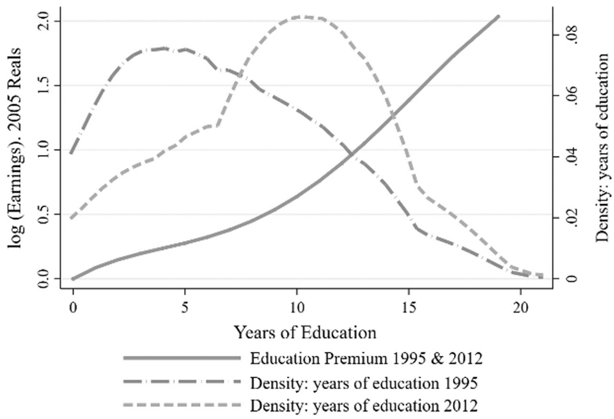
Table 4. Decomposition Results. Changes in Inequality. Gini

	2012–1995	2003–1995	2012–2003
Overall			
Gini coefficient (G_Y), end of period	40.82**	46.74**	40.82**
	[0.076]	[0.073]	[0.077]
Gini coefficient (G_Y), beginning of period	49.77**	49.77**	46.74**
	[0.071]	[0.078]	[0.067]
Difference	−8.95**	−3.03**	−5.92**
	[0.104]	[0.107]	[0.102]
Composition effect	1.06**	0.39**	0.89**
	[0.061]	[0.045]	[0.048]
Structure effect	−10.01**	−3.42**	−6.81**
	[0.109]	[0.103]	[0.099]
Composition effects			
Education	3.21**	1.13**	2.48**
	[0.059]	[0.042]	[0.045]
Potential experience	−0.32**	−0.27**	−0.11**
	[0.012]	[0.011]	[0.009]
Formality status	−0.87**	0.02	−0.94**
	[0.017]	[0.014]	[0.017]
Race and gender	−0.11**	−0.03**	−0.09**
	[0.012]	[0.007]	[0.008]
Region and urban	−0.20**	−0.12**	−0.07**
	[0.014]	[0.012]	[0.010]
Economic sector	−0.64**	−0.34**	−0.37**
	[0.024]	[0.020]	[0.017]
Structure effects			
Education	0.25	1.03**	−1.18**
	[0.285]	[0.275]	[0.295]
Potential experience	−4.71**	−1.15+	−3.50**
	[0.599]	[0.677]	[0.514]
Formality status	0.82**	1.03**	−0.16
	[0.108]	[0.122]	[0.101]
Race and gender	−1.52**	−0.99**	−0.52**
	[0.180]	[0.173]	[0.184]
Region and urban	−0.82**	−0.22	−0.60*
	[0.293]	[0.298]	[0.284]
Economic sector	3.09*	1.23	1.93
	[1.486]	[1.570]	[1.523]
Constant	−7.12**	−4.33*	−2.79+
	[1.675]	[1.786]	[1.672]
Combined effects			
Education	3.46**	2.16**	1.30**
	[0.290]	[0.277]	[0.299]
Potential experience	−5.03**	−1.43*	−3.60**
	[0.599]	[0.677]	[0.514]
Formality status	−0.05	1.04**	−1.10**
	[0.108]	[0.122]	[0.101]
Race and gender	−1.64**	−1.02**	−0.61**
	[0.183]	[0.173]	[0.186]
Region and urban	−1.02**	−0.35	−0.68*
	[0.292]	[0.299]	[0.284]
Economic sector	2.45+	0.89	1.56
	[1.487]	[1.571]	[1.523]
N	484,054	477,775	533,827

Source: Pesquisa Nacional por Amostra de Domicílios (PNAD).

Note: Standard errors are robust, and for combined effects are calculated with the delta method. **, *, and + denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. The Gini coefficient is expressed in percentage points and goes from 0 (perfect equality) to 100 (perfect inequality). See the note in table 3 for details on the individual effects included in each category. The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD sample, and the year 2012 included 2011 and 2012 PNAD sample.

Figure 4. The Paradox of Progress.



Source: Pesquisa Nacional por Amostra de Domicílios PNAD.

Note: The solid line draws the predictions of a quartic polynomial in education in a regression that includes a quartic polynomial in potential experience, two dummies for formality status, two race dummies, a gender dummy, four region dummies, sector dummies and an indicator of work in rural areas. The regression uses pooled data from the PNAD for the years 1995 and 2012. Kernel density functions use a bandwidth h = 2.0.

and appear to have operated in both subperiods, although the spatial structure effect was not significant during 1995–2003.

The premia associated with unobserved skills, captured in the framework by the constant term under the structure effect, also fell markedly and contributed substantially to reducing inequality: –7.12 Gini points. This is consistent with a single index model of residual inequality, where the price of unobserved skills falls if the premium on observable skills declines (Acemoglu 2002).

The contribution of changes in the returns to schooling, and experience merits additional discussion. For the 1995–2012 period, changes in the schooling premium had mildly positive, but statistically insignificant, effects on inequality. When these are added to the positive composition effects of education discussed above, the total effect of education was strongly inequality-enhancing (by 3.5 Gini points; see bottom panel, column 1 in table 4).

There were important differences across subperiods: Between 1995 and 2003, the total contribution of education to inequality is stronger (2.2 Gini points according to the estimates in column 2 of table 4), because the positive effect of compositional changes is reinforced by a positive structure effect. As noted in section 4, this is a period when the high school premium with respect to basic or no education is falling, but the premium of tertiary education with respect to high school is rising. This second effect turns out to dominate, at least in terms of the Gini coefficient.

The 2003–2012 interval, on the other hand, is characterized by a reduction in all schooling premia. The between-group differences across all schooling levels declined, as captured by a negative and significant structure effect (–1.2 in column 3 of table 4). Although the total effect of education (combining composition and structure) remains positive in this latter subperiod, reflecting a large composition effect, it is lower than in the earlier period. These inequality-increasing combined effects of education contrast sharply with the emphasis on educational dynamics as the main drivers of inequality reduction that characterized the earlier literature described in section 2.

It is potential experience that turns out to be the main driving force behind Brazil's inequality reduction in 1995–2012. Both the composition and structure effects of potential experience reported in table 4

large urban areas on the one hand (where industries were exposed to the largest negative tariff shocks in the 1980s and 1990s), and the rest of the country on the other.

are negative, although the structure effect was much larger in magnitude. This large inequality-reducing structure effect comes predominantly from the second subperiod, 2003–2012, reflecting the sharp reduction of the experience premium previously discussed (see fig. 3, panel B). The last subsection returns to the issue of falling returns to experience and uses the RAIS data to show that a similar temporal decline can be observed for returns to *actual* experience in the formal sector. In addition, although there are separate cohort effects, this study shows that the period effects that are described here – that is, a temporal decline in returns – are robust to controlling for cohort and age effects.

Percentiles and Percentile Ratios

However influential it may be, the Gini coefficient is but one measure of earnings dispersion. The study has also performed RIF-regression decompositions for each quantile of the earnings distribution, that is letting $\upsilon = q_\tau,\ \tau \in (0, 1)$. These are the decompositions in eqs. (5) to (7), when the $\hat{\beta}_j^{\upsilon}$ are obtained from the estimates in eq. (3): $\hat{\beta}_j^{\upsilon} = \hat{\beta}_j^{q_\tau}$. In practice, 99 percentiles were used, and point estimates are presented graphically in fig. 5. The curves plotting observed log earnings differences in fig. 5 are essentially earnings growth incidence curves (see Ravallion and Chen 2003), and the curves denoting either endowment or structure effects are the corresponding counterfactual growth incidence curves (see Ferreira 2012).

The top panel of fig. 5 presents the overall decomposition of observed log income differences at each percentile into composition and structure effects. These headline results are very consistent with what was learned from the decomposition of the Gini. First, for the entire period, the overall structure and endowment components had offsetting effects on earnings inequality: the structure effect is mostly downward-sloping (inequality-reducing), whereas the endowment effect is (mildly) upward-sloping along the distribution. Thus, changes in the distribution of labor force characteristics were generally unequalizing, and the observed decline in earnings inequality (captured by the growth incidence curves) is driven by changes in the structure of pay.[38]

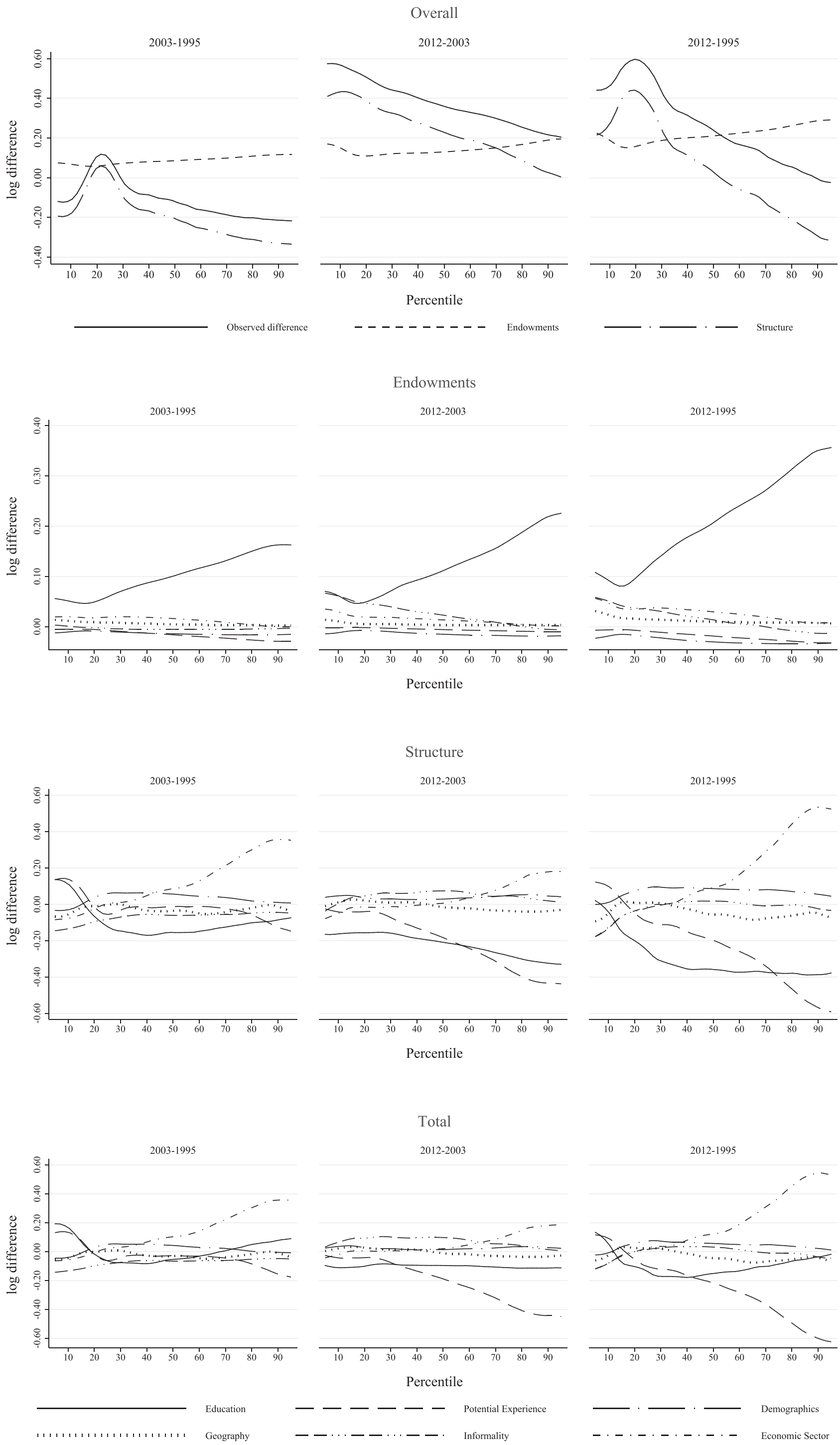
A further decomposition of the earnings changes at each percentile into detailed structure and composition effects (middle panels in fig. 5) reveals that the only upward-sloping endowment effect curve is the one for education. This is true in both subperiods and is consistent with the analysis for the Gini coefficient in table 3, panel 2. All other variables have moderately downward-sloping composition effect curves, but their combined effect is insufficient to fully offset the education effect.

There is more action in the structure effect graphs in fig. 5. In both subperiods, but particularly the second, the counterfactual growth incidence curve for the potential experience structure effect is downward sloping, contributing to a fall in inequality. The structure of returns to economic sectors (arising from conditional industry wage-premia) operates in the opposite direction. Their curves are upward sloping in both subperiods, but particularly the former. This too is consistent with findings from the Gini decomposition, though the effect appears more marked here. The structure effect for education is U-shaped in 1995–2003, consistent with the earlier observation that returns to completed high school fell relative to no education, while the premium for a completed college education rose relative to high school in that first subperiod. During 2003–2012, the structure effects of education were downward-sloping throughout, and thus unambiguously inequality-decreasing.

The remaining factors played relatively minor roles, but the gaps associated with demographic and spatial factors tended to be equalizing. When endowment and structure effects for each factor are added (bottom panel of fig. 5), economic sector premia are the main unequalizing factor in this disaggregated analysis, just as potential experience was the main factor driving inequality reduction.

38 The only part where structure effects were unequalizing was at the very bottom of the distribution, up to the 20^{th}–25^{th} percentile. This feature is much more marked in the first subperiod, 1995–2003, and is consistent with a minimum wage spike that lifted earnings substantially around the minimum wage, but not by as much for those further down the distribution (see Ferreira, Firpo, and Messina 2017).

Figure 5. Decomposing Changes in Earnings by Percentile. (a) Overall. (b) Endowments. (c) Structure. (d) Total



Source: Pesquisa Nacional por Amostra de Domicílios PNAD.

Note: The graph shows the generalized kernel-weighted local polynomial Oaxaca decomposition for each centile of the earnings distribution. See notes in table 3 for details on the individual effects included in each category. The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 includes 2011 and 2012 PNAD samples.

These conclusions, based on visual inspection, are broadly confirmed by a formal analysis based on selected percentile ratios. In particular, the study looks at the log 90-10 percentile ratio, a measure of top-to-bottom gaps, and at the log 90-50 and log 50-10 ratios, which are informative of inequality developments at the top and bottom of the distribution, respectively. Results of the RIF decomposition applied to these three inequality measures are presented (as log differences) in table 5.

As in table 4, the top panel of table 5 shows the evolution of the three inequality measures in 1995–2012 and in the two subperiods, and how structure and endowment effects contributed to the observed patterns. For 1995–2012, the log difference of the 90-10 range declined by an impressive 0.45 log points, confirming the sharp inequality reduction in Brazil. Inequality declined during both subperiods, and the reduction took place both at the top and at the bottom. The top panel of table 5 also confirms the finding from the Gini decomposition that the overall composition effect was always inequality-increasing during the period (except for p50-p10 during 2003–2012).

The decomposition of the detailed determinants of the evolution of the 90-10 percentile ratio, reported in the lower panels of table 5, confirms some of the patterns discussed for the evolution of the Gini coefficient, but places even greater emphasis on the importance of the falling experience premium. Reductions in the experience premium explain even more than the observed changes in p90-p10 inequality for the period as a whole and compensate for other factors that were inequality-augmenting, such as the distribution of economic sector premia. The reduction of the schooling premium also had an important equalizing role, but in line with the previous discussion of different effects on different parts of the distribution, depending on the period. During 1995–2003, reductions of the schooling premium contributed to wage compression in the lower half of the distribution (a decline of the p50-p10), while during 2003–2012 they only contributed to declines of the p90-p10. This equalizing force was partly offset by an inequality-increasing composition effect, as before. The results for race, gender, regional, and urban-rural gaps are less clear-cut for these percentile ratios than they were for the Gini index.

Robustness

Wage Inequality Decompositions

This subsection assesses the robustness of the findings in a number of different ways. First, it investigates changes in the sample (e.g., restricting it to men only, or excluding self-employed workers), different weighting schemes, and alternative definitions of earnings. Due to space constraints, we conduct this robustness analysis only for the Gini regressions discussed above. Table S1.1 in the supplementary online appendix, (available with this article at *The World Bank Economic Review* website) presents results for the 1995–2012 period, and tables S1.2 and S1.3 show the corresponding results for the subperiods 1995–2003 and 2003–2012. Column (1) re-estimates the Gini RIF regression without using sampling weights. Column (2) excludes female workers. Column (3) uses hourly wages instead of monthly earnings, and column (4) restricts the sample to full-time workers – those working 35 hours per week or more. Finally, column (5) presents results for wage employees only, thus excluding the self-employed from the sample.

Results are to a large extent stable across specifications. The main finding that declining returns to experience were the most important driver of falling wage inequality survives in all five specifications. So do the inequality-increasing composition effects of education on inequality in both periods, and the inequality-reducing structure effect of education during 2003–2012. The equalizing effects of declining racial and gender wage gaps are also consistent across specifications. By contrast, the combined effects of declining regional and urban/rural disparities on inequality reduction are not robust; in particular, they lose significance in the hourly wage specifications.

It is also interesting to note that both the strongly unequalizing structure effect of the economic sector variable and the strongly equalizing structure effect of unobserved skills (captured by the constant) are robust across specifications, *except* for the male-only regression. This suggests, among other things, the

Table 5. Decomposition Results. Changes in log(p90/p10), log(p90/p50), and log(p50/p10).

	2012–1995			2003–1995			2012–2003		
	p90-p10	p90-p50	p50-p10	p90-p10	p90-p50	p50-p10	p90-p10	p90-p50	p50-p10
Overall									
Log percentile ratio end of period	1.80**	1.09**	0.71**	2.17**	1.18**	0.99**	1.80**	1.09**	0.71**
	[0.010]	[0.020]	[0.018]	[0.005]	[0.005]	[0.008]	[0.005]	[0.024]	[0.022]
Log percentile ratio beginning period	2.25**	1.31**	0.94**	2.25**	1.31**	0.94**	2.17**	1.18**	0.99**
	[0.007]	[0.005]	[0.004]	[0.005]	[0.005]	[0.003]	[0.006]	[0.007]	[0.003]
Difference	−0.45**	−0.22**	−0.22**	−0.09**	−0.14**	0.05**	−0.36**	−0.09**	−0.27**
	[0.015]	[0.020]	[0.017]	[0.006]	[0.006]	[0.008]	[0.009]	[0.026]	[0.023]
Composition effect	0.11**	0.08**	0.03**	0.05**	0.03**	0.02**	0.04**	0.06**	−0.02**
	[0.004]	[0.004]	[0.002]	[0.003]	[0.002]	[0.002]	[0.003]	[0.002]	[0.002]
Structure effect	−0.56**	−0.31**	−0.25**	−0.13**	−0.17**	0.03**	−0.41**	−0.15**	−0.25**
	[0.017]	[0.020]	[0.017]	[0.006]	[0.007]	[0.008]	[0.009]	[0.026]	[0.023]
Composition effects									
Education	0.27**	0.16**	0.12**	0.11**	0.06**	0.05**	0.16**	0.11**	0.05**
	[0.004]	[0.003]	[0.002]	[0.004]	[0.002]	[0.001]	[0.003]	[0.002]	[0.002]
Potential Experience	−0.03**	−0.01**	−0.01**	−0.02**	−0.01**	−0.01**	−0.01**	−0.00**	−0.00**
	[0.001]	[0.000]	[0.000]	[0.001]	[0.001]	[0.001]	[0.001]	[0.000]	[0.000]
Formality status	−0.06**	−0.03**	−0.03**	−0.00**	0.00*	−0.01**	−0.07**	−0.03**	−0.04**
	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Race and gender	−0.02**	−0.00**	−0.01**	−0.01**	−0.00**	−0.00**	−0.01**	−0.00**	−0.00**
	[0.001]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]
Region and urban	−0.01**	−0.00**	−0.01**	−0.01**	−0.00**	−0.01**	−0.01**	0	−0.01**
	[0.001]	[0.001]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]
Economic sector	−0.04**	−0.02**	−0.02**	−0.02**	−0.02**	−0.00**	−0.03**	−0.01**	−0.02**
	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Structure effects									
Education	−0.36**	−0.04+	−0.32**	−0.27**	0.03	−0.30**	−0.09**	−0.09**	0
	[0.019]	[0.022]	[0.023]	[0.018]	[0.019]	[0.019]	[0.025]	[0.012]	[0.022]
Potential Experience	−0.66**	−0.38**	−0.29**	−0.33**	−0.16**	−0.17**	−0.33**	−0.21**	−0.12**
	[0.030]	[0.032]	[0.026]	[0.043]	[0.042]	[0.031]	[0.020]	[0.028]	[0.026]
Formality status	0.10**	−0.06**	0.15**	0.11**	−0.03**	0.14**	−0.01	−0.03**	0.03**
	[0.011]	[0.004]	[0.005]	[0.006]	[0.007]	[0.005]	[0.007]	[0.005]	[0.005]
Race and gender	0.05**	−0.03**	0.08**	0.08**	−0.03**	0.10**	−0.03*	0	−0.03**
	[0.010]	[0.008]	[0.005]	[0.011]	[0.010]	[0.007]	[0.013]	[0.011]	[0.006]
Region and urban	0	0.01	−0.01	0.07**	0.04**	0.03*	−0.07**	−0.03*	−0.04**
	[0.016]	[0.016]	[0.009]	[0.018]	[0.015]	[0.014]	[0.013]	[0.012]	[0.012]
Economic sector	0.62**	0.43**	0.20**	0.50**	0.36**	0.14**	0.13*	0.07	0.06**
	[0.069]	[0.067]	[0.022]	[0.062]	[0.066]	[0.022]	[0.064]	[0.048]	[0.021]
Constant	−0.31**	−0.24**	−0.07	−0.30**	−0.38**	0.08+	−0.01	0.14*	−0.15**
	[0.075]	[0.086]	[0.051]	[0.096]	[0.075]	[0.043]	[0.065]	[0.066]	[0.045]
Combined effects									
Education	−0.09**	0.11**	−0.20**	−0.16**	0.09**	−0.25**	0.07**	0.02+	0.05*
	[0.020]	[0.021]	[0.023]	[0.018]	[0.019]	[0.020]	[0.026]	[0.012]	[0.022]
Potential Experience	−0.69**	−0.39**	−0.30**	−0.35**	−0.17**	−0.18**	−0.34**	−0.22**	−0.12**
	[0.029]	[0.032]	[0.026]	[0.043]	[0.041]	[0.031]	[0.020]	[0.028]	[0.025]
Formality status	0.03**	−0.09**	0.12**	0.11**	−0.03**	0.13**	−0.07**	−0.06**	−0.01*
	[0.010]	[0.004]	[0.005]	[0.006]	[0.007]	[0.005]	[0.007]	[0.005]	[0.005]
Race and gender	0.03**	−0.04**	0.07**	0.07**	−0.03**	0.10**	−0.04**	−0.01	−0.03**
	[0.010]	[0.008]	[0.005]	[0.011]	[0.010]	[0.007]	[0.013]	[0.011]	[0.006]
Region and urban	−0.02	0.01	−0.02**	0.07**	0.04**	0.03*	−0.08**	−0.03*	−0.05**
	[0.016]	[0.016]	[0.009]	[0.018]	[0.015]	[0.014]	[0.013]	[0.012]	[0.012]
Economic sector	0.59**	0.41**	0.18**	0.48**	0.34**	0.14**	0.1	0.06	0.04+
	[0.068]	[0.067]	[0.022]	[0.062]	[0.066]	[0.022]	[0.064]	[0.048]	[0.022]
Observations	484,054	484,054	484,054	477,775	477,775	477,775	533,827	533,827	533,827

Source: Pesquisa Nacional por Amostra de Domicílios (PNAD).

Note: Standard errors in brackets. **, *, and + denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. Standard errors are bootstrapped with 100 replications. The log(p90/p10) is the difference between the 90th percentile and the 10th percentile of log monthly earnings. The series log(p90/p50) and log(p50/p10) are calculated analogously. See the note in table 3 for details on the individual effects included in each category. The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 included 2011 and 2012 PNAD samples.

intriguing possibility that the closing of gender gaps observed in the Brazilian labor market might have occurred differentially across sectors of the economy – something that must be left for future research.

Next, the study performs two robustness tests with respect to changes in the headline specification, concerning formality status and the role of minimum wages. It may be argued that a worker's choice of formal status is likely to be affected by wage differentials across sectors, rendering controls for self-employment and formal/informal employee potentially endogenous. We feel that the omitted variable problems that may arise from omitting such an important institutional factor from the analysis would outweigh this concern, but investigate the alternative here. The first three columns of table S1.4 thus exclude the indicator variables for formality status in the overall period (column 1) and the two subperiods (columns 2 and 3). Results for the remaining variables in the model are qualitatively similar to those in the baseline specification, where formality status is controlled for.[39]

Some recent work has attributed an important role in the wage inequality decline to the direct and indirect effects of a rising real minimum wage policy (Maurizio and Vázquez 2016; Engbom and Moser 2017; Alvarez et al. 2018). Indeed, while real mean and median earnings increased by 14 percent and 43 percent respectively between 1995 and 2012, the real minimum wage increased by 103 percent. The growth of the minimum wage was particularly strong between 2003 and 2012 (77 percent). The RIF framework does not naturally lend itself to the study of the effects of the minimum wage on inequality. However, given its prominent place in the recent literature, it is important to assess whether this study's findings survive after attempting to control for the minimum wage, however imperfectly. To investigate this issue, the study includes in columns 4 to 6 of table S1.4 an indicator variable for workers at or above the national minimum wage in any given year.[40] The main findings are essentially unchanged when adding this control.

In addition, for whatever it is worth, column (6) suggests that the rise in the minimum wage did contribute to falling inequality in the 2003–2012 subperiod, in line with earlier work.[41] Engbom and Moser (2017), for example, estimate an equilibrium labor market search model with heterogeneous workers and firms on matched employer-employee data (the RAIS dataset, described in section 3) and find large effects of the increases in Brazil's real minimum wages on the wage distribution in the formal sector of the labor market, during 1996–2012.[42]

Probing the Decline in the Experience Premium

Although they were neglected by previous studies of the determinants of wage inequality in Brazil, changes in the experience premium played a prominent role. In the PNAD the study only observes potential labor market experience, defined as age-education-6. Actual labor market experience cannot be constructed because it is not possible to follow individual workers over time. This gives rise to two potential concerns. First, if the relationship between potential and actual experience is not stable over time (say, if because of increasing female labor force participation people on average spend less time out of the labor force as time goes by), then what the study is interpreting as a change in returns may in fact be a change in the quantity of actual experience (per unit of potential experience). Second, if cohorts differ from one another

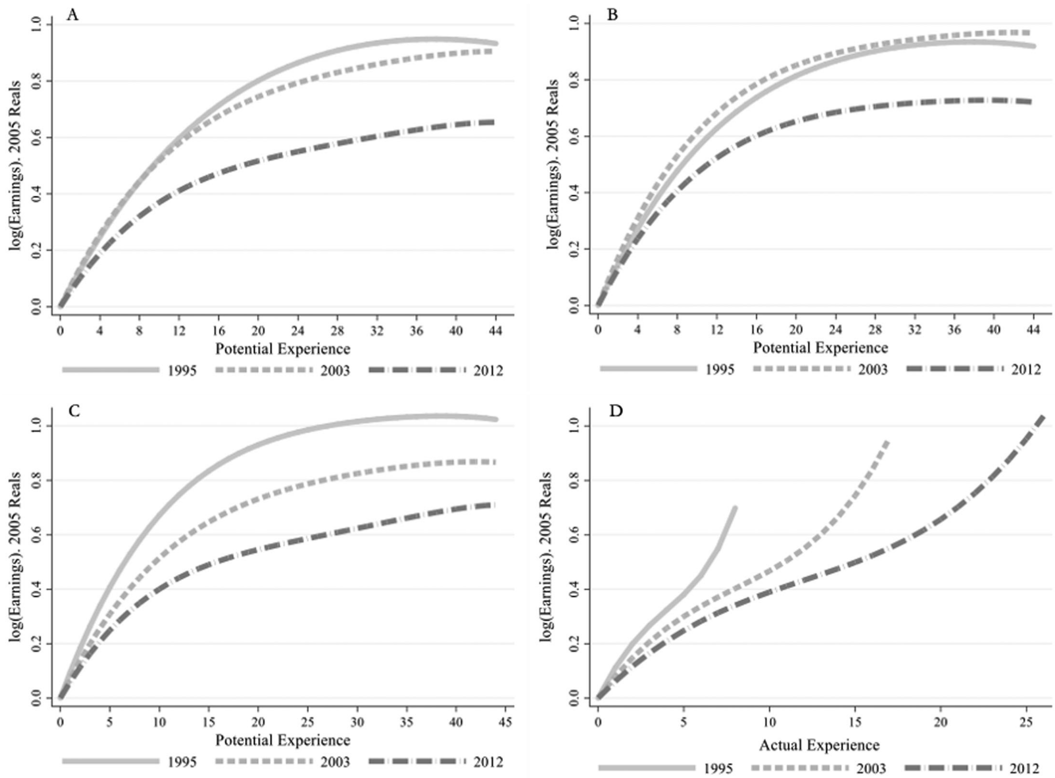
39 But note that the estimate for the structure effect of education for 1995–2012 becomes positive and significant when formality is omitted.

40 Minimum wage values at each point in time are obtained from the ILOSTAT Database.

41 However, during the slow-growth period of 1995–2003 minimum wage increases were associated with a higher Gini. As discussed in Ferreira, Firpo, and Messina (2017), this seems to be driven by the fact that in a period marked by a much softer labor market, the minimum wage increases translated into higher levels of noncompliance and rising self-employment.

42 Using a completely different empirical strategy, Brito (2015) also finds large minimum wage effects on the inequality decline, although she also incorporates the indirect effect of higher minimum wages on social assistance and social security benefits, which are indexed to the minimum wage.

Figure 6. Returns to Experience: Actual and Potential, Formal and Informal, RAIS versus PNAD. (a) Panel A. PNAD. Formal Sector Workers. (b) Panel B. PNAD. Informal and Self-employed Workers. (c) Panel C. RAIS. Potential Experience. (d) Panel D. RAIS. Actual Experience



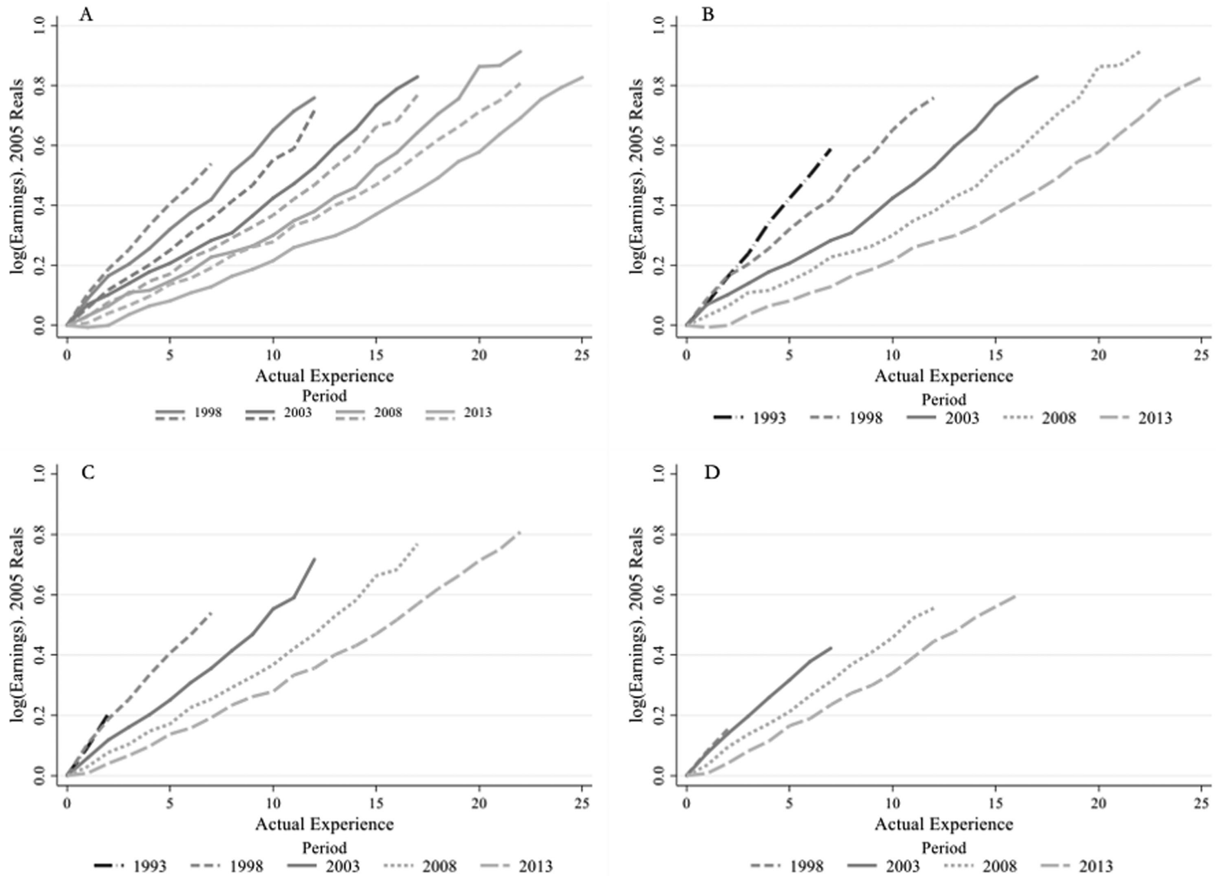
Source: Pesquisa Nacional por Amostra de Domicílios – PNAD (panels A, B) and Relação Anual de Informações Sociais – RAIS (panels C, D).
Note: Panel A. All formal employees of ages 18–65. Panel B. All informal (including self-employment) employees of ages 18–65. Panel C. All formal employees of ages 18–65. Panel D. All formal employees older than 18 who were born after 1968. All regressions include a quartic polynomial in education, a quartic polynomial in years of potential (panels A, B, and C) or actual (panel D) labor market experience, gender, region (northeast, north, southeast, south, and center-west), and sector dummies. Panels A and B additionally include race dummies ("Mestiço," Afro-Brazilian, white, and others) and a dummy variable for living in an urban area. Economic sectors in PNAD are Agriculture, Fishing, Mining and Quarrying, Manufacturing Industries, Electricity, Gas and Water, Construction, Trade, Hotels and Restaurants, Transport and Storage, Real Estate, Public Administration, Teaching, Social Health, Community Services, Domestic Service, and Extra-Territorial Org. Economic sectors in RAIS are Agriculture, Commerce, Construction, Mining, Industry, Public Services, Public administration, Services, and Others.

in systematic ways (e.g., in the quality of education they received) then, as the cohort composition of the labor force changes over time, the consequences of those cohort differences may be misinterpreted as changes in pay structure.

The study seeks to address these two concerns by re-examining the behavior of returns to experience in the RAIS administrative dataset, which was described in section 3. RAIS data do not cover informal or self-employed workers. So, the study first compares the evolution of returns to *potential experience* from 1995 to 2003 and 2012 for formal and informal sector workers, estimated separately in the PNAD data. These separate earnings-experience profiles are shown in panels A and B of fig. 6. Panel C in fig. 6 depicts the same profile, still using potential experience, but estimated for all workers in the RAIS samples for the same years (and thus only for the formal sector). Finally, panel D restricts the RAIS sample to workers born after 1968, and uses their actual experience instead.[43]

43 The age restriction makes it possible to compute the entire work histories of these workers, and hence their actual years of experience in the formal sector. See section 3 for a discussion.

Figure 7. Actual Labor Experience Premiums by Birth Cohort and Year in the Formal Sector. (a) Panel A. Cohort 1968–1972 vs. 1973–1977. (b) Panel B. Cohort 1968–1972. (c) Panel C. Cohort 1973–1977. (d) Panel D. Cohort 1978–1982



Source: Relação Anual de Informações Sociais – RAIS – yearly information from 1986 to 2013.

Note: Formal employees older than 18 who are born after 1968. Panel A: The solid and dashed lines correspond to 1968–1972 and 1973–1977 cohorts, respectively. Panels B, C, and D correspond to 1968–1972, 1973–1977, and 1978–1982 birth cohorts, respectively. Each line represents the coefficients of dummy variables for every year of actual labor experience (with respect to 0 years) in yearly regressions that include a quartic polynomial in education, a dummy variable for every year of age, gender dummy, region (northeast, north, southeast, south, and center-west), and sector dummies (Agriculture, Commerce, Construction, Mining, Industry, Public Services, Public administration, Services, Others) as controls.

A comparison of panels A and B suggests that the decline in returns to potential experience in the first subperiod may have been less pronounced (and possibly non-existent) for informal and self-employed workers. But the 2003–2012 decline is very marked in both sectors in the PNAD data, and so is the overall decline. In the RAIS data (panel C), the decline in returns to potential experience is similar to that observed in the PNAD (but more marked during 1995–2003). When using actual experience (and restricting the age range of the sample) in panel D, returns still decline markedly over time. The main difference between panels C and D – i.e., moving from potential to actual experience – is that returns become much less concave. In fact, they become convex at higher experience levels.[44] Thus, returns to actual experience declined. Still, movements over time in the actual experience premium probably reflect a combination of age, time, and cohort effects that the study tries to disentangle next.

A great advantage of working with actual experience is that it is possible to separate age effects from movements in the experience premium. Panel A of fig. 7 looks at less heterogeneous sets of workers, by plotting the earnings–actual experience profiles for two different cohorts (workers born in 1968–1972; and those born in 1973–1977), at four different points in time: 1998, 2003, 2008, and 2013. For greater comparability, these are five-year cohorts, examined at five-year intervals. Each line represents the coefficients of dummy variables for every year of actual labor experience in regressions that include a quartic polynomial in education, a dummy variable for every year of age, a gender dummy, region, and sector controls. Thus, the cohort-specific experience profiles are purged of age effects.

Panel A of fig. 7 reveals the presence of both cohort and period (or time) effects. The decline of the experience premium over time is evident for the two cohorts (solid line for 1968–1972; dashed line for 1973–1977): returns are lower in later years. Cohort effects are also present but operate in the opposite direction: at each point in time (same color), the younger cohort has higher returns for any given amount of actual experience than the older cohort. Thus, the secular decline in the returns to experience is not a cohort effect.

The cohort-specific time effect of a secular decline in returns to actual experience in Brazil's formal sector is illustrated even more starkly by panels B, C, and D of fig. 7, where each panel contains log earning – actual experience profiles for a single cohort: panel B for the 1968–1972 cohort; panel C for the 1973–1977 cohort; and panel D for the 1978–1982 cohort. In each panel, the lines correspond to the estimated returns in a particular year: 1993, 1998, 2003, 2008, and 2013 after the same set of control variables are purged. It is immediately obvious that returns to actual experience are falling over time rather markedly, for all three cohorts.

The evidence portrayed graphically in fig. 7 suggests that the fall in returns to labor market experience, identified as a key proximate driver of the decline in wage inequality in the PNAD sample, is a real phenomenon. It is certainly not an age effect, or an artifact of spurious differences across cohorts. It is not related to using data on potential rather than on actual experience. It is present in both the formal and informal sectors, in public as well as in private employment, and for men as well as women.

6. Conclusions

After decades of stable or rising income inequality, the period between 1995 and 2012 saw a steady decline in income dispersion in Brazil (at least when top incomes not captured by household surveys are ignored). While increases in the volume and improvements in the targeting of social transfers have played a role in that decline, perhaps its most important driver has been a reduction in inequality in labor earnings:

44 In results not shown to save space, it is found that these patterns are remarkably similar for males and females, when estimated separately by gender. They are also qualitatively similar for public and private sector workers, though the decline is somewhat more marked in the private sector.

between 1995 and 2012 the Gini coefficient for earnings fell by 18 percent, and other measures, such as the Theil index and the p90-p10 ratio, declined by between 30 percent and 40 percent.

The dominant narrative in the literature has so far attributed this decline primarily to educational dynamics: a substantial increase in years of schooling among working-age adults has translated into a rising supply of skills, followed by a decline in the returns to those skills in the labor market (revealing, presumably, that relative demand for skills has failed to keep pace with supply). Those falling returns to education were often credited with driving the decline in wage inequality.

The present analysis draws on RIF regression-based decompositions, which allow for separately investigating the relative roles of a broad set of potential determinants, including labor market experience, changes in the relative extent of formal employment; changes in demographic characteristics of the labor force (chiefly race and gender); and in the sectoral and spatial distribution of employment.

The study finds that the decline in earnings inequality between 1995 and 2012 was driven primarily by changes in the structure of remuneration in the Brazilian labor market, rather than directly by changes in the distribution of worker characteristics. But the main equalizing change in the structure of remuneration was the reduction in the experience – not the education – premium. Sixty percent (92 percent) of the reduction in the Gini (p90-p10 earnings ratio) during 2003–2012 can be attributed to reductions in the returns to potential experience. This decline in the experience premium is observed across the formal, informal, public and private sectors; is robust to using data on actual experience in the formal sector using RAIS; and is observed for multiple workers' birth cohorts. It did not act alone: declines in racial, gender, regional, and urban-rural conditional wage gaps contributed to lower wage inequality. But it certainly was the main driver, and since the overall supply of experience did not change much over the period, it must have reflected a decline in demand for it in the labor market.

The underlying causes of this reduction in the demand for experience are beyond the scope of this paper but invite further research. A smaller gap between older and younger cohorts has been documented elsewhere, and some have attributed it to changes in technology that are biased towards younger workers. Behaghel and Greenan (2010), for example, document that the uptake of information technology training opportunities among workers in high-productivity French firms is lower for older than for younger workers – a phenomenon they relate to age-biased technical change.

There may be other explanations: sharp changes in terms of trade after 2003 altered the demand for labor in Brazil (Messina and Silva 2018). Such a process generates employment and wage opportunities for those workers able to move from "losing" to "winning" sectors and regions. But worker mobility may be costlier for experienced workers, because they may face more difficulties to retrain, or a higher opportunity cost in general.[45] Given how robust the evidence on falling returns to labor market experience in Brazil is, and its importance in accounting for lower wage inequality, a better understanding of its underlying causes certainly seems like an important research priority.

Offsetting the impact of lower returns to experience on the wage distribution, the period of study saw a substantial inequality-augmenting effect of the increase in years of schooling across the population. This is the so-called paradox of progress effect, whereby a rightward movement in the distribution of years of schooling shifts population density to steeper segments of the earnings-education profile, leading to wider earnings gaps. Thus, contrary to the conventional view, the educational upgrade, fully considered, was actually inequality-increasing in Brazil over this period. There was a reduction in the schooling premium, and it did have important equalizing effects, particularly during 2003–2012, but it was more than offset by the inequality-augmenting compositional changes. This suggests that otherwise highly desirable educational policies may have unintended mechanical consequences on wage inequality. Inequality-averse

45 Dix-Carneiro (2014) finds some evidence of greater welfare losses and adjustment costs (from trade liberalization) for older workers, and that these are persistent over time.

policymakers that are aware of this mechanism may be able to better design compensatory policies to mitigate the inequality increases.

References

Acemoglu, Daron. 2002. “Technical Change, Inequality and the Labour Market.” *Journal of Economic Literature* 40: 7–72.

Acemoglu, Daron, and David Autor. 2011. “Skills, Tasks and Technologies: Implications for Employment and Earnings.” *Handbook of Labor Economics* 4: 1043–1171.

Acosta, Pablo, Guillermo Cruces, Sebastián Galiani, and Leonardo Gasparini. 2019. “Educational Upgrading and Returns to Skills in Latin America: Evidence from a Supply-Demand Framework.” *Latin America Economic Review* 28 (18): 1–20.

Adão, Rodrigo. 2015. “Worker Heterogeneity, Wage Inequality, and International Trade: Theory and Evidence from Brazil.” Job market paper. Massachusetts Institute of Technology, Cambridge, MA, USA.

Alvaredo, Facundo, and Leonardo Gasparini. 2015. “Recent Trends in Inequality and Poverty in Developing Countries.” In *Handbook of Income Distribution*, edited by A. B. Atkinson and F. Bourguignon, Vol. 2: 697–805. Amsterdam, the Netherlands: Elsevier.

Alvaredo, Facundo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. 2018. *World Inequality Report 2018*. Cambridge: Harvard University Press.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser. 2018. “Firms and the Decline of Earnings Inequality in Brazil.” *American Economic Journal: Macroeconomics* 10 (1): 149–89.

Azevedo, João P., Gabriela Inchauste, and Viviane Sanfelice. 2013. “Decomposing the Recent Inequality Decline in Latin America.” Working Paper No. 6715. World Bank, Washington, DC.

Barros, Ricardo, Mirela Carvalho, Samuel Franco, and Rosane Mendonça. 2010. “Markets, the State and the dynamics of inequality in Brazil.” In *Declining Inequality in Latin America: A Decade of Progress?* edited by Luis F. López Calva and Nora Lustig, 134–74. Washington, DC: Brookings Institution and UNDP.

Behaghel, Luc, and Nathalie Greenan. 2010. “Training and Age-Biased Technical Change.” *Annals of Economics and Statistics* 99/100: 317–42.

Blinder, Alan S. 1973. “Wage Discrimination: Reduced Form and Structural Estimates.” *Journal of Human Resources* 8: 436–55.

Bourguignon, François, Francisco Ferreira, and Nora Lustig. 2005. *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. Washington, DC: Oxford University Press and World Bank Publications.

Brito, Alessandra. 2015. “O Papel do Salário Mínimo na Redução da Desigualdade na Distribuição de Renda no Brasil entre 1995 e 2013.” Doctoral dissertation, Universidade Federal Fluminense, Niterói, Brazil.

Costa, Francisco, Jason Garred, and João Paulo Pessoa. 2016. “Winners and Losers from a Commodities-for-Manufactures Trade Boom.” *Journal of International Economics* 102: 50–69.

Cowell, Frank A., and Stephen P. Jenkins. 1995. “How Much Inequality Can We Explain? A Methodology and an Application to the USA.” *Economic Journal* 105 (429): 421–30.

Cruz, Gabriela F. D., and Rudi Rocha. 2018 “Efeitos do FUNDEF/B sobre frequência escolar, fluxo escolar e trabalho infantil: uma análise com base nos Censos de 2000 e 2010.” *Estudos Econômicos (São Paulo)* 48 (1): 39–75.

De Koning, Jaap, and Arie Gelderblom. 2004. “ICT and Older Workers: No Unwrinkled Relationship.” *International Journal of Manpower* 27 (5): 467–90.

Dix-Carneiro, Rafael. 2014. “Trade Liberalization and Labor Market Dynamics.” *Econometrica* 82 (3): 825–85.

Dix-Carneiro, Rafael, and Brian K. Kovak. 2017. “Trade Liberalization and Regional Dynamics.” *American Economic Review* 107 (10): 2908–46.

Engbom, Niklas, and Christian Moser. 2017. “Earnings Inequality and the Minimum Wage: Evidence from Brazil.” Working Paper No. 6393. CESifo Economic Studies, Center for Economic Studies and Ifo Institute for Economic Research, Munich, Germany.

Essama-Nssah, Boniface, and Peter J. Lambert. 2012. “Chapter 6 Influence Functions for Policy Impact Analysis.” In *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber*, in *Research on Economic Inequality*, edited by John A. Bishop and Rafael Salas, Vol. 20: 135–59. Somerville, MA: Emerald Group Publishing Limited.

Ferreira, Francisco H. G. 2012. "Distributions in Motion: Economic Growth, Inequality, and Poverty Dynamics." In *The Oxford Handbook of the Economics of Poverty*, edited by Philip N. Jefferson, 427–62. New York: Oxford University Press.

Ferreira, Francisco H. G., Sergio Firpo, and Julián Messina. 2017. "*Ageing poorly? Accounting for the Decline in Earnings Inequality in Brazil, 1995–2012*." Working Paper No. 8018. The World Bank Washington, DC.

Ferreira, Francisco H. G., Phillippe G. Leite, and Julie A. Litchfield. 2008. "The Rise and Fall of Brazilian Inequality: 1981–2004." *Macroeconomic Dynamics* 12 (2): 199–230.

Ferreira, F. H. G., J. Messina, J. Rigolini, L. F. López-Calva, M. A. Lugo, and R. Vakis. 2013. *Economic Mobility and the Rise of the Latin American Middle Class*. Washington, DC: The World Bank.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77 (3): 953–73.

Fortin, Nicole M., Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition Methods in Economics." *Handbook of Labor Economics* 4: 1–102.

Friedberg, Leora. 2003. "The Impact of Technological Change on Older Workers: Evidence from Data on Computer Use." *Industrial and Labor Relations Review* 56 (3): 511–29.

Gardeazabal, Javier, and Arantza Ugidos. 2004. "More on Identification in Detailed Wage Decompositions." *Review of Economics and Statistics* 86 (4): 1034–6.

Gasparini, Leonardo, Sebastian Galiani, Guillermo Cruces, and Pablo Acosta. 2011. " Educational Upgrading and Returns to Skills in Latin America: Evidence from a Supply-Demand Framework, 1990–2010. Policy Research working paper, no. WPS 5921. World Bank. https://openknowledge.worldbank.org/handle/10986/3696.

Goldin, Claudia, and Lawrence Katz. 2008. *The Race Between Education and Technology*. Cambridge: Harvard University Press.

Hampel, Frank. 1974. "The Influence Curve and Its Role in Robust Estimation." *Journal of the American Statistical Association* 69 (346): 383–93.

Jacobson, Margaret, and Filippo Occhino. 2012. "Labor's Declining Share of Income and Rising Inequality." Economic Commentary, 2012–13 (Federal Reserve Bank of Cleveland).

Jaume, David. 2017. "The Labor Market Effects of an Educational Expansion: A Theoretical Model with Applications to Brazil." Working Paper No. 220. Universidad Nacional de La Plata. Buenos Aires, Argentina.

Knight, John B., and Richard H. Sabot. 1983. "Educational Expansion and the Kuznets Effect." *American Economic Review* 73 (5): 1132–6.

Lam, David. 1999. "Generating Extreme Inequality: Schooling, Earnings, and Intergenerational Transmission of Human Capital in South Africa and Brazil." Research Report 99-439. Population Studies Center, University of Michigan, Ann Arbor.

López-Calva, Luis F., and Nora Lustig. 2010. *Declining Inequality in Latin America: A Decade of Progress?* Washington, DC: Brookings Institution and UNDP.

Lustig, Nora, Luis F. López-Calva, and Eduardo Ortiz-Juárez. 2013. "Declining Inequality in Latin America in the 2000s: The Cases of Argentina, Brazil, and Mexico." *World Development* 44: 129–41.

Maurizio, Roxana. 2014. *Labor Formalization and Declining Inequality in Argentina and Brazil in the 2000s: A Dynamic Approach*. Geneva: ILO.

Maurizio, Roxana, and Gustavo Vázquez. 2016. "Distribution Effects of the Minimum Wage in Four Latin American Countries: Argentina, Brazil, Chile and Uruguay." *International Labour Review* 155 (1): 97–131.

Messina, Julian, and Joana Silva. 2018. *Wage Inequality in Latin America. Understanding the Past to Prepare for the Future*. Washington, DC: World Bank Publications.

Messina, Julian, and Joana Silva. 2019. *Twenty Years of Wage Inequality in Latin America*. Washington, DC: The World Bank.

Mookherjee, Dilip, and Anthony F. Shorrocks. 1982. "A Decomposition Analysis of the Trend in UK Income Inequality." *Economic Journal* 92 (368): 886–902.

Morelli, Salvatore, Timothy Smeeding, and J. Thompson. 2015. "Post-1970 Trends in Within-Country Inequality and Poverty: Rich and Middle-Income Countries." In *Handbook of Income Distribution*, edited by A. B. Atkinson and F. Bourguignon, Vol. 2: 593–6. Amsterdam, the Netherlands: Elsevier.

Ñopo, Hugo. 2012. *New Century, Old Disparities: Gender and Ethnic Earnings Gaps in Latin America and the Caribbean*. Washington, DC: World Bank Publications.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14 (3): 693–709.

Oaxaca, Ronald, and Michael R. Ransom. 1999. "Identification in Detailed Wage Decompositions." *Review of Economics and Statistics* 81 (1): 154–7.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge: The Belknap Press of Harvard University Press.

Ravallion, Martin, and Shaohua Chen. 2003. "Measuring Pro-Poor Growth." *Economics Letters* 78: 93–9.

Stiglitz, Joseph. 2012. *The Price of Inequality: How Today's Divided Society Endangers Our Future*. New York: W.W. Norton & Company.

Wang, Yang, Nora Lustig, and Otavio Bartalotti. 2016. "Decomposing Changes in Male Wage Distribution in Brazil." *Research in Labor Economics* 44: 49–78.

World Bank. 2016. *Taking on Inequality: Poverty and Shared Prosperity Report 2016*. Washington, DC: World Bank Publications.

Yun, Myeong-Su. 2005. "A Simple Solution to the Identification Problem in Detailed Wage Decompositions." *Economic Inquiry* 43: 766–72.

Supplementary Online Appendix

Labor Market Experience and Falling Earnings Inequality in Brazil: 1995–2012

Francisco H. G. Ferreira, Sergio P. Firpo, and Julián Messina

S1. Robustness Checks

Table S1.1. Robustness Decomposition Results. Changes in Inequality. Gini. 2012–1995.

	Monthly earnings without sampling weights (1)	Monthly earnings males (2)	Hourly earnings (3)	Hourly earnings, full time workers (more than 35 hours) (4)	Removing self-employed (5)
Overall					
Gini coefficient, end of period	41.192**	39.975**	41.731**	39.896**	39.211**
	[0.065]	[0.095]	[0.075]	[0.088]	[0.088]
Gini coefficient, beginning period	49.851**	48.612**	50.097**	49.431**	48.382**
	[0.071]	[0.087]	[0.069]	[0.080]	[0.083]
Difference	−8.659**	−8.638**	−8.366**	−9.535**	−9.171**
	[0.096]	[0.128]	[0.102]	[0.119]	[0.121]
Composition effect	1.323**	0.457**	1.105**	1.571**	2.233**
	[0.058]	[0.075]	[0.059]	[0.072]	[0.073]
Structure effect	−9.982**	−9.095**	−9.470**	−11.107**	−11.404**
	[0.103]	[0.133]	[0.109]	[0.123]	[0.127]
Composition effects					
Education	3.215***	2.652***	2.878***	4.056***	4.153**
	[0.055]	[0.071]	[0.059]	[0.074]	[0.073]
Potential experience	−0.303***	−0.328***	−0.318***	−0.410***	−0.314**
	[0.011]	[0.016]	[0.012]	[0.015]	[0.015]
Formality status	−0.736***	−0.687***	−0.547***	−0.775***	−0.629**
	[0.016]	[0.019]	[0.014]	[0.018]	[0.016]
Race and gender	−0.107***	−0.0455***	−0.130***	−0.231***	−0.198**
	[0.010]	[0.011]	[0.011]	[0.015]	[0.014]
Region and urban	−0.103***	−0.145***	−0.127***	−0.146***	−0.031*
	[0.014]	[0.019]	[0.013]	[0.015]	[0.014]
Economic sector	−0.643***	−0.990***	−0.651***	−0.923***	−0.747**
	[0.022]	[0.031]	[0.023]	[0.029]	[0.030]
Structure effects					
Education	−0.00103	0.337	−0.436	−0.294	0.119
	[0.265]	[0.314]	[0.291]	[0.311]	[0.342]
Potential experience	−4.938***	−5.243***	−4.729***	−5.017***	−4.647**
	[0.557]	[0.852]	[0.601]	[0.698]	[0.640]
Formality status	0.841***	0.838***	0.631***	0.481***	0.516**
	[0.096]	[0.135]	[0.106]	[0.108]	[0.073]
Race and gender	−1.407***	−0.818***	−1.176***	−1.551***	−1.844**
	[0.172]	[0.115]	[0.175]	[0.181]	[0.217]
Region and urban	−0.966**	−0.793*	−0.0229	−0.392	−0.800*
	[0.300]	[0.362]	[0.287]	[0.324]	[0.331]
Economic sector	4.568**	−0.927	9.455***	6.971***	4.784**
	[1.402]	[2.124]	[1.473]	[1.887]	[1.560]
Constant	−8.079***	−2.489	−13.19***	−11.31***	−9.531**
	[1.579]	[2.364]	[1.662]	[2.086]	[1.787]
Combined effects					
Education	3.214***	2.989***	2.442***	3.763***	4.271**
	[0.265]	[0.321]	[0.297]	[0.318]	[0.348]
Potential experience	−5.241***	−5.571***	−5.046***	−5.426***	−4.961**
	[0.558]	[0.852]	[0.601]	[0.698]	[0.639]
Formality status	0.104	0.151	0.0835	−0.293**	−0.113
	[0.097]	[0.134]	[0.104]	[0.106]	[0.073]

Table S1.1. Continued

	Monthly earnings without sampling weights (1)	Monthly earnings males (2)	Hourly earnings (3)	Hourly earnings, full time workers (more than 35 hours) (4)	Removing self-employed (5)
Race and gender	−1.514***	−0.863***	−1.306***	−1.782***	−2.042**
	[0.173]	[0.118]	[0.178]	[0.186]	[0.220]
Region and urban	−1.069***	−0.938**	−0.150	−0.538	−0.832*
	[0.300]	[0.361]	[0.286]	[0.323]	[0.330]
Economic sector	3.926**	−1.917	8.803***	6.048**	4.037**
	[1.402]	[2.126]	[1.473]	[1.887]	[1.561]
N	484,054	286,294	481,017	390,680	369,348

Source: Pesquisa Nacional por Amostra de Domicílios (PNAD).
Note: Robust standard errors in brackets. **, *, and + denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. Standard errors are calculated with the delta method. The Gini coefficient is expressed in percentage points and goes from 0 (perfect equality) to 100 (perfect inequality). Full-time workers in column (5) are those working more than 35 hours in the reference week. See the note in table 3 for details on the individual effects included in each category. The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 includes 2011 and 2012 PNAD samples.

Table S1.2. Robustness Decomposition Results. Changes in Inequality. Gini. 2003–1995

	Monthly earnings without sampling weights (1)	Monthly earnings males (2)	Hourly earnings (3)	Hourly earnings, full-time workers (more than 35 hours) (4)	Removing self-employed (5)
Overall					
Gini coefficient, end of period	47.011**	45.905**	48.202**	47.181**	45.023**
	[0.068]	[0.090]	[0.075]	[0.088]	[0.085]
Gini coefficient, beginning period	49.851**	48.612**	50.097**	49.431**	48.382**
	[0.071]	[0.095]	[0.076]	[0.088]	[0.092]
Difference	−2.841**	−2.708**	−1.896**	−2.250**	−3.358**
	[0.098]	[0.131]	[0.107]	[0.125]	[0.126]
Composition effect	0.537**	0.223**	0.437**	0.662**	0.815**
	[0.045]	[0.059]	[0.047]	[0.059]	[0.056]
Structure effect	−3.378**	−2.931**	−2.333**	−2.912**	−4.173**
	[0.093]	[0.125]	[0.102]	[0.116]	[0.120]
Composition effects					
Education	1.035***	0.932***	1.157***	1.685***	1.605**
	[0.041]	[0.053]	[0.045]	[0.057]	[0.055]
Potential experience	−0.257***	−0.272***	−0.267***	−0.328***	−0.336**
	[0.010]	[0.013]	[0.010]	[0.014]	[0.015]
Formality status	0.110***	0.0940***	0.0462***	−0.0342**	0.172**
	[0.014]	[0.015]	[0.010]	[0.012]	[0.014]
Race and gender	−0.0353***	−0.00426	−0.0504***	−0.0845***	−0.071**
	[0.007]	[0.004]	[0.006]	[0.008]	[0.009]
Region and urban	0.0535***	−0.0425**	−0.0845***	−0.0611***	−0.063**
	[0.012]	[0.015]	[0.011]	[0.012]	[0.013]
Economic sector	−0.369***	−0.483***	−0.364***	−0.516***	−0.491**
	[0.019]	[0.025]	[0.020]	[0.024]	[0.026]

Table S1.2. Continued

	Monthly earnings without sampling weights (1)	Monthly earnings males (2)	Hourly earnings (3)	Hourly earnings, full-time workers (more than 35 hours) (4)	Removing self-employed (5)
Structure effects					
Education	1.091***	1.750***	1.357***	1.665***	1.594**
	[0.252]	[0.298]	[0.274]	[0.306]	[0.344]
Potential experience	–0.609	–0.813	–1.350	–1.295	–0.999
	[0.600]	[0.959]	[0.691]	[0.806]	[0.725]
Formality status	1.162***	1.026***	1.028***	0.681***	0.555**
	[0.106]	[0.152]	[0.119]	[0.123]	[0.083]
Race and gender	–1.159***	–0.506***	–1.050***	–1.260***	–1.357**
	[0.168]	[0.107]	[0.171]	[0.174]	[0.210]
Region and urban	–0.835**	–0.661	–0.428	–0.610	0.001
	[0.307]	[0.365]	[0.303]	[0.338]	[0.335]
Economic sector	2.138	–3.933	6.563***	1.831	3.152+
	[1.473]	[2.113]	[1.623]	[2.078]	[1.650]
Constant	–5.167**	0.206	–8.454***	–3.923	–7.119**
	[1.659]	[2.396]	[1.841]	[2.319]	[1.910]
Combined effects					
Education	2.126***	2.681***	2.515***	3.350***	3.199**
	[0.255]	[0.302]	[0.278]	[0.312]	[0.347]
Potential experience	–0.865	–1.085	–1.617*	–1.623*	–1.335+
	[0.600]	[0.960]	[0.691]	[0.806]	[0.725]
Formality status	1.272***	1.120***	1.075***	0.646***	0.727**
	[0.107]	[0.153]	[0.120]	[0.124]	[0.085]
Race and gender	–1.194***	–0.510***	–1.101***	–1.345***	–1.428**
	[0.168]	[0.107]	[0.171]	[0.175]	[0.210]
Region and urban	–0.782*	–0.704	–0.513	–0.671*	–0.062
	[0.308]	[0.365]	[0.303]	[0.338]	[0.335]
Economic sector	1.769	–4.416*	6.199***	1.315	2.660
	[1.474]	[2.114]	[1.624]	[2.079]	[1.650]
N	477,775	288,006	476,930	383,776	355,465

Source: Pesquisa Nacional por Amostra de Domicílios (PNAD).

Note: Robust standard errors in brackets. **, * and + denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. Standard errors are calculated with the delta method. The Gini coefficient is expressed in percentage points and goes from 0 (perfect equality) to 100 (perfect inequality). Full-time workers in column (5) are those working more than 35 hours in the reference week. See the note in table 3 for details on the individual effects included in each category. The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD sample, and the year 2012 includes 2011 and 2012 PNAD sample.

Table S1.3. Robustness Decomposition Results. Changes in Inequality. Gini. 2012–2003.

	Monthly earnings without sampling weights (1)	Monthly earnings males (2)	Hourly earnings (3)	Hourly earnings, full-time workers (more than 35 hours) (4)	Removing self-employed (5)
Overall					
Gini coefficient, end of period	41.192**	39.975**	41.731**	39.896**	39.211**
	[0.065]	[0.096]	[0.076]	[0.089]	[0.089]
Gini coefficient, beginning period	47.011**	45.905**	48.202**	47.181**	45.023**
	[0.068]	[0.084]	[0.070]	[0.082]	[0.079]
Difference	−5.818**	−5.930**	−6.470**	−7.285**	−5.813**
	[0.094]	[0.128]	[0.103]	[0.121]	[0.119]
Composition effect	0.994**	0.535**	1.040**	1.358**	1.640**
	[0.048]	[0.062]	[0.047]	[0.060]	[0.059]
Structure effect	−6.812**	−6.465**	−7.510**	−8.643**	−7.452**
	[0.092]	[0.122]	[0.102]	[0.116]	[0.113]
Composition effects					
Education	2.611***	2.207***	2.270***	3.059***	3.031**
	[0.045]	[0.057]	[0.046]	[0.059]	[0.056]
Potential experience	−0.109***	−0.117***	−0.0990***	−0.139***	−0.049**
	[0.009]	[0.010]	[0.008]	[0.011]	[0.012]
Formality status	−0.911***	−0.829***	−0.627***	−0.760***	−0.847**
	[0.016]	[0.020]	[0.014]	[0.017]	[0.017]
Race and gender	−0.0878***	−0.0543***	−0.104***	−0.184***	−0.138**
	[0.007]	[0.008]	[0.008]	[0.010]	[0.010]
Region and urban	−0.154***	−0.0990***	−0.0387***	−0.0857***	0.028**
	[0.009]	[0.016]	[0.009]	[0.011]	[0.010]
Economic sector	−0.356***	−0.572***	−0.362***	−0.532***	−0.385**
	[0.017]	[0.018]	[0.014]	[0.018]	[0.021]
Structure effects					
Education	−1.523***	−1.900***	−2.342***	−2.646***	−1.959**
	[0.271]	[0.328]	[0.308]	[0.334]	[0.356]
Potential experience	−4.267***	−4.369***	−3.330***	−3.664***	−3.577**
	[0.473]	[0.720]	[0.538]	[0.618]	[0.547]
Formality status	−0.257**	−0.139	−0.365***	−0.179	0.006
	[0.091]	[0.131]	[0.103]	[0.104]	[0.072]
Race and gender	−0.232	−0.299*	−0.101	−0.254	−0.475*
	[0.175]	[0.119]	[0.183]	[0.193]	[0.222]
Region and urban	−0.133	−0.135	0.401	0.219	−0.797*
	[0.291]	[0.355]	[0.286]	[0.323]	[0.318]
Economic sector	2.512	3.071	2.966*	5.265**	1.762
	[1.397]	[2.293]	[1.468]	[1.873]	[1.586]
Constant	−2.913	−2.694	−4.739**	−7.383***	−2.412
	[1.539]	[2.470]	[1.638]	[2.051]	[1.767]
Combined effects					
Education	1.088***	0.307	−0.0724	0.413	1.072**
	[0.272]	[0.334]	[0.311]	[0.338]	[0.361]
Potential experience	−4.376***	−4.486***	−3.429***	−3.803***	−3.626**
	[0.474]	[0.720]	[0.538]	[0.619]	[0.547]
Formality status	−1.168***	−0.968***	−0.991***	−0.940***	−0.840**
	[0.092]	[0.130]	[0.103]	[0.104]	[0.072]
Race and gender	−0.320	−0.353**	−0.205	−0.438*	−0.614**
	[0.175]	[0.121]	[0.185]	[0.195]	[0.223]
Region and urban	−0.287	−0.234	0.362	0.133	−0.769*
	[0.291]	[0.355]	[0.285]	[0.323]	[0.318]

Table S1.3. Continued

	Monthly earnings without sampling weights (1)	Monthly earnings males (2)	Hourly earnings (3)	Hourly earnings, full-time workers (more than 35 hours) (4)	Removing self-employed (5)
Economic sector	2.156	2.499	2.604	4.733*	1.377
	[1.397]	[2.293]	[1.468]	[1.873]	[1.586]
N	533,827	311,868	531,549	428,740	408,749

Source: Pesquisa Nacional por Amostra de Domicílios (PNAD).

Note: Robust standard errors in brackets. **, *, and + denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. Standard errors are calculated with the delta method. The Gini coefficient is expressed in percentage points and goes from 0 (perfect equality) to 100 (perfect inequality). Full-time workers in column (5) are those working more than 35 hours in the reference week. See the note in table 3 for details on the individual effects included in each category. The year 1995 includes 1995 and 1996 PNAD samples, the year 2003 includes 2002 and 2003 PNAD samples, and the year 2012 includes 2011 and 2012 PNAD samples.

Table S1.4. Robustness Decomposition Results. Changes in Inequality. Gini

	Excluding formality status			Minimum wage and formality status		
	2012–1995 (1)	2003–1995 (2)	2012–2003 (3)	2012–1995 (4)	2003–1995 (5)	2012–2003 (6)
Overall						
Gini coefficient, end of Period	40.819**	46.742**	40.819**	40.819**	46.742**	40.819**
	[0.076]	[0.073]	[0.077]	[0.075]	[0.072]	[0.076]
Gini coefficient, beginning period	49.767**	49.767**	46.742**	49.767**	49.767**	46.742**
	[0.071]	[0.078]	[0.067]	[0.071]	[0.078]	[0.067]
Difference	−8.948**	−3.025**	−5.923**	−8.948**	−3.025**	−5.923**
	[0.104]	[0.107]	[0.102]	[0.103]	[0.106]	[0.102]
Composition effect	1.516**	0.179**	1.647**	4.013**	2.847**	1.784**
	[0.059]	[0.043]	[0.046]	[0.066]	[0.053]	[0.054]
Structure effect	−10.464**	−3.204**	−7.570**	−12.961**	−5.872**	−7.707**
	[0.110]	[0.104]	[0.099]	[0.106]	[0.100]	[0.094]
Composition effects						
Education	2.969**	1.016**	2.353**	4.114**	1.613**	3.105**
	[0.059]	[0.041]	[0.045]	[0.060]	[0.044]	[0.047]
Potential experience	−0.331**	−0.269**	−0.108**	−0.395**	−0.322**	−0.132**
	[0.013]	[0.011]	[0.010]	[0.013]	[0.012]	[0.011]
Minimum wage				1.134**	1.839**	−0.671**
				[0.029]	[0.032]	[0.032]
Formality status				−0.319**	0.001	−0.226**
				[0.012]	[0.006]	[0.011]
Race and gender	−0.106**	−0.017*	−0.089**	−0.274**	−0.113**	−0.194**
	[0.012]	[0.007]	[0.008]	[0.012]	[0.008]	[0.009]
Region and urban	−0.237**	−0.135**	−0.094**	−0.026**	−0.028**	0.006
	[0.016]	[0.014]	[0.012]	[0.009]	[0.006]	[0.005]
Economic sector	−0.778**	−0.416**	−0.415**	−0.222**	−0.143**	−0.104**
	[0.025]	[0.021]	[0.017]	[0.022]	[0.018]	[0.014]
Structure effects						
Education	0.961**	1.546**	−0.985**	−0.163	2.144**	−2.911**
	[0.288]	[0.276]	[0.300]	[0.258]	[0.249]	[0.263]
Potential experience	−4.354**	−1.227+	−3.081**	−3.122**	0.692	−3.754**
	[0.602]	[0.679]	[0.516]	[0.577]	[0.652]	[0.488]
Minimum wage				0.069**	0.348**	−0.313**
				[0.026]	[0.024]	[0.032]

Table S1.4. Continued

	Excluding formality status			Minimum wage and formality status		
	2012–1995 (1)	2003–1995 (2)	2012–2003 (3)	2012–1995 (4)	2003–1995 (5)	2012–2003 (6)
Formality status				–1.138**	–0.806**	–0.426**
				[0.112]	[0.125]	[0.108]
Race and gender	–1.659**	–0.988**	–0.671**	–1.587**	–1.415**	–0.139
	[0.181]	[0.174]	[0.185]	[0.177]	[0.169]	[0.180]
Region and urban	–0.441	0.101	–0.550+	–1.268**	–0.697*	–0.576*
	[0.294]	[0.300]	[0.286]	[0.285]	[0.289]	[0.275]
Economic sector	2.958*	1.169	1.841	2.162	0.266	1.921
	[1.482]	[1.564]	[1.517]	[1.472]	[1.548]	[1.502]
Constant	–7.930**	–3.805*	–4.124*	–7.914**	–6.404**	–1.509
	[1.671]	[1.776]	[1.665]	[1.646]	[1.749]	[1.635]
Combined effects						
Education	3.930**	2.562**	1.368**	3.951**	3.757**	0.194
	[0.293]	[0.278]	[0.304]	[0.263]	[0.252]	[0.268]
Potential experience	–4.685**	–1.496*	–3.189**	–3.516**	0.369	–3.886**
	[0.601]	[0.679]	[0.516]	[0.577]	[0.652]	[0.488]
Minimum wage				1.203**	2.187**	–0.984**
				[0.040]	[0.040]	[0.044]
Formality status				–1.457**	–0.805**	–0.652**
				[0.111]	[0.125]	[0.106]
Race and gender	–1.765**	–1.005**	–0.760**	–1.862**	–1.528**	–0.333+
	[0.184]	[0.174]	[0.187]	[0.180]	[0.169]	[0.182]
Region and urban	–0.678*	–0.034	–0.644*	–1.294**	–0.724*	–0.570*
	[0.294]	[0.300]	[0.285]	[0.285]	[0.289]	[0.275]
Economic sector	2.180	0.753	1.427	1.940	0.123	1.817
	[1.483]	[1.564]	[1.517]	[1.472]	[1.549]	[1.502]
N	484,054	477,775	533,827	484,054	477,775	533,827

Source: Pesquisa Nacional por Amostra de Domicílios (PNAD).

Note: Robust standard errors in brackets. **, *, and + denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. Standard errors are calculated with the delta method. The Gini coefficient is expressed in percentage points and goes from 0 (perfect equality) to 100 (perfect inequality). See the note in table 3 for details on the individual effects included in each category. The year 1995 includes 1995 and 1996 samples, the year 2003 includes 2002 and 2003 samples, and the year 2012 includes 2011 and 2012 samples.