

Arabage, Amanda Cappellazzo; Souza, André Portela

**Article**

## Wage dynamics and inequality in the Brazilian formal labor market

EconomiA

**Provided in Cooperation with:**

The Brazilian Association of Postgraduate Programs in Economics (ANPEC), Rio de Janeiro

*Suggested Citation:* Arabage, Amanda Cappellazzo; Souza, André Portela (2019) : Wage dynamics and inequality in the Brazilian formal labor market, EconomiA, ISSN 1517-7580, Elsevier, Amsterdam, Vol. 20, Iss. 3, pp. 153-190, <https://doi.org/10.1016/j.econ.2019.10.001>

This Version is available at:

<https://hdl.handle.net/10419/266943>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

# Wage dynamics and inequality in the Brazilian formal labor market

Amanda Cappellazzo Arabage<sup>a</sup>, André Portela Souza<sup>b,\*</sup>

<sup>a</sup> FGV EESP Clear – Center for Learning on Evaluation and Results for Brazil and Lusophone Africa, Av. Paulista 542, 3rd floor, São Paulo, SP 01310-000, Brazil

<sup>b</sup> Sao Paulo School of Economics, Fundação Getulio Vargas (FGV EESP), Rua Itapeva 474, 12th floor, Sao Paulo, SP 0331-000, Brazil

Received 5 February 2019; received in revised form 9 August 2019; accepted 29 October 2019

Available online 16 January 2020

## Abstract

This paper investigates the decrease in wage inequality among working age men in the Brazilian formal labor market by decomposing it into its permanent and transitory components. The permanent component refers to workers' productivity characteristics such as education and abilities, whereas the transitory component relates to noise caused by economic instability. We use data from the Annual Reports of Social Information (RAIS) from 1994 to 2016, covering periods of both economic instability and stability. Our proposed model includes year and cohort specific effects in both components and we use minimum distance methods to estimate our parameters of interest. Overall we observe a downward trend in wage inequality levels for most cohorts during the entire period. However, different sources are associated with this decrease in different sub-periods. This decline can be attributed to reductions in the transitory component from 1994 to 2005 (related to the economic stabilization process) and in the permanent component from 2010 to 2016. Moreover, our results suggest that the permanent component has an important role in explaining the wage inequality level in Brazil, especially considering older cohorts, even though its share in terms of total variance has decreased over the 2010s. Finally, we show that age and education are accountable for a great share of permanent inequality, with this percentage being even greater for younger cohorts, although their relative importance also decreased over the years.

*JEL classification:* J13; J16; O5

*Keywords:* Wage inequality; Variance decomposition; Formal labor market

© 2019 The Authors. Production and hosting by Elsevier B.V. on behalf of National Association of Postgraduate Centers in Economics, ANPEC. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

[s.subramanian.2@elsevier.com](mailto:s.subramanian.2@elsevier.com)

## 1. Introduction

Various studies have documented the decrease in income inequality in Latin America over the last decades (e.g. López-Calva and Lustig, 2010; Azevedo et al., 2013). This decrease is highly associated to decreases in labor earnings inequality, but their causes remain subject of debate. Lustig et al. (2013), based on other studies for Argentina, Brazil,

\* Corresponding author.

E-mail addresses: [amanda.arabage@fgv.br](mailto:amanda.arabage@fgv.br) (A.C. Arabage), [andre.portela.souza@fgv.br](mailto:andre.portela.souza@fgv.br) (A.P. Souza).

Peer review under responsibility of National Association of Postgraduate Centers in Economics, ANPEC.

<https://doi.org/10.1016/j.econ.2019.10.001>

1517-7580 © 2019 The Authors. Production and hosting by Elsevier B.V. on behalf of National Association of Postgraduate Centers in Economics, ANPEC. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

and Mexico, presents evidence that a very important factor in explaining the downward trend in inequality is the fall in the skill premium which they argue seems to be associated to both increases in the supply of skilled labor and decreases in the relative demand for this type of worker.

Manacorda et al. (2010) argues that the decline in the skill premium in Argentina, Brazil, Chile, Colombia, and Mexico during the 1980s and 1990s can be explained by a substantial increase in the supply of workers with secondary-level education that led to reduced wages of this group relative to workers with primary-level education. There is also evidence that increased openness to trade had an important role in declining skill premium in the 1990s (e.g. Gonzaga et al., 2006; Ferreira et al., 2007), as well as the boom in commodity prices, technological diffusion and skill mismatches that could have favored non-tertiary educated workers in the 2000s (Gasparini et al., 2011).

We aim to contribute to the debate about the Brazilian wage inequality decrease by investigating the dynamics of wages and its covariance structure. We rely on a large body of literature focused on evaluating the roles played by the permanent and transitory components of wages in explaining the evolution of inequality over time. Besides helping understand the functioning of the labor market, provided evidence may also contribute for public policy design.

In this literature, the permanent component refers to workers' productivity characteristics such as education and abilities, whereas the transitory component relates to noise caused by economic instability. Wage inequality changes over time driven by permanent or transitory components have different implications. For example, a decrease in inequality resulting mainly from a decrease in the permanent component is associated with higher mobility in the labor market (i.e. workers are able to change positions in the wage distribution over their lifetimes) and both short and long-run decreases in inequality (Ramos, 2003; Santos and Souza, 2007; Gustavsson, 2008). This could be the case if there was a change in demand in favor of less skilled workers with lower lifetime earnings (due to increased openness to trade, for example) that would cause changes in the wage distribution (in particular, in the positions occupied by each worker) so that it would lead to a decrease in the permanent component of variances and, ultimately, in inequality.

By contrast, a decrease in the transitory component is related to inequality decreases in the short-run and have little or no effect on long-run inequality (Baker et al., 2003; Gustavsson, 2008). Possible causes for changes in the transitory component include macroeconomic shocks that affect workers differently, changes in competition and weaker worker-firm attachment (Haider, 2001), changes in regulations or in union activities and even changes in the demand for temporary jobs and for self-employment (Moffitt and Gottschalk, 2012). In the case of technological shocks, both components might change, as the permanent component can be affected if the shock alters works abilities and/or the returns of skills and human capital and the transitory component can be affected if the shock modifies the returns for job occupations.

These different implications of the permanent and transitory components should also be kept in mind when designing public policy for reducing inequality. If there is evidence that the permanent component of wage inequality is predominant, policies oriented toward enhancing worker qualifications (e.g. formal and technical education, trainings) might be recommended.

On the other hand, if the transitory component of wage inequality has relatively greater importance, policies focused on correcting market imperfections and on mitigating the effects of negative shocks that workers may suffer in the labor market may be more suitable (Cappellari, 2004; Santos and Souza, 2007). Some examples could be providing safety nets and unemployment insurance.

Evidence on the contributions of permanent and transitory components to earnings inequality is available for several countries such as United States (e.g. Lillard and Willis, 1978; MaCurdy, 1982; Abowd and Card, 1989; Baker, 1997; Gottschalk et al., 1994; Haider, 2001; Moffitt and Gottschalk, 2012; DeBacker et al., 2013; Lochner et al., 2018), Canada (Baker et al., 2003; Beach et al., 2010), Great Britain (Dickens, 2000; Ramos, 2003; Kalwij and Alessie, 2007; Li, 2012), Germany (Myck et al., 2009; Bartels and Bönke, 2010; Bönke et al., 2015), Italy (Cappellari, 2004), Spain (Cervini-Plá and Ramos, 2012), Sweden (Gustavsson, 2008), Luxembourg (Sologon and O'Donoghue, 2010), Finland (Kässi, 2014), Norway (Blundell et al., 2014), European Union countries (Sologon and O'Donoghue, 2009, 2012), Venezuela (Freije et al., 2002) and China (Zhao, 2007). Considering the Brazilian case, Santos and Souza (2007) provides evidence for the State of São Paulo and Arabage and Souza (2015)<sup>1</sup> for Brazil.

<sup>1</sup> An earlier version of this paper is registered as Arabage (2013).

We investigate the decrease in wage inequality of working age men in the Brazilian formal labor market by decomposing it into its permanent and transitory components. We use data from the Annual Reports of Social Information (RAIS) that covers the period from 1994 to 2016, meaning that we observe periods of both economic instability (e.g. second half of the 1990s) and stability (e.g. mid-2000s). Individuals in our sample are divided into thirteen birth-years cohorts and in our model we account for year and cohort specific effects in both permanent and transitory components.

Although our results are restricted to the formal labor market, in the next section we illustrate that it covers the majority of workers in Brazil and that the trend in terms of inequality evolution is relatively similar to that for informal workers. Regardless of this limitation, our results shed some light on the roles played by each inequality component in changing wage inequality in Brazil.

Our results suggest that the permanent component has an important role in explaining the wage inequality level in Brazil for all considered cohorts, with its relative importance being greater for the older ones, even though its share in terms of total variance has decreased over the 2010s. Considering the contribution of each component for changes in inequality, the wage inequality reduction from 1994 to 2005 was completely driven by reductions in the transitory component which are related to the economic stabilization process. From 2005 to 2010, there was an increase in wage inequality for all cohorts that was caused by transitory variance increases, with younger cohorts experiencing higher inequality growth. Finally, from 2010 to 2016, for the cohorts that experienced a decrease in wage inequality this can be entirely attributed to the reduction in the permanent component.

The paper is divided into seven sections, including this introduction. Section 2 presents stylized facts about the economic performance and evolution of wage inequality in Brazil in the last decades. Section 3 presents our proposed model and describes the estimation procedure. Section 4 describes our data, sample, and cohorts' definitions and also provides descriptive statistics, including the evolution of wage inequality per cohort. Section 5 displays our results, and Section 6 presents discussions about the changes in inequality and the role played by age and education. Lastly, Section 7 concludes.

## 2. Context and wage inequality in Brazil

Brazil experienced strong macroeconomic instability during the 1980s and 1990s with high levels of inflation and a stagnant economy. Various attempts were made in order to control inflation, but it was only in 1994 with the Real Plan that these efforts succeeded. The plan and other economic policies in the second half of the 1990s included the introduction of a new currency (the Real), privatizations, government reforms, increased openness to trade, liberalizations in the financial sector, and inflation-targeting policies (de Souza, 2012). Some of these structural reforms were also observed in other countries in Latin America in the same period (see Edwards, 1996; Lora, 2012).

In the 2000s, Brazil experienced better economic performance with higher growth rates (an average of approximately 4% annually, with the exception of 2009, the year of the global economic crisis) as a result from favorable institutional and macroeconomic environment (high commodities prices boosted Brazilian exports, foreign investment was elevated and expansionary fiscal policies were introduced (de Souza, 2012)). In the 2010s, however, Brazil has experienced economic recession, high inflation rates, and unemployment growth in a context of an economic and political crisis and corruption scandals. Signs of economic recovery recently started to show.

Since 1994 and especially during the 2000s, there was an expressive earnings inequality decrease in Brazil, considering both total household income and labor income (Lustig et al., 2013). Ferreira et al. (2016) shows that the household per capita income's Gini coefficient fell from 0.59 in 1995 to 0.52 in 2012, while the labor income's Gini coefficient had an ever greater decline from 0.50 in 1995 to 0.40 in 2012. The authors also provide evidence that this decline in wage inequality is not sensitive to the inequality measure used, as they find similar trends when considering the Theil (T) index and the 90-10 percentile ratio. Barros et al. (2010) attributes this wage inequality reduction to the increase in educational attainment and the declining schooling wage premia during the 2000s, with the latter having greater importance. Ferreira et al. (2016) presents evidence that inequality reductions in other dimensions such as gender, race, geographical region and formal/informal work as well as the rise in minimum wages also contributed to the inequality decline. Souza et al. (2018) presents evidence that there was a reduction in inequality from 1995 to 2005 due to both: (i) increases in labor income related to formalization, especially among lower-income workers; and (ii) increases in non-labor income related to the implementation of cash transfers social policies.

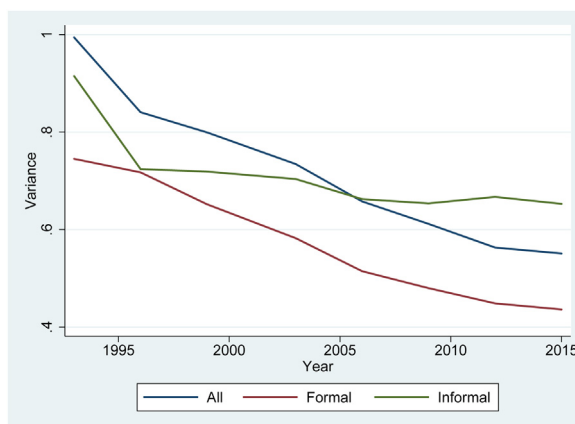


Fig. 1. Log of real wages' variance – Brazil, selected years.

Source: Elaborated by the authors. Data: PNAD.

We use data from the Brazilian National Household Survey (PNAD/IBGE) for selected years to investigate the evolution of wage inequality in the formal and informal labor markets<sup>2</sup> using the variance of the log of real wages as a measure of inequality. Results are presented in Fig. 1 and reinforce the decrease in inequality over these years, which is observed both in the formal and informal labor markets, although is more pronounced for the former. It is worth mentioning that the majority of male workers in Brazil are in the formal labor market and during the analyzed period their proportion increased from 72% in 1993 to 78% in 2015.<sup>3</sup>

In this paper, we analyze the evolution of wage inequality focusing on the formal labor market.<sup>4</sup> We consider the 1994–2016 period and use the wage variance as an inequality measure, which is decomposed it into its permanent and transitory components in order to shed some light on the which was the relative importance of each one in explaining the observed inequality decrease.

### 3. Model and estimation

Models in this literature decompose covariances into two components: permanent and transitory. By modeling a labor income function based on permanent and transitory elements, variances and covariances of earnings can be analyzed over time, which allows the importance of each component to be assessed as well as their contributions to changes in inequality over time. To avoid participation self-selection bias issues, we consider only economically active males aged 25–64, which is discussed further in Section 4.

#### 3.1. Covariance structure

We start by presenting the *canonical model* of this literature, which was first presented in Lillard and Willis (1978) and is described in further detail in Ramos (2003). Although it is a very simple and restricted model, it is helpful for the understanding of the intuition of the analysis.

Let  $y_{it}$  be the log wage of individual  $i$  in year  $t$  and consider that it corresponds to the sum of a time-invariant individual component  $\mu_i$  and a transitory component  $v_{it}$ . As described in Ramos (2003), the permanent component refers to characteristics assumed to be persistent over the analyzed period, whether measurable (e.g. education) or

<sup>2</sup> Selected years are 1993, 1996, 1999, 2003, 2006, 2009, 2012 and 2015. The sample includes only men who were 25–64 years old and who were working during the reference week of the Brazilian National Household Survey (PNAD/IBGE). Formal waged workers are those with a formal labor contract, who work for the government or who are in the military. All wage values are in 2015 Reais (R\$), using the INPC deflator for the PNAD income, as proposed by Corseuil and Foguel (2002).

<sup>3</sup> Sample definition information is available in the previous footnote.

<sup>4</sup> We focus on the formal labor market because of the need to follow the same individuals through time.

not (e.g. abilities, preferences). The transitory component refers to economic instability and is assumed to be serially uncorrelated (white noise). We have:

$$y_{it} = \mu_i + v_{it} \quad (1)$$

$$\mu_i \sim iid(\bar{\mu}_i, \sigma_\mu^2) \quad (2)$$

$$v_{it} \sim iid(0, \sigma_v^2) \quad (3)$$

Assuming both components are orthogonal ( $Cov(\mu_i, v_{it}) = 0$ ), we have the following covariance structure:

$$Cov(y_{it}, y_{is}) = \begin{cases} \sigma_\mu^2 + \sigma_v^2 & \text{for } t = s \\ \sigma_\mu^2 & \text{for } t \neq s \end{cases} \quad (4)$$

We use the variance of log wages as a measure of inequality, following papers in the inequality decomposition into permanent and transitory components literature. We notice that the variance of the permanent component ( $\sigma_\mu^2$ ) that represents the persistent dispersion of wages fully determines covariances between different time periods, as the transitory component is assumed to be white noise.

Because this model makes use of very strong hypothesis and restrictions about the dynamic structure of wages, most studies in this literature proposed more sophisticated and complex models. The main modifications are: (i) including cohort shifters in one or both components (e.g. Baker et al., 2003; Ramos, 2003; Kalwij and Alessie, 2007; Moffitt and Gottschalk, 2012); (ii) in the permanent component, including year-specific factor loadings (e.g. Gustavsson, 2008; Kässi, 2014) or considering a random walk model (e.g. Ramos, 2003) or a random growth model (e.g. Cappellari, 2004; Moffitt and Gottschalk, 2011); and (iii) in the transitory component, consider there to be serial correlation and treat it as a time series process such as an AR( $p$ ) or an ARMA( $p, q$ ) (e.g. MaCurdy, 1982; Baker, 1997; Dickens, 2000; Haider, 2001; Ramos, 2003; Baker et al., 2003; Moffitt and Gottschalk, 2012).

In this paper, we consider cohort and year-specific shifters in both components. Specifically, we have:

$$y_{ict} = \rho_t \lambda_c \mu_i + \gamma_t \kappa_c v_{it} \quad (5)$$

where  $\rho_t$  refers to year  $t$  permanent component shifter,  $\lambda_c$  refers to cohort  $c$  permanent component shifter,  $\mu_i$  is still the time-invariant individual component,  $\gamma_t$  refers to year  $t$  transitory component shifter,  $\kappa_c$  refers to cohort  $c$  transitory component shifter and  $v_{it}$  is the transitory component. The shifters of the permanent component relate to the skill premium, as they capture human capital returns.

For the transitory component, we investigate three possibilities: (i) consider it to be white noise (*canonical model with factor loadings*); (ii) assume it follows an AR(1) process (*AR(1) model with factor loadings*); and (iii) assume it follows an ARMA(1, 1) process (*ARMA(1, 1) model with factor loadings*). In all cases considered, we assumed that: (a)  $\mu_i \sim iid(\bar{\mu}_i, \sigma_\mu^2)$ ; (b) there is no correlation between shifters ( $\rho_t, \gamma_t, \lambda_c$  and  $\kappa_c$ ) which are also uncorrelated to  $\mu_i$  and  $v_{it}$ ; and (c)  $Cov(\mu_i, v_{it}) = 0$ . In cases (2) and (3) we also assume that the time series process in the transitory component is stationary.

For the canonical model with factor loadings (possibility (i)), we also have that  $v_{it} \sim iid(0, \sigma_v^2)$ . In this case, the covariance structure is given by:

$$Cov(y_{ict}, y_{ics}) = \begin{cases} \rho_t^2 \lambda_c^2 \sigma_\mu^2 + \gamma_t^2 \kappa_c^2 \sigma_v^2 & \text{for } t = s \\ \rho_t \rho_s \lambda_c^2 \sigma_\mu^2 & \text{for } t \neq s \end{cases} \quad (6)$$

In this equation,  $\rho_t^2 \lambda_c^2 \sigma_\mu^2$  is the permanent component and  $\gamma_t^2 \kappa_c^2 \sigma_v^2$  is the transitory component of inequality when  $t = s$ .

For the AR(1) model with factor loadings (possibility (ii)), the additional assumptions are:

$$v_{it} = \phi v_{it-1} + \epsilon_{it} \quad (7)$$

$$\epsilon_{it} \sim N(0, \sigma_\epsilon^2) \quad (8)$$

$$Cov(\epsilon_{it}, \epsilon_{is}) = 0 \quad \text{for } t \neq s \quad (9)$$

Table 1  
Summary of estimated models.

Model	Equation
(1) Canonical model	(4)
(2) Canonical model with factor loadings	(6)
(3) Model with factor loadings and the transitory component following an AR(1) process	(11)
(4) Model with factor loadings and the transitory component following an ARMA(1, 1) process	(15)

Source: Elaborated by the authors.

Including parameter  $\sigma_{v_{-1}}^2$ , which corresponds to the variance of the transitory component ( $v_{it}$ ) in the period immediately before the first period in our analysis, we have:

$$y_{ict} = \rho_t \lambda_c \mu_i + \gamma_t \kappa_c \left( \phi^{t+1} v_{i,-1} + \sum_{j=0}^t \phi^{t-j} \epsilon_{ij} \right) \tag{10}$$

In this case, the covariance structure is given by:

$$\begin{aligned} \text{Var}(y_{ict}) &= \rho_t^2 \lambda_c^2 \sigma_\mu^2 + \gamma_t^2 \kappa_c^2 \phi^{2(t+1)} \sigma_{v_{-1}}^2 + \gamma_t^2 \kappa_c^2 \sum_{j=0}^t \phi^{2(t-j)} \sigma_\epsilon^2 \\ \text{Cov}(y_{ict}, y_{ict+s}) &= \rho_t \rho_{t+s} \lambda_c^2 \sigma_\mu^2 + \gamma_t \gamma_{t+s} \kappa_c^2 \phi^{2(t+1)+s} \sigma_{v_{-1}}^2 + \gamma_t \gamma_{t+s} \kappa_c^2 \sum_{j=0}^t \phi^{2(t-j)+s} \sigma_\epsilon^2 \end{aligned} \tag{11}$$

Thus, the permanent component is  $\rho_t^2 \lambda_c^2 \sigma_\mu^2$  in  $\text{Var}(y_{ict})$  and  $\rho_t \rho_{t+s} \lambda_c^2 \sigma_\mu^2$  in  $\text{Cov}(y_{ict}, y_{ict+s})$ . Finally, for the ARMA(1, 1) model with factor loadings (possibility (iii)), the additional assumptions are:

$$v_{it} = \phi v_{it-1} + \epsilon_{it} + \theta \epsilon_{it-1} \tag{12}$$

$$\epsilon_{it} \sim N(0, \sigma_\epsilon^2) \tag{13}$$

$$E(\epsilon_{it}, \epsilon_{it-j}) = \begin{cases} \sigma_\epsilon^2 & \text{for } j = 0 \\ 0 & \text{for } j \neq 0 \end{cases} \tag{14}$$

In this case, the covariance structure is given by:

$$\begin{aligned} \text{Var}(y_{ict}) &= \rho_t^2 \lambda_c^2 \sigma_\mu^2 + \gamma_t^2 \kappa_c^2 \sigma_\epsilon^2 \frac{(1 + 2\phi\theta + \theta^2)}{(1 - \phi^2)} \\ \text{Cov}(y_{ict}, y_{ict-h}) &= \rho_t \rho_{t-h} \lambda_c^2 \sigma_\mu^2 + \gamma_t \gamma_{t-h} \kappa_c^2 \phi^{h-1} \sigma_\epsilon^2 \frac{(1 + \phi\theta)(\phi + \theta)}{(1 - \phi^2)} \text{ for } h \geq 1 \end{aligned} \tag{15}$$

Thus, the permanent component is  $\rho_t^2 \lambda_c^2 \sigma_\mu^2$  in  $\text{Var}(y_{ict})$  and  $\rho_t \rho_{t-h} \lambda_c^2 \sigma_\mu^2$  in  $\text{Cov}(y_{ict}, y_{ict-h})$ .

Table 1 summarizes the four estimated models.

To analyze and decompose wage inequality, we can consider wages directly or the residuals of a regression of wages on age and education per year, which allows us to access inequality within age and education groups. In this paper we present results for both wages and residuals so that we can make comparisons between them and also compare our results to those of other countries.

### 3.2. Estimation

We use minimum distance methods to estimate our parameters of interest, following most studies in this literature, so that we can decompose wage inequality into its permanent and transitory components.

The log of real hourly wages (or the residuals of its regression on age and education) is used for calculating<sup>5</sup> the observed covariance matrix  $C$  and the covariance matrix of observed covariances  $V$ . As shown in the previous subsection, the elements of the covariance matrix for wages can be modeled in various ways. Consider any of the previously mentioned models with a vector of parameters  $b$  (whose size is smaller than  $T$ , the number of time periods in the analysis) so that  $m = f(b)$  is the vector of theoretical moments predicted by the model. Using minimum distance methods, we can estimate  $b$  by minimizing:

$$(m - f(b))' A(m - f(b)) \quad (16)$$

where  $A$  is a positive definite weighting matrix. Hence,  $b$  is chosen to reduce the distance between the observed moments  $m$  and the theoretical moments predicted by the model  $f(b)$  as close to zero as possible. For the case of the AR(1) model with factor loadings, the expression to be minimized is indicated in Appendix B as an example.

Following [Dickens \(2000\)](#), let  $R^{-1}$  be the inverse matrix of  $R = PVP'$ , where  $P = I - F(F'AF)^{-1}F'A$  and  $F = F(b^*)$  is the Jacobian matrix assessed in  $b^*$  (the estimated value of  $b$ ):

$$F(b^*) = \frac{\partial f(b)}{\partial b}(b^*) \quad (17)$$

Inference is made based on the following statistic, which, under the null hypothesis that  $m = f(b^*)$  (equivalent to the specification used being correct), has a chi-squared asymptotic distribution (as demonstrated by [Newey, 1985](#)):

$$n[(m - f(b^*))' R^-(m - f(b^*))] \sim \chi_h^2 \quad (18)$$

where  $n$  is the number of individuals in the sample,  $R^-$  corresponds to the generalized inverse<sup>6</sup> of matrix  $R$ , and  $h$  is the number of degrees of freedom, representing the difference between the size of  $m$  and the rank of the Jacobian matrix assessed in  $b^*(F(b^*))$ . This statistic may be used to test the general structure of the model and to compare different methods.

Standard errors of estimated parameters can be obtained by calculating the asymptotic variance of  $b^*$ , thus allowing for  $t$  tests to be conducted for each of the estimated coefficients. According to [Chamberlain \(1984\)](#), under some conditions,  $b^*$  converges to the true value of  $b$  and  $\sqrt{n}(b^* - b)$  converges in distribution to  $N(0, \Omega)$ , where  $\Omega = (F'AF)^{-1}F'AVAF(F'AF)^{-1}$ . We can treat  $b^*$  as if:

$$b^* \sim N(b, \Omega/n) \quad (19)$$

such that the asymptotic variance of  $b^*$  can be estimated using:

$$A \text{Var}(b^*) = [(F'AF)^{-1}F'AVAF(F'AF)^{-1}]/n \quad (20)$$

The choice of  $A$  determines the type of minimum distance to be estimated. One alternative would be the optimal minimum distance method, where  $A = V^{-1}$  and  $V^{-1}$  is the inverse matrix of  $V$ , thereby minimizing:

$$(m - f(b))' V^{-1}(m - f(b)) \quad (21)$$

This method is referred to as optimal because it minimizes the asymptotic variance of  $b^*$ , reducing it to  $\Omega = (F'V^{-1}F)^{-1}$ . However, most studies in this literature prefer the equally weighted minimum distance, where  $A = I$  thereby minimizing  $(m - f(b))'(m - f(b))$  and attributing equal weights for all moments, as there is evidence that the optimal minimum distance method can be biased in small samples ([Altonji and Segal, 1996](#)).

<sup>5</sup> Calculation details are available in Appendix A.

<sup>6</sup> [Rao \(1962\)](#).



As we use an unbalanced panel in which different numbers of observations are used for estimating each moment, attributing equal weights to all moments may not be appropriate. We follow Haider (2001) and choose  $A = \Pi^{-1} \Pi^{-1}$ , where:

$$\Pi = \begin{bmatrix} \frac{n_1}{n} & 0 & \dots & 0 \\ 0 & \frac{n_2}{n} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{n_L}{n} \end{bmatrix} \quad (22)$$

and  $n_l$  corresponds to the number of observations used to calculate moment  $l$ , and  $n$  is the total number of observations. Therefore, larger weights are attributed to moments for which there are proportionally more available observations. In the next section we describe our data and the cohort structure we use.

## 4. Data

### 4.1. Sample

We use the Annual Reports of Social Information (RAIS) data from the Brazilian Ministry of Labor and Employment (MTE), which provides information about the formal labor market only. All firms that legally hire employees must report wages<sup>7</sup> of all their workers to the Ministry on an annual basis,<sup>8</sup> in addition to job contract information (e.g. admission date, occupation, number of hours worked) and individual characteristics such as gender, age, and education.

The RAIS dataset contains information on all individuals that had at least one job during the reference year. As long as individuals stay in the formal labor market, it is possible to follow them through time using their unique identification number from the Brazilian Social Integration Program (PIS). We consider only the main job of each individual.<sup>9</sup> One drawback of using RAIS is that it refers to the formal labor market only, but as the analysis requires longitudinal wage data that allows tracking of individuals over long periods of time, unfortunately, no such data is available for the informal labor market in Brazil. On the other hand, as RAIS is based on administrative data reported by employers, it is less subject to measurement errors than self-reported data on earnings or wages, which is a particularly interesting advantage in this context.

We use data for all years from 1994 to 2016 in our analysis. The reason for starting in 1994 is minimizing problems related to measurement errors since before 1994 wages were reported in different currencies and in some years there was very high inflation, with nominal wages changing monthly or even weekly. We consider only men<sup>10</sup> aged 25–64, as we opt for excluding individuals whose level of education is highly likely to change (those under 25) and individuals that could have retired already according to the age criteria at the time (those over 64). These criteria follow the vast majority of studies in this literature.

In order to define our sample, first, using the PIS identification number, we made a list of all male individuals who participated in the formal labor market at least once from 1994 to 2016 (i.e. had at least one job contract so that there is information about them in RAIS). This list contains more than 69 million individuals. The second step was to draw a random sample of individuals of this list, keeping 5% of them. The final step was to keep in the final adjusted sample only individuals who met the age criteria and for whom there was available wage information<sup>11</sup> in at least one year of

<sup>7</sup> In the RAIS dataset there is no available information about variable compensations such as bonus and equity, so unfortunately these will not be included in the analysis.

<sup>8</sup> Executive Order n. 76.9000 of September 23, 1975.

<sup>9</sup> The main job selection followed the criteria suggested by the Brazilian Ministry of Labor and Employment. In case an individual has more than one job, the main one is considered to be the one which is still active on December 31st or the one that was terminated the last. In case an individual has more than one active job, the main one is considered to be the one with the oldest admission date.

<sup>10</sup> The decision to investigate wage inequality among men is based on the fact that the participation of women in the labor market has peculiarities outside the scope of this study (e.g. need for childcare and domestic work, wage discrimination, cultural reasons).

<sup>11</sup> Wage on December 31st.

Table 2  
Sample selection.

Step	Number of individuals
List of men based on RAIS data	69,366,140
Random sample (5%)	3,468,307
Final adjusted sample	2,412,302

Source: Elaborated by the authors. Data: RAIS.

Table 3  
Cohorts' description.

Cohort number	Birth years	Years observed	Ages observed	Available covariance	Number of individuals
1	1930–1934	1994–1998	60–64	15	13,817
2	1935–1939	1994–2003	55–64	55	30,500
3	1940–1944	1994–2008	50–64	120	53,313
4	1945–1949	1994–2013	45–64	210	89,676
5	1950–1954	1994–2016	40–64	276	137,414
6	1955–1959	1994–2016	35–61	276	184,566
7	1960–1964	1994–2016	30–56	276	235,489
8	1965–1969	1994–2016	25–51	276	278,526
9	1970–1974	1995–2016	25–46	253	306,852
10	1975–1979	2000–2016	25–41	153	337,183
11	1980–1984	2005–2016	25–36	78	362,564
12	1985–1989	2010–2016	25–31	28	307,656
13	1990–1994	2015–2016	25–26	3	74,746
	Total			2,019	2,412,302

Source: Elaborated by the authors. Data: RAIS.

the considered period. [Table 2](#) presents the total number of individuals in each step of the sample selection process. Our final sample consists of 2,412,302 individuals.

For the analysis, an unbalanced panel with a total of 16,752,034 individual-year observations was built based on our sample of individuals. The reasons the panel is unbalanced are: (i) individuals may be out of the formal labor market in some of the years considered (e.g. unemployed, informal worker or self-employed); and (ii) observations (defined as the pair individual-year) are only kept in the panel as long the individual meets the age criteria in that year (i.e. observations of individuals who turn 65 during our period of analysis are not included in the panel, even if there is available information).

We sort individuals into thirteen cohorts based on birth years. [Table 3](#) presents the definition of cohorts as well as the years for which each one is observed, possible ages during this period, number of covariances (moments) calculated and the number of individuals. Cohorts 5, 6, 7 and 8 are the ones that can be observed during all years of the analyzed period, each contributing with 276 moments. In total, we have 2019 covariances to be used in the analysis.

For each cohort, in [Table 4](#) we present the total number of individuals actually observed per year as well as the percentage of these regarding the total number of individuals in the sample for the specific cohort. Taking cohort 1 as an example, in 1994, we observe 11,415 individuals which corresponds to 83% of individuals in this cohort (the total is 13,817, as stated in [Table 3](#)). For cohorts 1–5, we notice that this percentage decreases substantially over the years for which they are observed, as part of the individuals in these groups age out. On the other hand, for cohorts 9–13, this percentage increases over the first few years for which they are observed as part of individuals in these groups start meeting the age criteria (turn 25). In general, the percentages decrease over the years for all cohorts and present relatively similar patterns.

Table 4  
Observed individuals (total and percentage).

Year	Cohort												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Birth years	1930–1934	1935–1939	1940–1944	1945–1949	1950–1954	1955–1959	1960–1964	1965–1969	1970–1974	1975–1979	1980–1984	1985–1989	1990–1994
1994	11,415 83%	21,441 70%	34,445 65%	54,835 61%	78,497 57%	98,741 53%	117,190 50%	126,539 45%					
1995	9,150 66%	20,806 68%	33,387 63%	54,290 61%	79,146 58%	100,000 54%	118,978 51%	129,880 47%	26,325 9%				
1996	6,642 48%	19,147 63%	31,347 59%	51,434 57%	76,829 56%	98,635 53%	118,913 50%	131,240 47%	53,854 18%				
1997	4,186 30%	18,016 59%	29,653 56%	49,205 55%	75,122 55%	98,790 54%	119,773 51%	134,548 48%	84,161 27%				
1998	2,069 15%	16,478 54%	27,532 52%	45,911 51%	71,517 52%	95,906 52%	118,178 50%	133,646 48%	112,787 37%				
1999		15,009 49%	25,726 48%	43,397 48%	68,612 50%	93,405 51%	116,509 49%	132,532 48%	141,903 46%				
2000		11,958 39%	24,733 46%	42,417 47%	67,947 49%	93,836 51%	118,172 50%	135,690 49%	146,676 48%	29,895 9%			
2001		9,030 30%	23,800 45%	41,308 46%	66,880 49%	93,303 51%	118,370 50%	136,589 49%	149,433 49%	62,622 19%			
2002		4,032 13%	15,386 29%	29,025 32%	50,072 36%	72,214 39%	93,194 40%	108,037 39%	119,727 39%	77,089 23%			
2003		1,970 6%	14,353 27%	27,479 31%	48,972 36%	71,671 39%	93,570 40%	108,388 39%	121,507 40%	107,057 32%			
2004			13,024 24%	26,205 29%	47,927 35%	71,767 39%	94,754 40%	111,176 40%	125,395 41%	142,340 42%			

2005	10,091 19%	25,038 28%	47,082 34%	71,570 39%	95,808 41%	112,870 41%	127,654 42%	147,084 44%	31,099 9%		
2006	7,636 14%	23,589 26%	45,767 33%	70,885 38%	96,135 41%	113,621 41%	129,987 42%	152,630 45%	66,555 18%		
2007	5,075 10%	22,518 25%	44,801 33%	70,905 38%	96,892 41%	115,457 41%	132,656 43%	157,030 47%	106,060 29%		
2008	2,593 5%	21,570 24%	44,399 32%	71,225 39%	98,182 42%	117,159 42%	135,281 44%	161,828 48%	146,351 40%		
2009		19,884 22%	42,406 31%	69,684 38%	97,802 42%	117,654 42%	136,313 44%	164,564 49%	185,737 51%		
2010		16,368 18%	41,519 30%	70,016 38%	99,331 42%	120,497 43%	139,791 46%	170,223 50%	195,798 54%	37,993 12%	
2011		7,690 9%	25,631 19%	46,910 25%	71,086 30%	93,650 34%	119,559 39%	160,043 47%	197,262 54%	78,274 25%	
2012		5,739 6%	25,525 19%	45,429 25%	65,851 28%	80,979 29%	94,777 31%	117,195 35%	138,269 38%	81,567 27%	
2013		4,194 5%	34,808 25%	63,961 35%	94,560 40%	117,924 42%	137,642 45%	169,162 50%	199,310 55%	157,767 51%	
2014			32,079 23%	61,214 33%	92,501 39%	116,173 42%	136,554 45%	167,815 50%	198,682 55%	196,125 64%	
2015			23,934 17%	55,262 30%	85,763 36%	109,106 39%	129,597 42%	158,965 47%	188,948 52%	187,511 61%	33,784 45%
2016			16,704 12%	48,807 26%	78,941 34%	102,140 37%	122,470 40%	151,164 45%	180,516 50%	179,783 58%	65,402 87%

Source: Elaborated by the authors. Data: RAIS.

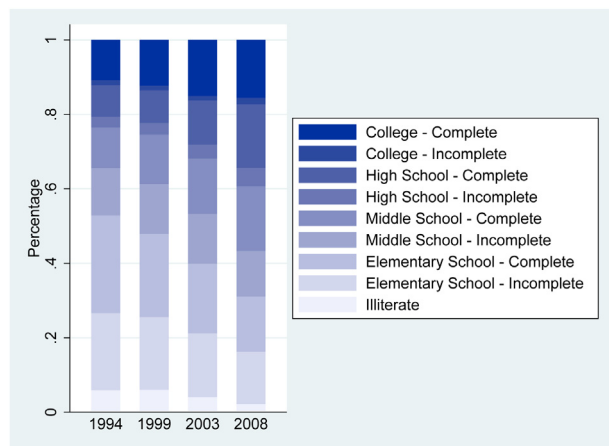


Fig. 2. Education profile of cohort 3 (birth years 1940–1944) over the years.

Source: Elaborated by the authors. Data: RAIS.

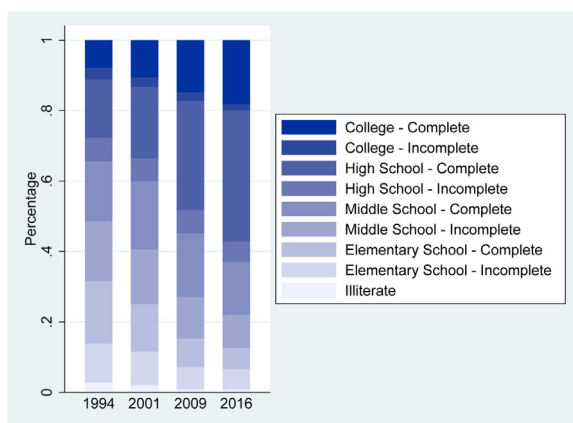


Fig. 3. Education profile of cohort 7 (birth years 1960–1964) over the years.

Source: Elaborated by the authors. Data: RAIS.

#### 4.2. Descriptive statistics

For the analysis, we use: (i) the log of real hourly wages<sup>12</sup>; and (ii) the residuals obtained from regressions of the log of real hourly wages on age and education dummies, which are estimated separately for each cohort and year using ordinary least squares. In both cases, we compute covariances at all possible lag lengths separately for each cohort. Pooling covariance elements of all cohorts gives us 2019 moments (as presented in Table 3). In this subsection, we present descriptive statistics considering this cohort structure.

We start by addressing the evolution of education profiles of selected cohorts over the years, which present similar patterns to those observed for other cohorts.<sup>13</sup> Figs. 2–4 present the evolution of the education profile over the years considered in our analysis by showing the percentage of individuals per level of schooling for cohorts 3 (birth years 1940–1944), 7 (birth years 1960–1964) and 11 (birth years 1980–1984), respectively. In all cases, we notice that the

<sup>12</sup> We use individuals' monthly wages on December 31st reported in minimum wage units. As minimum wage values vary over the years, workers' wages were converted to Brazilian currency units (Reais) using the nominal value of the minimum wage from December of the corresponding year. Hourly real wages were calculated by dividing monthly wages by contracted monthly working hours. Real values were obtained using the Brazilian National Consumer Price Index (*Índice Nacional de Preços ao Consumidor* (INPC) – IBGE) and all values are reported in 2016 Reais (R\$).

<sup>13</sup> In this subsection, some cohorts and years were selected for the sake of brevity. Information for other cohorts and years are available upon request.

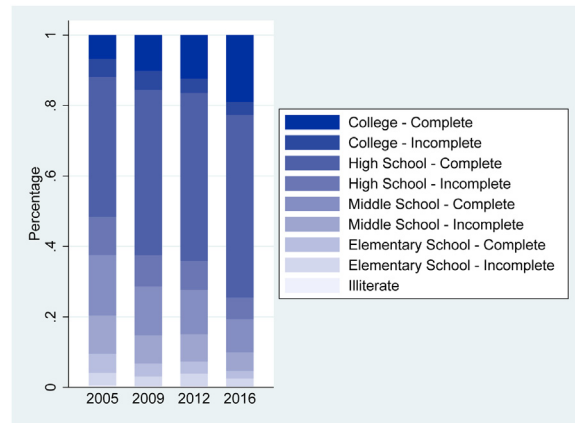


Fig. 4. Education profile of cohort 11 (birth years 1980–1984) over the years.

Source: Elaborated by the authors. Data: RAIS.

percentage of individuals with at most elementary school (either incomplete or complete) decreased substantially over the years while the percentage of individuals who completed high school increased significantly. Another important aspect is the fact that younger cohorts present much higher levels of education than older cohorts considering both high school and college completion. This fact highlights the importance of considering cohort effects in the proposed model.

Table 5 presents means and standard deviations of the log of real hourly wages per cohort and year. For all cohorts, we notice an increase in the mean of log of real hourly wages over the years, except for cohorts 4–9 (birth years 1945–1974) in 2011, when we observe a decrease even though it is not persistent (from 2012 onwards the upward trend is resumed). This decrease in the mean of log of real hourly wages in 2011 might be related to the historically low readjustment of the minimum wage due to the decrease of the gross domestic product in Brazil in 2009 as a reflect of the global economic crisis.<sup>14</sup> Over the analyzed period, standard deviations decreased for older cohorts (e.g. 2, 4, and 6) and increased for younger cohorts (e.g. 9, 10 and 11). When comparing different cohorts, we notice that the mean of log of real hourly wages tends to be shorter for younger cohorts (e.g. 12 and 13) and for the oldest ones (e.g. 1 and 2), in line with what was expect considering the formal labor market, in particular, experience return and retirement possibility.<sup>15</sup>

Tables 6–8 present log of real hourly wages' covariance matrices for cohorts 3 (birth years 1940–1944), 7 (birth years 1960–1964) and 11 (birth years 1980–1984). For cohorts 3 and 7 we observe a decrease in variances and covariances over the years, meaning that there was a inequality reduction for these groups. For cohort 11, however, we observe an increase in the variance and covariances over the years, what can be interpreted as a inequality rise. Besides different evolution patterns, differences in the levels of covariances between cohorts is also remarkable: variances and covariances are much higher among older cohorts. Covariance matrices for residual for these cohorts present similar trends and are available on Appendix C. One aspect to highlight is the difference between levels of covariances of log of real hourly wages and residuals: as expected, levels of covariances of residuals are much shorter as they are net of age and education contributions to permanent inequality.

Figs. 5 and 6 present the evolution of variances of the log of real hourly wages and residuals (respectively) over the analyzed period for each cohort.<sup>16</sup> Patterns are similar in the two figures. For cohorts 1 and 2 (birth years 1930–1939), which are the oldest ones, variances are relatively flat over the years. For cohorts 3–9 (birth years 1940–1974) and 13

<sup>14</sup> By law, the minimum wage readjustment is based on the inflation rate from the previous year and the economic growth experienced two years before. Given the economic contraction in 2009, the readjustment rule had to be adapted.

<sup>15</sup> In Brazil, people can retire based on age or the total number of years they have worked. We excluded individuals who could have retired based on the age criteria from our analysis. It is possible that among the older individuals in our sample are those who have not yet worked enough years to retire (e.g. longer unemployment periods or informal work over their lifetimes) and so decide to stay in the formal labor market, which could be correlated to smaller wages.

<sup>16</sup> For each cohort, a kernel-weighted local polynomial smooth plot of the variance with local mean smoothing was used.

Table 5  
Descriptive statistics: log of real hourly wages.

Year	Cohort												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Birth Years	1930–1934	1935–1939	1940–1944	1945–1949	1950–1954	1955–1959	1960–1964	1965–1969	1970–1974	1975–1979	1980–1984	1985–1989	1990–1994
1994	2.04 (1.01)	2.1 (1.04)	2.28 (1.09)	2.43 (1.1)	2.44 (1.08)	2.35 (1.03)	2.19 (0.96)	1.99 (0.87)	–	–	–	–	–
1995	2.15 (0.99)	2.17 (1.01)	2.34 (1.06)	2.5 (1.09)	2.54 (1.07)	2.46 (1.02)	2.31 (0.96)	2.13 (0.87)	1.98 (0.8)	–	–	–	–
1996	2.17 (1)	2.18 (1.01)	2.32 (1.05)	2.5 (1.08)	2.55 (1.06)	2.47 (1.01)	2.33 (0.95)	2.16 (0.87)	2.02 (0.79)	–	–	–	–
1997	2.2 (0.99)	2.2 (1)	2.32 (1.05)	2.49 (1.07)	2.55 (1.06)	2.5 (1.01)	2.37 (0.95)	2.2 (0.87)	2.05 (0.79)	–	–	–	–
1998	2.23 (1.01)	2.2 (0.99)	2.32 (1.04)	2.47 (1.06)	2.55 (1.06)	2.5 (1.01)	2.38 (0.95)	2.23 (0.87)	2.06 (0.78)	–	–	–	–
1999	–	2.17 (1)	2.26 (1.03)	2.41 (1.05)	2.49 (1.04)	2.44 (0.99)	2.34 (0.94)	2.19 (0.87)	2.01 (0.78)	–	–	–	–
2000	–	2.19 (0.99)	2.28 (1.02)	2.41 (1.04)	2.48 (1.03)	2.45 (0.99)	2.36 (0.94)	2.22 (0.87)	2.06 (0.79)	1.93 (0.72)	–	–	–
2001	–	2.24 (1.01)	2.29 (1.03)	2.41 (1.04)	2.48 (1.03)	2.45 (0.99)	2.36 (0.94)	2.24 (0.88)	2.08 (0.8)	1.95 (0.72)	–	–	–
2002	–	2.18 (0.97)	2.24 (1.02)	2.38 (1.04)	2.47 (1.04)	2.46 (1.01)	2.36 (0.97)	2.22 (0.89)	2.06 (0.81)	1.91 (0.72)	–	–	–
2003	–	2.19 (0.95)	2.23 (0.99)	2.36 (1.02)	2.47 (1.03)	2.46 (1.01)	2.36 (0.96)	2.23 (0.88)	2.08 (0.8)	1.92 (0.7)	–	–	–
2004	–	–	2.26 (0.99)	2.37 (1.02)	2.48 (1.04)	2.48 (1.01)	2.39 (0.96)	2.26 (0.89)	2.12 (0.81)	1.95 (0.71)	–	–	–

2005	–	–	2.31	2.4	2.51	2.5	2.42	2.29	2.17	2.02	1.91	–	–
	–	–	(0.99)	(1.01)	(1.02)	(1)	(0.96)	(0.88)	(0.81)	(0.72)	(0.65)	–	–
2006	–	–	2.4	2.48	2.56	2.56	2.49	2.36	2.25	2.11	2	–	–
	–	–	(0.99)	(1.01)	(1.02)	(1)	(0.95)	(0.88)	(0.81)	(0.73)	(0.65)	–	–
2007	–	–	2.44	2.5	2.58	2.57	2.51	2.4	2.29	2.18	2.05	–	–
	–	–	(0.98)	(1)	(1.01)	(0.99)	(0.95)	(0.88)	(0.82)	(0.75)	(0.66)	–	–
2008	–	–	2.45	2.52	2.59	2.59	2.53	2.43	2.33	2.23	2.1	–	–
	–	–	(0.98)	(1.01)	(1.01)	(0.99)	(0.95)	(0.89)	(0.83)	(0.77)	(0.68)	–	–
2009	–	–	–	2.56	2.62	2.62	2.57	2.47	2.38	2.29	2.16	–	–
	–	–	–	(1)	(1.01)	(0.99)	(0.95)	(0.88)	(0.82)	(0.77)	(0.68)	–	–
2010	–	–	–	2.6	2.64	2.64	2.6	2.51	2.42	2.34	2.23	2.12	–
	–	–	–	(1.01)	(1)	(0.98)	(0.95)	(0.88)	(0.82)	(0.77)	(0.7)	(0.61)	–
2011	–	–	–	2.36	2.37	2.38	2.39	2.37	2.37	2.37	2.3	2.18	–
	–	–	–	(0.84)	(0.84)	(0.82)	(0.81)	(0.79)	(0.78)	(0.78)	(0.72)	(0.63)	–
2012	–	–	–	2.65	2.67	2.66	2.62	2.54	2.46	2.39	2.31	2.19	–
	–	–	–	(1.01)	(1)	(0.99)	(0.94)	(0.88)	(0.82)	(0.77)	(0.72)	(0.62)	–
2013	–	–	–	2.7	2.71	2.72	2.7	2.64	2.58	2.53	2.46	2.31	–
	–	–	–	(0.98)	(0.96)	(0.95)	(0.92)	(0.87)	(0.82)	(0.79)	(0.74)	(0.63)	–
2014	–	–	–	–	2.72	2.73	2.72	2.67	2.62	2.57	2.51	2.35	–
	–	–	–	–	(0.95)	(0.94)	(0.92)	(0.87)	(0.83)	(0.8)	(0.75)	(0.64)	–
2015	–	–	–	–	2.69	2.69	2.69	2.65	2.6	2.55	2.5	2.35	2.22
	–	–	–	–	(0.95)	(0.94)	(0.92)	(0.87)	(0.83)	(0.8)	(0.76)	(0.66)	(0.56)
2016	–	–	–	–	2.71	2.69	2.7	2.66	2.61	2.56	2.52	2.37	2.23
	–	–	–	–	(0.95)	(0.92)	(0.91)	(0.87)	(0.83)	(0.8)	(0.76)	(0.66)	(0.56)

Source: Elaborated by the authors. Data: RAIS.

Note: Mean of the log of real hourly wages. Standard deviations in parentheses.



Table 6  
Covariance matrix of the log of real hourly wages for cohort 3 (birth years 1940–1944).

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
1994	1.184														
1995	1.031	1.128													
1996	0.984	1.035	1.100												
1997	0.963	1.010	1.038	1.096											
1998	0.943	0.987	1.008	1.035	1.076										
1999	0.932	0.974	0.993	1.015	1.022	1.061									
2000	0.913	0.953	0.974	0.994	1.005	1.021	1.047								
2001	0.912	0.947	0.963	0.978	0.988	1.006	1.007	1.054							
2002	0.906	0.914	0.930	0.938	0.951	0.968	0.968	0.995	1.032						
2003	0.865	0.885	0.895	0.900	0.912	0.919	0.926	0.942	0.953	0.974					
2004	0.867	0.882	0.887	0.896	0.907	0.913	0.917	0.934	0.945	0.942	0.980				
2005	0.865	0.873	0.890	0.890	0.900	0.915	0.909	0.931	0.937	0.931	0.948	0.978			
2006	0.851	0.861	0.875	0.879	0.891	0.900	0.896	0.918	0.921	0.925	0.938	0.948	0.971		
2007	0.850	0.842	0.849	0.868	0.886	0.900	0.904	0.915	0.950	0.934	0.946	0.950	0.944	0.969	
2008	0.866	0.856	0.858	0.865	0.874	0.898	0.900	0.901	0.938	0.930	0.931	0.945	0.930	0.940	0.963

Source: Elaborated by the authors. Data: RAIS.

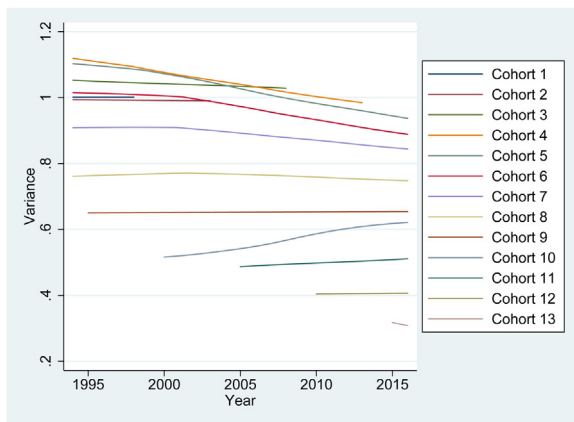


Fig. 5. Variance of the log of real hourly wages – all cohorts.

Source: Elaborated by the authors. Data: RAIS.

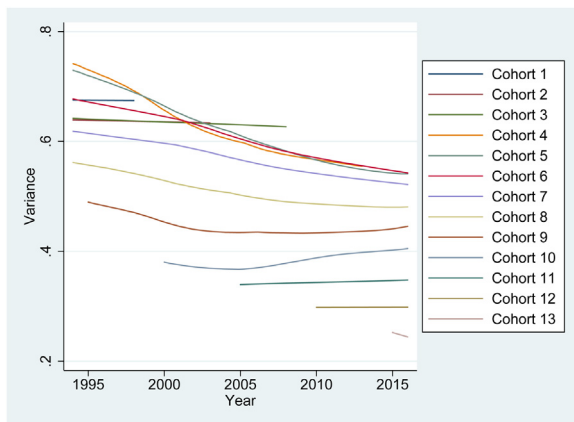


Fig. 6. Variance of the residuals – all cohorts.

Source: Elaborated by the authors. Data: RAIS.

Table 7  
Covariance matrix of the log of real hourly wages for cohort 7 (birth years 1960–1964).

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	
1994	0.928																							
1995	0.808	0.917																						
1996	0.766	0.828	0.902																					
1997	0.743	0.798	0.831	0.910																				
1998	0.727	0.776	0.800	0.840	0.909																			
1999	0.704	0.752	0.772	0.808	0.832	0.887																		
2000	0.686	0.734	0.749	0.784	0.807	0.825	0.888																	
2001	0.676	0.717	0.731	0.768	0.787	0.801	0.824	0.892																
2002	0.704	0.737	0.748	0.782	0.797	0.811	0.833	0.880	0.937															
2003	0.701	0.736	0.745	0.777	0.794	0.800	0.821	0.856	0.880	0.923														
2004	0.699	0.733	0.738	0.771	0.789	0.795	0.814	0.846	0.874	0.892	0.929													
2005	0.685	0.718	0.723	0.753	0.772	0.776	0.796	0.827	0.854	0.873	0.886	0.913												
2006	0.673	0.708	0.712	0.743	0.761	0.763	0.782	0.816	0.846	0.861	0.876	0.881	0.911											
2007	0.665	0.697	0.702	0.733	0.748	0.749	0.768	0.800	0.833	0.849	0.862	0.865	0.881	0.901										
2008	0.660	0.690	0.697	0.726	0.742	0.745	0.762	0.794	0.829	0.843	0.856	0.857	0.872	0.879	0.905									
2009	0.653	0.684	0.689	0.718	0.735	0.736	0.752	0.784	0.820	0.831	0.843	0.846	0.861	0.867	0.880	0.896								
2010	0.647	0.678	0.684	0.712	0.732	0.733	0.747	0.778	0.815	0.824	0.836	0.837	0.853	0.856	0.872	0.878	0.894							
2011	0.492	0.514	0.515	0.532	0.550	0.546	0.554	0.551	0.586	0.601	0.602	0.606	0.606	0.614	0.622	0.630	0.633	0.648						
2012	0.660	0.685	0.689	0.712	0.735	0.730	0.750	0.781	0.817	0.827	0.837	0.836	0.851	0.853	0.865	0.871	0.880	0.600	0.891					
2013	0.617	0.642	0.648	0.678	0.697	0.692	0.710	0.738	0.777	0.787	0.798	0.799	0.809	0.813	0.821	0.827	0.834	0.603	0.867	0.849				
2014	0.614	0.640	0.646	0.673	0.694	0.688	0.706	0.734	0.776	0.786	0.795	0.796	0.807	0.808	0.818	0.822	0.830	0.590	0.864	0.834	0.850			
2015	0.615	0.638	0.644	0.671	0.690	0.685	0.702	0.733	0.778	0.786	0.794	0.795	0.805	0.807	0.815	0.819	0.825	0.574	0.856	0.827	0.837	0.850		
2016	0.603	0.629	0.633	0.657	0.679	0.674	0.690	0.720	0.761	0.768	0.777	0.778	0.792	0.790	0.801	0.804	0.811	0.561	0.844	0.809	0.823	0.827	0.837	

Source: Elaborated by the authors. Data: RAIS.

Table 8

Covariance matrix of the log of real hourly wages for cohort 11 (birth years 1980–1984).

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
2005	0.416											
2006	0.389	0.429										
2007	0.385	0.409	0.440									
2008	0.374	0.402	0.422	0.462								
2009	0.365	0.390	0.409	0.434	0.463							
2010	0.360	0.384	0.401	0.426	0.440	0.489						
2011	0.342	0.370	0.389	0.412	0.424	0.462	0.513					
2012	0.336	0.363	0.377	0.398	0.407	0.441	0.469	0.515				
2013	0.347	0.370	0.387	0.407	0.415	0.447	0.475	0.488	0.546			
2014	0.349	0.375	0.389	0.408	0.413	0.446	0.472	0.484	0.525	0.563		
2015	0.351	0.377	0.388	0.406	0.414	0.445	0.469	0.480	0.517	0.544	0.578	
2016	0.347	0.375	0.385	0.402	0.408	0.440	0.461	0.473	0.504	0.529	0.550	0.571

Source: Elaborated by the authors. Data: RAIS.

(birth years 1990–1994), we observe a decrease in variances from the first to the last year they are observed, although for some cohorts (8 and 9 (birth years 1965–1974)) there were increases in some of the years. Finally, for cohorts 10–12 (birth years 1975–1989), we observe an increase in variances, with cohort 10 (birth years 1975–1979) being the one with the highest inequality rise. Once again, these differences between cohorts' trends emphasize the importance of considering cohort effects in the proposed model.

Figs. 7 and 8 show trends in selected covariances at different lag orders for each cohort<sup>17</sup> for the log of real hourly wages and residuals, respectively.<sup>18</sup> In all cases (i.e. for all cohorts and both variables), low-order covariances present similar trends to those observed for variances while high-order covariances tend to become flatter. The gap between variances and covariances of various lag orders is smaller when considering the log of real hourly wages than residuals, but in both cases there is not much variation between most cohorts in terms of the size of observed gaps.

Regarding the relative magnitude of covariance to variances, we notice that all covariances are large when compared to variances, with this ratio (covariance/variance) decreasing as the lag order increases. As discussed in Dickens (2000), variances reflect both permanent and transitory components while higher order covariances tend to reflect the permanent component only, as transitory effects tend to fade with time. The relatively high levels of high-order covariances observed for all cohorts indicate that the permanent component probably answers for a large share of observed inequality in wages in the formal labor market in Brazil.

Finally, it is worth mentioning that the fact that variances and low-order covariances present either decreasing or increasing trends (depending on the cohort) while high-order covariances' trends are always flatter is a frequently observed pattern in this literature, which also serves as evidence that the transitory component in the proposed model should be accessed as a time series process such as an ARMA( $p, q$ ), as previously discussed (an AR component to address persistence and an MA component to help explain the rapid decline in low-order covariances (Moffitt and Gottschalk, 2011)).

## 5. Results

### 5.1. Estimation results and model selection

In this subsection we present estimation results for our proposed models. Four different models were estimated using minimum distance methods that were previously described: (1) a canonical model; (2) a canonical model with factor loadings (i.e. year and cohort shifters in each component); (3) a model with factor loadings and the transitory

<sup>17</sup> We omitted cohort 13 (birth years 1990–1994) as we can only observe it for three years.

<sup>18</sup> For each cohort, kernel-weighted local polynomial smooth plots of the covariances with local mean smoothing were used.

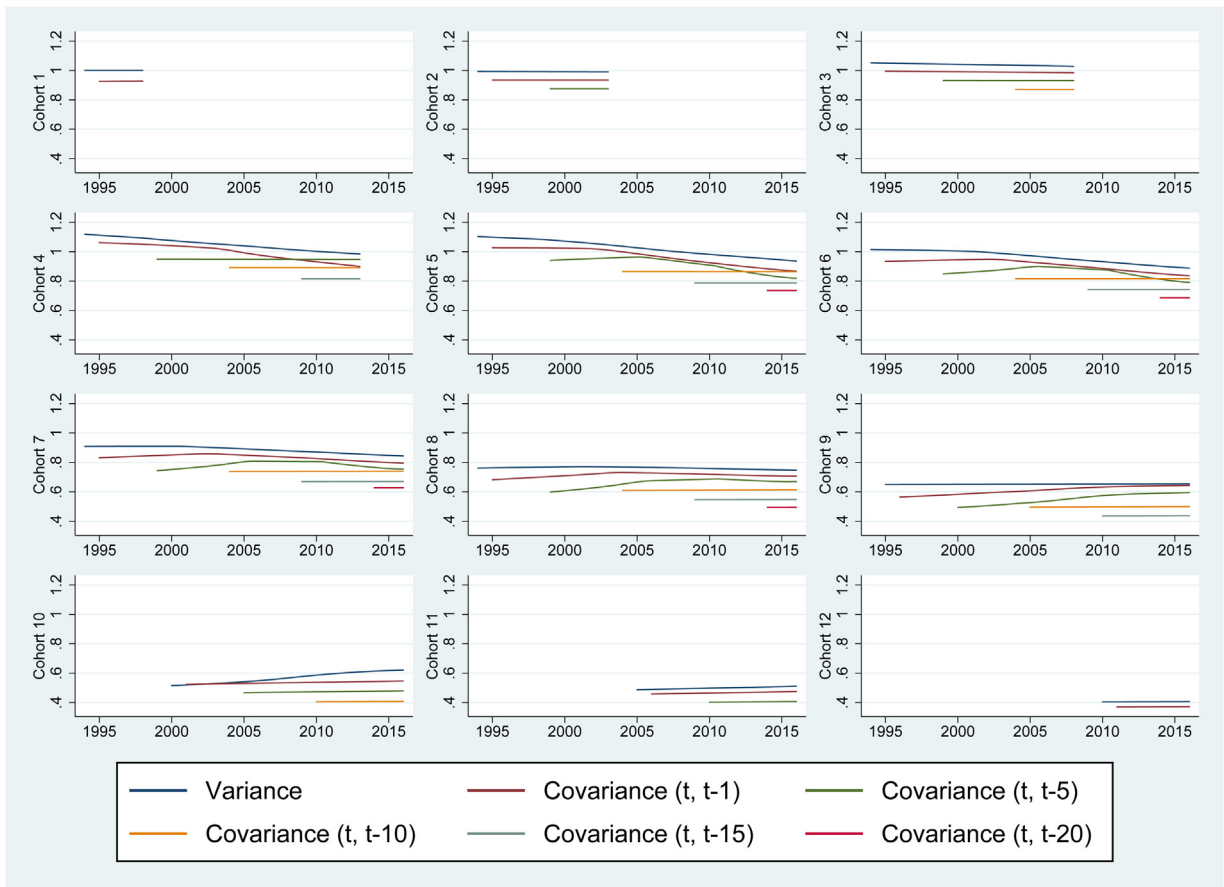


Fig. 7. Variance and selected covariances of the log of real hourly wages – all cohorts.

Source: Elaborated by the authors. Data: RAIS.

component following an AR(1) process; and (4) a model with factor loadings and the transitory component following an ARMA(1, 1) process. For models with factor loadings, cohort shifters for cohorts 1–3 and year shifters for the first year (1994) in both the permanent and transitory component were normalized to unity for identification, following the procedure used in this literature. The vast majority of parameters are statistically significant, with the exception of some cohort shifters for cohorts with few observed covariances (e.g. cohort 12).

We use models' Chi-squared statistics to compare them and select the preferred one: the model with the transitory component following an AR(1) process with factor loadings.<sup>19</sup> For both the log of real hourly wages and residual, this model is the option with the lowest chi-squared statistic.<sup>20</sup> Recall that this model includes: (i) year-specific shifters in the permanent component ( $\rho_t$ ); (ii) cohort shifters in the permanent component ( $\lambda_c$ ); (iii) year-specific shifters in the transitory component ( $\gamma_t$ ); (iv) cohort shifters in the transitory component ( $\kappa_c$ ); and (v) the assumption that the transitory component ( $v_{it}$ ) follows an AR(1) process.

The results for the log of real hourly wages and for the residuals are presented in Table 9. For the log of real hourly wages, we notice that year-specific shifters in the permanent component were relatively stable and close to unity during the first years of the period of analysis but then increased from the last years of the 1990s until 2008.

<sup>19</sup> Tables 13 and 14 in Appendix D present estimation results of these four models for the log of real hourly wages and residuals, respectively.

<sup>20</sup> Chi-squared statistics indicate rejection of the general model structures but the rejection of the null hypothesis at conventional significance levels is very common in this literature, as discussed in Ramos (2003). Therefore, we use this statistic for model comparison purposes only.

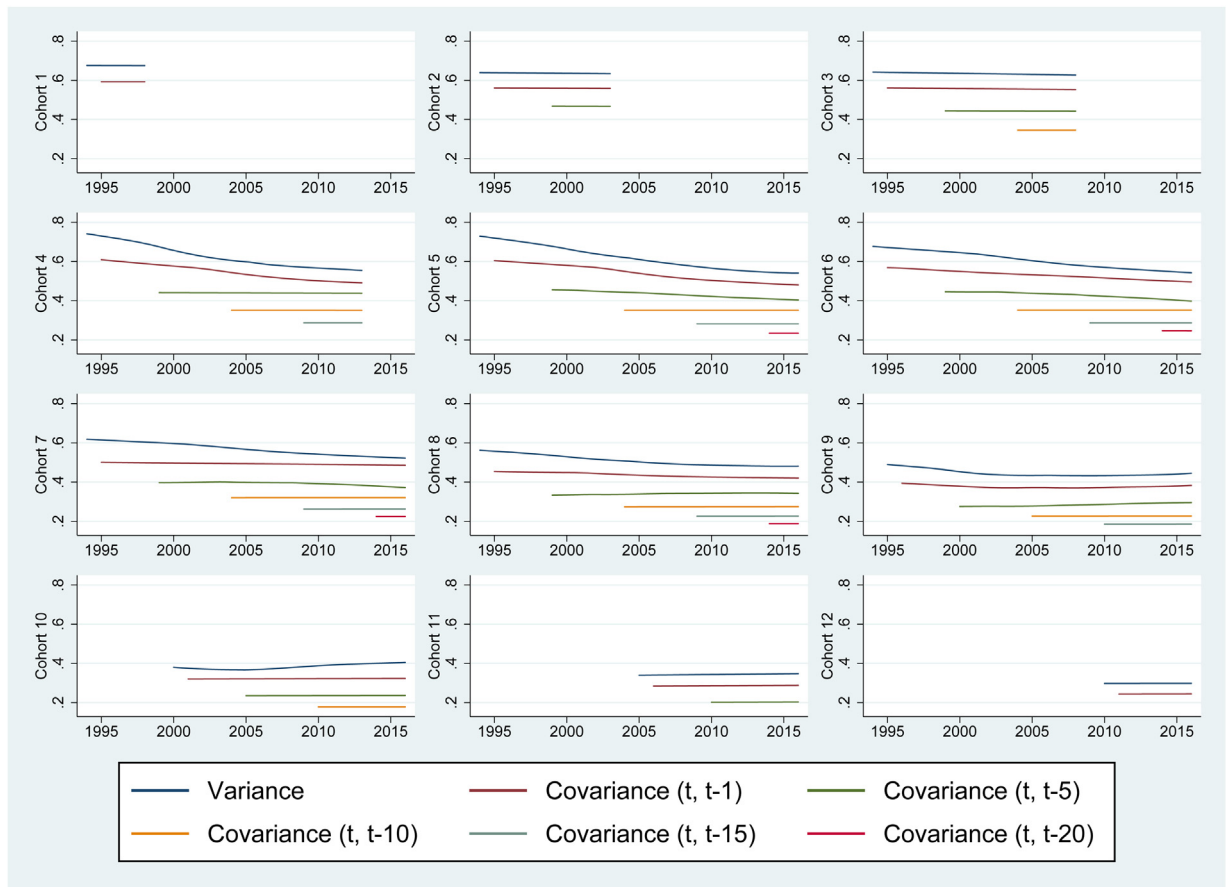


Fig. 8. Variance and selected covariances of the residuals – all cohorts.

Source: Elaborated by the authors. Data: RAIS.

Thereafter, they decreased substantially, especially from 2014 to 2016. This suggests that the permanent component had different roles over the analyzed period: until late 2000s it contributed to increasing wage inequality while it seems to be partially responsible for the wage inequality reduction in the last years. Still considering the permanent component, cohort shifters indicate that this component is much larger for older cohorts, as all estimates are below 1 and decrease substantially as we consider younger groups (i.e. those numbered with higher values). This pattern relates to the observed higher wage inequality among older cohorts presented in the previous section.

Considering the transitory component, as in most studies in this literature, we observe a highly persistent AR(1) process as the estimated parameter  $\phi$  is 0.906. This suggests that transitory shocks in wages need a long time to dissipate. For the year-specific shifters in this component, we notice they present a downward trend from the mid-1990s until 2003, then fluctuate until 2008. From 2009 onwards, there is an upward trend.<sup>21</sup> This suggests that the transitory component has contributed to decreasing wage inequality until the beginning of the 2000s, but in most recent years, however, it has been contributing to its increase. Finally, regarding cohort shifters for this component, we also notice that estimates are smaller for younger cohorts<sup>22</sup>, although variation among cohorts is not as high as in the case

<sup>21</sup> With the exception of 2011.

<sup>22</sup> Estimates for  $\kappa_{12}$  and  $\kappa_{13}$  are very different from the other cohort shifters, which we attribute to the fact that for this two cohorts we can only observe a reduced number of covariances, as described in Table 3.

Table 9  
Estimation results.

Parameter	Log wages		Residuals	
	Estimate	SE	Estimate	SE
<i>Permanent comp.</i>				
$\sigma_u^2$	0.754	0.0005	0.148	0.0006
$\rho_{1994}$	1.000	–	1.000	–
$\rho_{1995}$	0.996	0.0004	1.038	0.0010
$\rho_{1996}$	0.995	0.0003	0.974	0.0010
$\rho_{1997}$	1.009	0.0003	1.026	0.0011
$\rho_{1998}$	1.014	0.0003	0.965	0.0014
$\rho_{1999}$	1.033	0.0003	1.050	0.0011
$\rho_{2000}$	1.018	0.0003	0.988	0.0015
$\rho_{2001}$	1.029	0.0004	0.921	0.0021
$\rho_{2002}$	1.097	0.0003	1.298	0.0013
$\rho_{2003}$	1.101	0.0004	1.733	0.0019
$\rho_{2004}$	1.077	0.0004	1.374	0.0009
$\rho_{2005}$	1.098	0.0004	1.333	0.0010
$\rho_{2006}$	1.080	0.0004	1.257	0.0006
$\rho_{2007}$	1.098	0.0003	1.217	0.0009
$\rho_{2008}$	1.123	0.0003	1.235	0.0009
$\rho_{2009}$	1.040	0.0003	1.293	0.0010
$\rho_{2010}$	1.032	0.0003	1.222	0.0011
$\rho_{2011}$	0.757	0.0004	0.616	0.0017
$\rho_{2012}$	1.031	0.0004	1.289	0.0016
$\rho_{2013}$	0.980	0.0004	1.109	0.0014
$\rho_{2014}$	0.941	0.0003	1.124	0.0012
$\rho_{2015}$	0.931	0.0003	1.114	0.0013
$\rho_{2016}$	0.913	0.0003	1.089	0.0014
$\lambda_1$	1.000	–	1.000	–
$\lambda_2$	1.000	–	1.000	–
$\lambda_3$	1.000	–	1.000	–
$\lambda_4$	1.037	0.0005	0.988	0.0009
$\lambda_5$	1.027	0.0005	1.001	0.0010
$\lambda_6$	1.000	0.0005	1.036	0.0011
$\lambda_7$	0.949	0.0004	0.990	0.0010
$\lambda_8$	0.856	0.0004	0.907	0.0009
$\lambda_9$	0.744	0.0005	0.795	0.0011
$\lambda_{10}$	0.678	0.0005	0.683	0.0010
$\lambda_{11}$	0.620	0.0005	0.356	0.0032
$\lambda_{12}$	0.685	0.0007	0.011	0.0944
$\lambda_{13}$	0.424	0.0029	0.000	0.0000
<i>Transitory comp.</i>				
$\sigma_v^2$				
$\sigma_{v-1}^2$	0.122	0.0020	0.536	0.0018
$\sigma_\epsilon^2$	0.143	0.0004	0.074	0.0001
$\phi$	0.906	0.0007	0.915	0.0003
$\theta$				
$\gamma_{1994}$	1.000	–	1.000	–
$\gamma_{1995}$	0.830	0.0019	0.981	0.0011
$\gamma_{1996}$	0.801	0.0010	1.010	0.0010
$\gamma_{1997}$	0.639	0.0010	0.974	0.0006
$\gamma_{1998}$	0.675	0.0009	1.030	0.0006
$\gamma_{1999}$	0.584	0.0009	0.959	0.0004
$\gamma_{2000}$	0.585	0.0009	0.986	0.0005
$\gamma_{2001}$	0.610	0.0011	1.041	0.0006
$\gamma_{2002}$	0.176	0.0014	0.748	0.0011
$\gamma_{2003}$	0.012	0.0017	0.499	0.0018

Table 9 (Continued)

Parameter	Log wages		Residuals	
	Estimate	SE	Estimate	SE
$\gamma_{2004}$	0.381	0.0018	0.744	0.0007
$\gamma_{2005}$	0.265	0.0017	0.776	0.0008
$\gamma_{2006}$	0.352	0.0014	0.820	0.0008
$\gamma_{2007}$	0.269	0.0014	0.853	0.0007
$\gamma_{2008}$	0.130	0.0019	0.836	0.0009
$\gamma_{2009}$	0.515	0.0013	0.838	0.0009
$\gamma_{2010}$	0.534	0.0013	0.843	0.0010
$\gamma_{2011}$	0.277	0.0000	0.764	0.0010
$\gamma_{2012}$	0.531	0.0012	0.847	0.0009
$\gamma_{2013}$	0.564	0.0015	0.874	0.0010
$\gamma_{2014}$	0.626	0.0013	0.865	0.0010
$\gamma_{2015}$	0.633	0.0013	0.871	0.0011
$\gamma_{2016}$	0.659	0.0016	0.895	0.0012
$\kappa_1$	1.000	0.0011	1.000	–
$\kappa_2$	1.000	0.0012	1.000	–
$\kappa_3$	1.000	0.0013	1.000	–
$\kappa_4$	0.918	0.0017	1.046	0.0006
$\kappa_5$	0.867	0.0013	1.023	0.0006
$\kappa_6$	0.878	0.0011	0.987	0.0005
$\kappa_7$	0.927	0.0010	0.951	0.0004
$\kappa_8$	0.955	0.0010	0.904	0.0004
$\kappa_9$	0.900	0.0010	0.827	0.0006
$\kappa_{10}$	0.820	0.0008	0.799	0.0005
$\kappa_{11}$	0.966	0.0013	0.921	0.0008
$\kappa_{12}$	0.007	0.1512	0.910	0.0010
$\kappa_{13}$	–0.774	0.0029	0.822	0.0009
$\chi^2$ statistics		5,902,870		6,908,308
Degrees of freedom		1,951		1,951

Source: Elaborated by the authors. Data: RAIS.

of cohort shifters for the permanent component. Once again, this pattern relates to the observed higher wage inequality among older cohorts.

For the residuals, year-specific shifters in the permanent component were relatively stable and close to unity during the 1990s, then showed a substantial increase until 2003 followed by decrease onwards, although in a slower rate.<sup>23</sup> Cohort shifters for this component present a similar pattern to that observed for the log of real hourly wages: estimates are smaller for younger cohorts. In the case of the transitory component, we notice that year-specific shifters are relatively stable and close to unity until 1998, then present a downward trend in the beginning of the 2000s which is then followed by an upward trend until 2016. Cohort shifters for this component also present similar a trend to that observed for the log of real hourly wages.

## 5.2. Decomposition results

In order to access the relative contributions of each component to wage inequality, we decompose predicted variances into their permanent and transitory components using estimated parameters. Decomposition results are presented per cohort in Figs. 9 and 10 for the log of real hourly wages and residuals, respectively. For any given cohort and year, using bars, we present the percentages of total predicted variance that are attributed to the permanent and to the transitory

<sup>23</sup> With the exception of 2011.



Fig. 9. Log of real hourly wages' variance decomposition per cohort.

Source: Elaborated by the authors. Data: RAIS.

components, so that when we add both of them the result is always equal to 1. We also use these graphs to present the fit of our model considering variances (i.e. estimated predicted variance/actual observed variance per cohort and year). For most cohorts and years, our model seems to predict the evolution of variances relatively closely<sup>24</sup> and, in general, the model fit seems to be better for the log of real hourly wages than for residuals.

Considering the log of real hourly wages (Fig. 9), we notice that, except for cohort 13 (birth years 1990–1994), the permanent component predominates in terms of the share of total variance. Its relative importance, however, varied over the years. The permanent component was responsible for about 80% of total variance until 2001 for cohorts 1–7 (birth years 1930–1964) and for about 70% for cohorts 8–10 (birth years 1965–1979). From 2002 to 2008, we observe an increased relative importance of the permanent component for all cohorts (more than 90% in some cases). From 2009 onwards, however, there is a downward trend in its relative importance for all cohorts, except for cohort 12 (birth years 1985–1989).

For the residuals (Fig. 10), as expected, the relative importance of the permanent component is largely reduced when compared to that observed for the log of real hourly wages. This suggests that education and age are responsible for a substantial part of the permanent component regardless of the cohort or year considered, as discussed further in the next section. It is worth noting that the relative importance of the permanent component tends to be smaller for younger cohorts (e.g. 11, 12 and 13 (birth years 1980–1994)). In terms of the evolution of its relative importance over

<sup>24</sup> The main exceptions are cohorts with reduced number of observed covariances (e.g. cohorts 12 and 13 (birth years 1985–1994)) and the 2011 variances, which is related to the disparate smaller mean and variances of the log of real hourly wages observed in this year.





Fig. 10. Residuals' variance decomposition per cohort.

Source: Elaborated by the authors. Data: RAIS.

the years, we observe a similar pattern to that for the log of real hourly wages as it is relatively stable for all cohorts until 2001, then it gets increased during the 2000s with a peak in 2003. This is then followed by a downward trend in the last years although not as pronounced.

## 6. Discussion

### 6.1. Changes in inequality

Having observed distinct trends in wage inequality for different cohorts and decomposed it into its permanent and transitory components, we now proceed to assess the extent of the contribution of each component for changes in inequality from 1994 to 2016. We divide our period of analysis into 4 shorter intervals of 5–6 years each: 1994–2000, 2000–2005, 2005–2010, and 2010–2016. For each cohort and interval, we calculate the observed change per variance component as a percentage of the total estimated predicted variance of the first year of the interval. For example, for cohort 1 (birth years 1930–1934) and interval 1994–2000, the observed change in the permanent component is given by the total difference in this component between 2000 and 1994, divided by 1994's total estimated predicted variance.

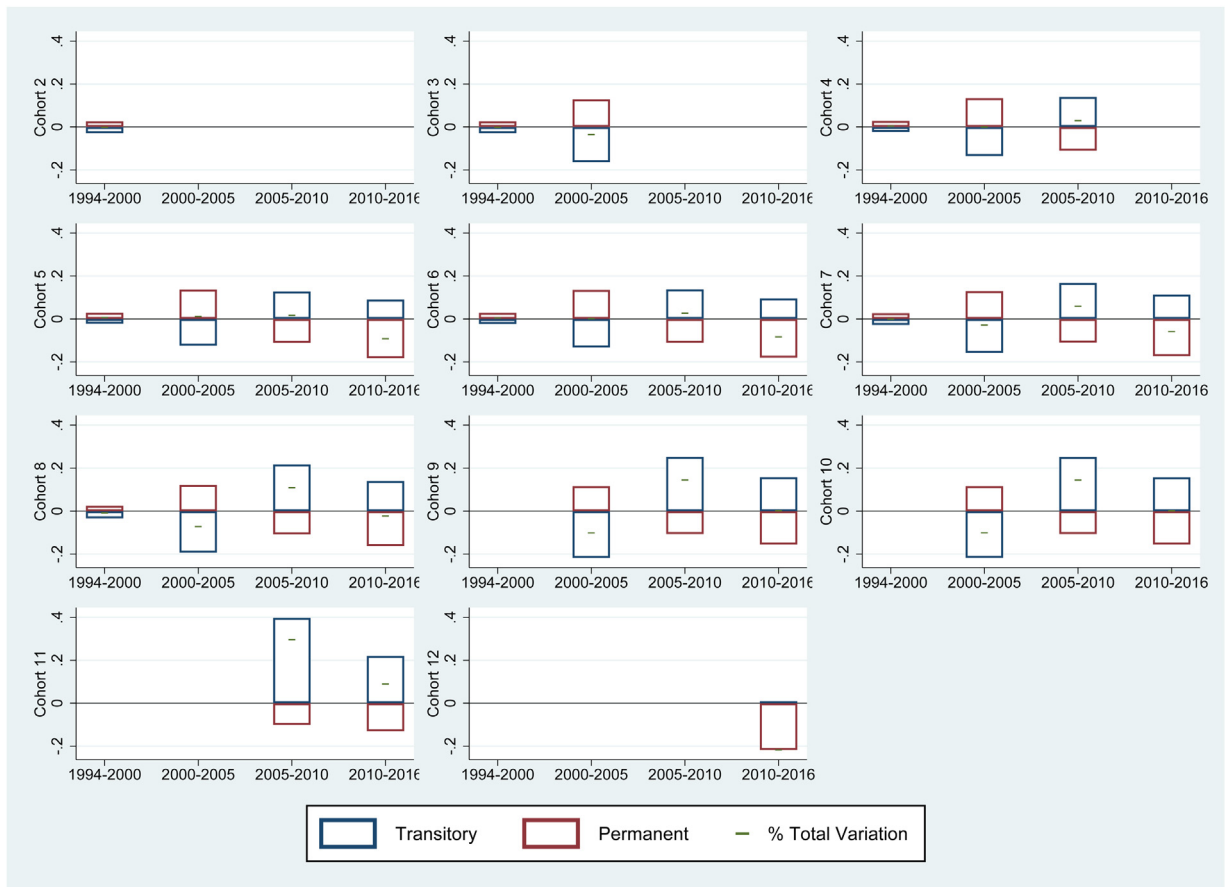


Fig. 11. Log of real hourly wages' changes in inequality and components' contribution per cohort.

Source: Elaborated by the authors. Data: RAIS.

Fig. 11 presents the total change in wage inequality (variance) per cohort<sup>25</sup> and time interval in green as well as observed changes per variance component (in red for the permanent component and in blue for the transitory component). For all cohorts, in the first and second considered intervals, we notice that the permanent component contributed to increasing wage inequality while the transitory component contributed to its reduction. In both cases, changes are much more expressive in the second (2000–2005) than in the first interval (1994–2000). Considering these two intervals, we can say that there was a wage inequality reduction for cohort 3 (birth years 1940–1944) and cohorts 7, 8, 9 and 10 (birth years 1960–1979) which was concentrated in 2000–2005 and completely driven by reductions in the transitory component, that were actually sufficient to offset increases in the permanent component in the same period. We attribute this changes to the stabilization process that took place in Brazil in the second half of the 1990s, which included the Real Plan that succeeded in reducing inflation, and other structural reforms (privatization, economic openness, restructuring of the financial sector, and restructuring of fiscal and monetary policy). The other cohorts experienced stability or a slight increase in wage inequality during this period, as a result of increases in the permanent component.

Results for the third and fourth intervals (2005–2010 and 2010–2016), however, are different. We now observe the permanent component contributing to wage inequality decrease while the transitory component contributed to its increase. From 2005 to 2010, we observe an increase in wage inequality for all cohorts caused completely by transitory

<sup>25</sup> Cohorts 1 (birth years 1930–1934) and 13 (birth years 1990–1994) are omitted as it is not possible to observe them during any of the considered intervals.

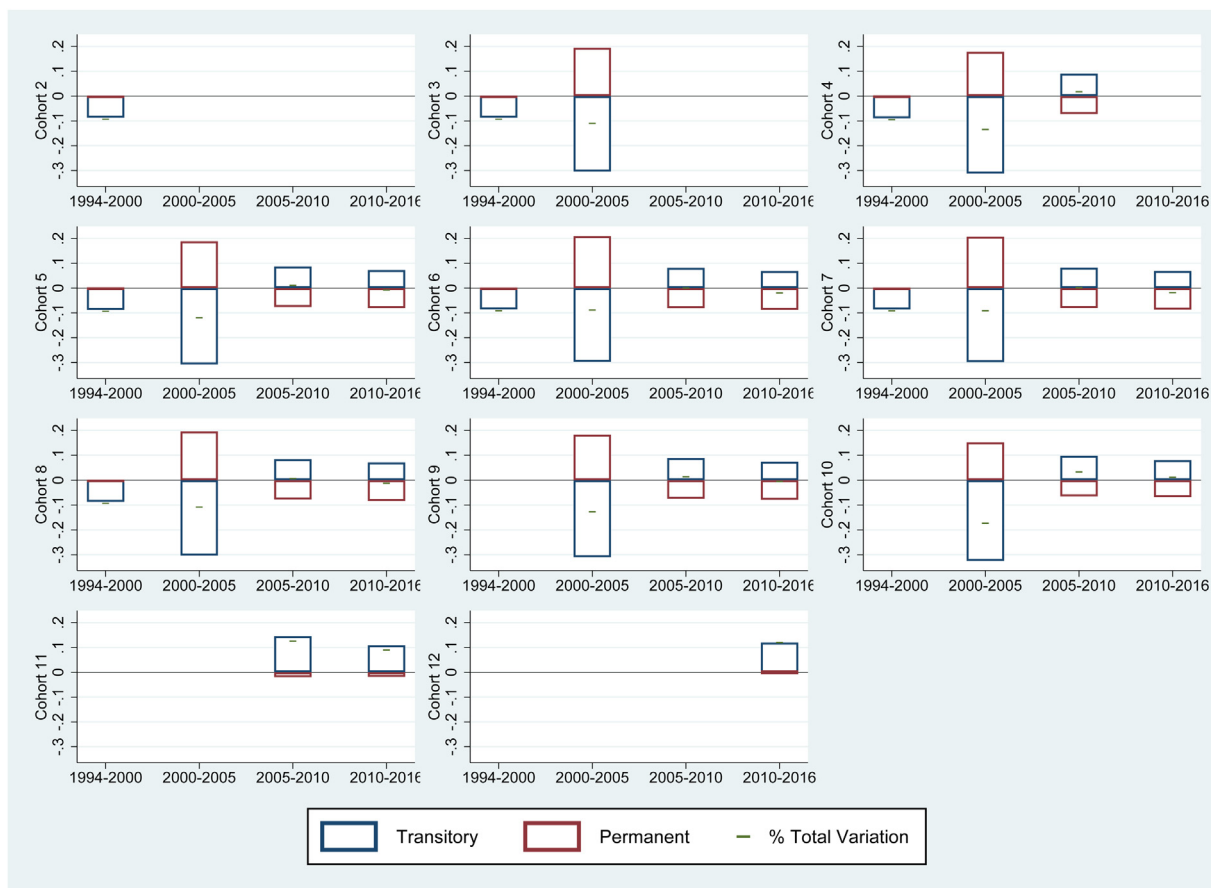


Fig. 12. Residuals' changes in inequality and components' contribution per cohort.

Source: Elaborated by the authors. Data: RAIS.

variance increases, with younger cohorts experiencing higher inequality increases (e.g. cohorts 11 and 12 (birth years 1980–1989)). In the last interval (2010–2016), distinct changes are observed for different cohorts. Cohorts 5–8 (birth years 1950–1969) and 12 (birth years 1985–1989) experienced a decrease in wage inequality, which can be entirely attributed to the reduction in the permanent component, that is associated with decreases in the returns on individuals' time-invariant characteristics (e.g. education and abilities). Cohorts 9 and 10 (birth years 1970–1979) present almost no variation in inequality as their permanent and transitory components' changes are balanced and of opposite signs. Cohort 11 (birth years 1980–1984) is the only one that experiences a wage inequality rise during this period due to transitory component increase.

Fig. 12 presents similar graphs for residuals. Overall, observed patterns are similar to those of the log of real hourly wages but with observed changes of lesser magnitude. For most cohorts, the permanent component contributed to wage inequality increase in the first two intervals and to its decrease in the last two. The transitory component contributed to wage inequality reduction until mid-2000s and to its rise from 2005 onwards.

As the residuals refer to wages net of age and education effects, it is important to highlight that changes in the permanent component of the variance of residuals are associated with variations in the returns on individuals' time-invariant attributes beyond those characteristics. The permanent component is also related to another important aspect of the labor market: mobility. Decreases in the permanent component are associated with more frequent swapping of positions between individuals in the distribution of wages over time, i.e. higher mobility, which seems to be the case for all cohorts in the last two intervals (2005–2010 and 2010–2016).



Fig. 13. Age and education's share in the permanent component per cohort.

Source: Elaborated by the authors. Data: RAIS.

## 6.2. The role of age and education in the level of permanent inequality

Our proposed model allows us to investigate the importance of observable permanent characteristics in the level of permanent inequality. In this subsection, we make use of our previous estimations of the permanent component of variances for the log of real hourly wages and residuals to calculate the share of this component of inequality level that can be attributed to age and education. This is particularly important as there is evidence that wage inequality in Latin America and, in particular, in Brazil, can be largely explained by these types of productivity characteristics and their returns (e.g. Reis and Barros, 1991; Lustig et al., 2013).

Let  $P_R$  be the permanent component of the residuals' inequality and  $P$  be the permanent component of inequality of the log of real hourly wages. The share of the permanent component attributed to age and education can be calculated as:

$$\text{Share} = 1 - \frac{P_R}{P} \quad (23)$$

Fig. 13 shows the evolution of this share over the years per cohort. For cohorts 1–9 (birth years 1930–1974), age and education were responsible for about 80% of total permanent variances from 1994 until early 2000s. The share decreased substantially in 2003 for all observed cohorts, but this is then followed by an upward trend until the early 2010s. Over the last years of our analyzed period, the share is once again relatively stable at approximately 70% (cohorts 5–10 (birth years 1950–1979)). For cohorts 12 and 13 (birth years 1985–1994), our results suggest that age and education are responsible for all permanent wage variances, but this result should be taken with caution as estimations

are less accurate for these groups as the number of available observed covariances is reduced. Comparing all cohorts, we can say that they present similar patterns in terms of the evolution of the share of the permanent component of wage inequality attributed to age and education, although results suggest this share tends to be larger for relatively younger cohorts (e.g. 10 and 11 (birth years 1975–1984)).

## 7. Conclusions

In this paper we used administrative data from the Brazilian formal labor market to investigate wage inequality among working age men from 1994 to 2016. We used the variances of the log of real hourly wages and of the residuals of its regression on age and education as measures of inequality. Following a similar approach of other papers in this literature, we decompose inequality into its permanent and transitory components.

We observe an overall decrease of inequality in this period for most cohorts. However, when considering each of the thirteen proposed birth-years cohorts and analyzing the evolution of wage inequality during the period, almost no change is observed for the oldest cohorts (1 and 2 (birth years 1930–1939)), a downward trend is evidenced for intermediate age cohorts (3–9 (birth years 1940–1974)) and an upward trend is found for some of the younger ones (10–12 (birth years 1975–1989)). Decomposition results of the log of real hourly wages suggest the permanent component has an important role in explaining wage inequality levels for all cohorts, with its relative importance being greater for the older ones, even though its share in terms of total variance has decreased over the 2010s.

In terms of the contribution of each component for changes in inequality, results differ depending on the time interval considered. From 1994 to 2005, there was a wage inequality reduction for cohort 3 (birth years 1940–1944) and cohorts 7, 8, 9 and 10 (birth years 1960–1979) that was completely driven by reductions in the transitory component which are related to the stabilization process of the late 1990s and its developments on the following years. The remaining cohorts experienced either stability or a slight increase in wage inequality, in this case as a result of increases in the permanent component. From 2005 to 2010, there was an increase in wage inequality for all cohorts caused by transitory variance increases, with younger cohorts (11 and 12 (birth years 1980–1989)) experiencing higher inequality growth. Finally, from 2010 to 2016, some of the intermediate age cohorts (5–8 (birth years 1950–1969) and 12 (birth years 1985–1989)) experienced a decrease in wage inequality, which can be entirely attributed to the reduction in the permanent component. This decrease in permanent inequality can be partly explained by decreases in year permanent component shifters ( $\rho$ ) and in the cohort shifters of younger cohorts ( $\lambda$ ). In the same sup-period, younger cohorts presented almost no change in inequality (cohorts 9 and 10 (birth years 1970–1979)) or experienced an inequality rise due to transitory component increase (cohort 11 (birth years 1980–1984)).

We can relate our findings to previous evidence in this literature by comparing our results those of similar studies for other countries, even though this should be made with caution because of methodological differences such as the kind of data source that was used, earnings/wages definition, sample design, proposed model, and estimation procedure. Starting with the United States, the country for which the majority of studies in this literature refer to, there is available evidence for the earnings inequality increase that was observed from the late 1960s until the 2000s. For the 1970s and 1980s, [Moffitt and Gottschalk \(2011\)](#) shows that both components equally contributed to the inequality increase. When extending the period of analysis to 2004, [Moffitt and Gottschalk \(2012\)](#) finds that the transitory component played a larger role in the earnings inequality increase in the 1970s and early 1980s which they attribute to economic instability. In the late 1980s, they show that both components contributed equally to the increase of inequality. Since the mid-1990s, however, their results suggest the permanent component accounted for two-thirds of the increase in earnings inequality. Using tax returns data from 1987 to 2009, [DeBacker et al. \(2013\)](#) also finds evidence that the earning inequality increase in the United States was mainly driven by increases in the permanent component.

A similar pattern was also found for Canada from 1976 to 1992 by [Baker et al. \(2003\)](#): earnings inequality grew substantially due to increases in both components, although the permanent component played a larger role. Studies for Great Britain find that both components had similar roles in the earnings inequality increase from the 1970s until the 1990s ([Dickens, 2000](#)). [Ramos \(2003\)](#) provides evidence that in the early 1990s, the permanent component predominated but lost its relative importance through the decade.

In a study considering 14 European Union countries in 1994–2001, [Sologon and O'Donoghue \(2009\)](#) provides evidence that for countries in which there was an earnings inequality increase, this was due to increases in the permanent component (Luxembourg, Italy, Greece and Finland) or in both components (Portugal and Netherlands). For countries that experienced a decrease in inequality, it was driven by the transitory component only (Germany, France, UK

and Ireland), the permanent component only (Belgium and Spain) or both (Denmark and Austria). The authors also highlight the fact that in all analyzed countries the permanent component presents greater relative importance in terms of total variance for older cohorts while the transitory component seems to play a larger role in explaining inequality in the case of younger cohorts, a pattern that is also observed by [Kalwij and Alessie \(2007\)](#) for Great Britain and in our case for Brazil.

While most studies in this literature investigate increases in wage inequality in developed countries, Brazil experienced a decrease in wage inequality in its formal labor market from 1994 to 2016 for most cohorts, with the exception of some of the younger ones. Another striking difference is the fact that the proportion of inequality change over time that can be attributed to each component is not as well-balanced in Brazil. In fact, our results suggest that, for most cohorts and years, changes in wage inequality were driven by one of the components exclusively. The transitory component was completely responsible for inequality decreases until mid-2000s while the permanent component was the one that caused wage inequality reductions for some cohorts in the 2010–2016 interval, which we attribute to overall increases in education attainment and declines in the schooling wage premia ([Barros et al., 2010](#); [Lustig et al., 2013](#); [Ferreira et al., 2016](#)) and also to a decline in the returns of non-observable permanent abilities.

Our findings contribute to the debate on the reasons for the observed decrease in wage inequality in Brazil, and suggest that the stabilization process that took place in the 1990s was particularly relevant for reducing inequality in the first half of our period of analysis. They also reveal the central role played by characteristics such as age and education in explaining wage inequality in Brazil: although their relative importance decreased over the years, they are still responsible for more than 50% of permanent inequality, with this percentage being even greater for younger cohorts.

Some final caveats are worth mentioning. As our results are restricted to working age men in the Brazilian formal labor market, generalizations should be made with caution, as we can expect that this group probably represents a lower bound of the transitory component of wage inequality, as discussed in [Santos and Souza \(2007\)](#). Although we include cohort effects in our models and use residuals, we do not address the possible consequences of formalization in inequality *per se* (i.e. how the entry of new people in the formal labor market changes its composition and how this may affect inequality), which we believe is an important topic for further research. Nonetheless, concerning public policy, our results suggest that there is still room for policies focused on formal or technical education so as to reduce permanent inequality as well as policies focused on promoting labor market stability as a way of reducing transitory wage inequality which has in fact risen in the last decade.

## Appendix A

The covariance matrix of the log of real hourly wages and the covariance matrix of the elements of this matrix can be estimated following [Abowd and Card \(1989\)](#), except for some modifications to adjust the method to the case of an unbalanced panel. Let  $y_i$  be the deviation of the log of real hourly wage of individual  $i$  at  $t$  from the mean log of wages for that year.  $T$  time periods are considered so that  $\tilde{y}_i$  is a  $1 \times T$  vector that contains information on the deviations in the log of the wage from the mean of each year for individual  $i$ :

$$\tilde{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix} \quad (24)$$

The covariance matrix,  $C$ , is a  $T \times T$  matrix whose elements,  $C_{jk}$ , where  $j \in [1, T]$  and  $k \in [1, T]$ , are given by:

$$C_{jk} = \frac{1}{n_{jk}} \sum_i \tilde{y}_{ij} \tilde{y}_{ik} \quad (25)$$

where  $n_{jk}$  stands for the number of individuals who contribute to that specific time period. To calculate the covariance matrix of this covariance matrix,  $V$ , let  $\tilde{m}_{ijk}$  be the vector of distinct elements of product  $\tilde{y}_{ij}\tilde{y}_{ik}$ . Using this notation, the elements of  $C$  can be estimated by:

$$m_{jk} = \frac{1}{n_{jk}} \sum_i \tilde{m}_{ijk} \tag{26}$$

Finally, define  $\tilde{s}_{ijk} = \tilde{m}_{ijk} - m_{jk}$ , so that the elements of the covariance matrix of covariances  $jk$  and  $mn$ ,  $V_{jkmn}$  where  $j \in [1, T]$ ,  $k \in [1, T]$ ,  $m \in [1, T]$  and  $n \in [1, T]$ , are defined by:

$$V_{jkmn} = \frac{1}{n_{jkmn}} \sum_i \tilde{s}_{ijk}\tilde{s}_{imn} \tag{27}$$

where  $n_{jkmn}$  stands for the number of individuals that contribute both to moments  $jk$  and  $mn$ . In other words,  $n_{jkmn}$  is the number of individuals for whom there is available information on wages at  $j, k, m$  and  $n$ . Using this approach, it is possible to calculate the observed covariance matrices,  $C$ , and the covariance matrix of covariance,  $V$ .

### Appendix B

Considering the case of the AR(1) model with factor loadings, the expression to be minimized is:

$$S'AS \tag{28}$$

where

$$S = \begin{pmatrix} \text{Var}(y_{c=1,t=1994}) - (\rho_{1994}^2 \lambda_1^2 \sigma_\mu^2 + \gamma_{1994}^2 \kappa_1^2 ((\phi^2 \sigma_{v-1}^2) + \sigma_\epsilon^2)) \\ \text{Cov}(y_{c=1,t=1995}, y_{c=1,t=1994}) - (\rho_{1994} \rho_{1995} \lambda_1^2 \sigma_\mu^2 + \gamma_{1994} \gamma_{1995} \kappa_1^2 ((\phi^3 \sigma_{v-1}^2) + (\phi \sigma_\epsilon^2))) \\ \text{Cov}(y_{c=1,t=1996}, y_{c=1,t=1994}) - (\rho_{1996} \rho_{1994} \lambda_1^2 \sigma_\mu^2 + \gamma_{1996} \gamma_{1994} \kappa_1^2 ((\phi^4 \sigma_{v-1}^2) + (\phi^2 \sigma_\epsilon^2))) \\ \vdots \\ \text{Var}(y_{c=13,t=2015}) - \left( \rho_{2015}^2 \lambda_{13}^2 \sigma_\mu^2 + \gamma_{2015}^2 \kappa_{13}^2 \left( (\phi^{44} \sigma_{v-1}^2) + \left( \sigma_\epsilon^2 \sum_{j=0}^{21} \phi^{2(21-j)} \right) \right) \right) \\ \text{Cov}(y_{c=13,t=2016}, y_{c=13,t=2015}) - \left( \rho_{2015} \rho_{2016} \lambda_{13}^2 \sigma_\mu^2 + \gamma_{2015} \gamma_{2016} \kappa_{13}^2 \left( (\phi^{45} \sigma_{v-1}^2) + \left( \sigma_\epsilon^2 \sum_{j=0}^{21} \phi^{2(21-j)+1} \right) \right) \right) \\ \text{Var}(y_{c=13,t=2016}) - \left( \rho_{2016}^2 \lambda_{13}^2 \sigma_\mu^2 + \gamma_{2016}^2 \kappa_{13}^2 \left( (\phi^{46} \sigma_{v-1}^2) + \left( \sigma_\epsilon^2 \sum_{j=0}^{22} \phi^{2(22-j)} \right) \right) \right) \end{pmatrix} \tag{29}$$

### Appendix C

Tables 10–12 present the covariance matrix of the residuals for cohorts 3, 7 and 11, respectively. Evolution patterns over the years are similar to those observed for the log of real hourly wages described in the Data section.

Table 10  
Covariance matrix of the residuals for cohort 3.

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
1994	0.789														
1995	0.628	0.726													
1996	0.574	0.623	0.702												
1997	0.533	0.577	0.611	0.686											
1998	0.498	0.541	0.566	0.604	0.659										
1999	0.476	0.517	0.540	0.569	0.587	0.639									
2000	0.454	0.494	0.516	0.545	0.560	0.585	0.632								
2001	0.440	0.476	0.495	0.523	0.537	0.558	0.576	0.637							
2002	0.381	0.400	0.412	0.435	0.449	0.465	0.481	0.514	0.588						
2003	0.354	0.383	0.388	0.407	0.419	0.431	0.446	0.470	0.497	0.573					
2004	0.350	0.367	0.372	0.392	0.404	0.415	0.426	0.450	0.474	0.509	0.578				
2005	0.338	0.352	0.361	0.377	0.387	0.402	0.409	0.430	0.452	0.482	0.513	0.582			
2006	0.336	0.346	0.339	0.362	0.371	0.392	0.396	0.417	0.436	0.464	0.491	0.524	0.581		
2007	0.325	0.329	0.320	0.344	0.362	0.378	0.397	0.403	0.435	0.449	0.476	0.505	0.522	0.586	
2008	0.311	0.327	0.314	0.321	0.345	0.358	0.375	0.375	0.415	0.441	0.461	0.477	0.496	0.511	0.568

Source: Elaborated by the authors. Data: RAIS.



Table 11  
Covariance matrix of the residuals for cohort 7.

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	
1994	0.671																							
1995	0.529	0.637																						
1996	0.483	0.539	0.631																					
1997	0.451	0.494	0.541	0.627																				
1998	0.418	0.458	0.494	0.536	0.609																			
1999	0.394	0.428	0.459	0.494	0.517	0.588																		
2000	0.373	0.405	0.431	0.465	0.485	0.514	0.588																	
2001	0.356	0.382	0.408	0.439	0.455	0.477	0.506	0.586																
2002	0.333	0.353	0.378	0.406	0.418	0.436	0.462	0.514	0.597															
2003	0.312	0.333	0.349	0.375	0.379	0.394	0.417	0.458	0.492	0.568														
2004	0.310	0.327	0.338	0.365	0.371	0.388	0.409	0.443	0.475	0.509	0.577													
2005	0.296	0.313	0.324	0.348	0.353	0.366	0.386	0.418	0.444	0.474	0.506	0.567												
2006	0.283	0.300	0.310	0.333	0.340	0.350	0.369	0.402	0.428	0.450	0.480	0.507	0.567											
2007	0.283	0.296	0.312	0.335	0.343	0.350	0.368	0.399	0.427	0.431	0.455	0.476	0.511	0.574										
2008	0.267	0.281	0.291	0.311	0.317	0.325	0.341	0.373	0.398	0.416	0.439	0.454	0.483	0.507	0.562									
2009	0.260	0.274	0.283	0.303	0.310	0.315	0.331	0.361	0.386	0.401	0.422	0.435	0.460	0.481	0.508	0.558								
2010	0.249	0.261	0.268	0.289	0.297	0.305	0.319	0.344	0.369	0.382	0.403	0.411	0.438	0.453	0.480	0.507	0.554							
2011	0.174	0.182	0.185	0.195	0.201	0.202	0.211	0.217	0.228	0.238	0.250	0.261	0.270	0.291	0.307	0.331	0.353	0.415						
2012	0.257	0.264	0.276	0.290	0.300	0.306	0.320	0.342	0.365	0.360	0.375	0.386	0.405	0.433	0.437	0.462	0.474	0.342	0.550					
2013	0.226	0.234	0.246	0.265	0.272	0.278	0.291	0.310	0.335	0.336	0.352	0.360	0.376	0.399	0.404	0.424	0.434	0.325	0.492	0.523				
2014	0.222	0.231	0.243	0.260	0.268	0.272	0.284	0.302	0.327	0.327	0.342	0.350	0.365	0.386	0.391	0.408	0.420	0.306	0.472	0.477	0.523			
2015	0.218	0.226	0.235	0.253	0.260	0.264	0.277	0.295	0.322	0.323	0.335	0.344	0.356	0.374	0.382	0.398	0.407	0.286	0.451	0.454	0.476	0.517		
2016	0.210	0.219	0.226	0.243	0.252	0.254	0.266	0.282	0.307	0.310	0.323	0.329	0.343	0.358	0.366	0.381	0.391	0.271	0.430	0.430	0.453	0.470	0.511	

Source: Elaborated by the authors. Data: RAIS.

Table 12  
Covariance matrix of the residuals for cohort 11.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
2005	0.308											
2006	0.255	0.311										
2007	0.234	0.262	0.319									
2008	0.213	0.237	0.266	0.328								
2009	0.198	0.218	0.241	0.271	0.328							
2010	0.186	0.204	0.223	0.249	0.274	0.339						
2011	0.161	0.182	0.201	0.223	0.244	0.280	0.348					
2012	0.161	0.177	0.191	0.208	0.224	0.251	0.281	0.351				
2013	0.158	0.173	0.189	0.204	0.216	0.240	0.265	0.292	0.365			
2014	0.157	0.172	0.185	0.198	0.207	0.230	0.250	0.272	0.315	0.375		
2015	0.152	0.168	0.180	0.191	0.201	0.221	0.238	0.258	0.293	0.325	0.378	
2016	0.148	0.166	0.174	0.184	0.193	0.212	0.226	0.244	0.272	0.299	0.326	0.372

Source: Elaborated by the authors. Data: RAIS.

## Appendix D

Tables 13 and 14 present estimation results of the four models for the log of real hourly wages and residuals, respectively.

Table 13  
Log of real hourly wages – estimation results.

Parameter	Canonical		Canonical with factor loadings		Factor loadings and AR(1)		Factor loadings and ARMA(1,1)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
<i>Permanent comp.</i>								
$\sigma_u^2$	0.857	0.0004	0.808	0.0002	0.754	0.0005	0.753	0.0004
$\rho_{1994}$			1.000	–	1.000	–	1.000	–
$\rho_{1995}$			1.008	0.0003	0.996	0.0004	0.996	0.0003
$\rho_{1996}$			1.023	0.0003	0.995	0.0003	0.998	0.0003
$\rho_{1997}$			1.037	0.0003	1.009	0.0003	1.012	0.0003
$\rho_{1998}$			1.068	0.0002	1.014	0.0003	1.018	0.0003
$\rho_{1999}$			1.043	0.0002	1.033	0.0003	1.039	0.0003
$\rho_{2000}$			1.034	0.0002	1.018	0.0003	1.025	0.0003
$\rho_{2001}$			1.054	0.0003	1.029	0.0004	1.038	0.0004
$\rho_{2002}$			1.067	0.0003	1.097	0.0003	1.098	0.0003
$\rho_{2003}$			1.022	0.0003	1.101	0.0004	1.102	0.0004
$\rho_{2004}$			1.088	0.0003	1.077	0.0004	1.084	0.0004
$\rho_{2005}$			1.091	0.0002	1.098	0.0004	1.102	0.0003
$\rho_{2006}$			1.085	0.0003	1.080	0.0004	1.088	0.0004
$\rho_{2007}$			1.076	0.0003	1.098	0.0003	1.105	0.0003
$\rho_{2008}$			1.065	0.0003	1.123	0.0003	1.128	0.0003
$\rho_{2009}$			1.127	0.0003	1.040	0.0003	1.060	0.0002
$\rho_{2010}$			1.128	0.0003	1.032	0.0003	1.051	0.0003
$\rho_{2011}$			0.797	0.0004	0.757	0.0004	0.768	0.0004
$\rho_{2012}$			1.110	0.0004	1.031	0.0004	1.049	0.0004
$\rho_{2013}$			1.053	0.0004	0.980	0.0004	0.998	0.0004
$\rho_{2014}$			1.018	0.0002	0.941	0.0003	0.962	0.0002
$\rho_{2015}$			0.999	0.0003	0.931	0.0003	0.952	0.0003
$\rho_{2016}$			0.976	0.0003	0.913	0.0003	0.934	0.0003
$\lambda_1$			1.000	–	1.000	–	1.000	–
$\lambda_2$			1.000	–	1.000	–	1.000	–
$\lambda_3$			1.000	–	1.000	–	1.000	–
$\lambda_4$			0.992	0.0004	1.037	0.0005	1.031	0.0005
$\lambda_5$			0.985	0.0004	1.027	0.0005	1.020	0.0005
$\lambda_6$			0.964	0.0004	1.000	0.0005	0.996	0.0004
$\lambda_7$			0.922	0.0004	0.949	0.0004	0.948	0.0004
$\lambda_8$			0.841	0.0004	0.856	0.0004	0.858	0.0004
$\lambda_9$			0.732	0.0004	0.744	0.0005	0.747	0.0005
$\lambda_{10}$			0.673	0.0004	0.678	0.0005	0.681	0.0005
$\lambda_{11}$			0.630	0.0004	0.620	0.0005	0.621	0.0005
$\lambda_{12}$			0.606	0.0004	0.685	0.0007	0.672	0.0006
$\lambda_{13}$			0.605	0.0004	0.424	0.0029	0.489	0.0013
<i>Transitory comp.</i>								
$\sigma_v^2$	0.104	0.0003	0.229	0.0002				
$\sigma_{v-1}^2$					0.122	0.0020		
$\sigma_\varepsilon^2$					0.143	0.0004	0.049	0.0001
$\phi$					0.906	0.0007	0.874	0.0007
$\theta$							–0.116	0.0022
$\gamma_{1994}$			1.000	–	1.000	–	1.000	–
$\gamma_{1995}$			0.862	0.0008	0.830	0.0019	1.118	0.0021
$\gamma_{1996}$			0.814	0.0008	0.801	0.0010	1.252	0.0020
$\gamma_{1997}$			0.720	0.0009	0.639	0.0010	1.088	0.0017

Table 13 (Continued)

Parameter	Canonical		Canonical with factor loadings		Factor loadings and AR(1)		Factor loadings and ARMA(1,1)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\gamma_{1998}$			0.651	0.0015	0.675	0.0009	1.194	0.0014
$\gamma_{1999}$			0.734	0.0005	0.584	0.0009	1.083	0.0014
$\gamma_{2000}$			0.714	0.0006	0.585	0.0009	1.115	0.0014
$\gamma_{2001}$			0.698	0.0008	0.610	0.0011	1.187	0.0015
$\gamma_{2002}$			0.391	0.0015	0.176	0.0014	0.342	0.0027
$\gamma_{2003}$			0.532	0.0016	0.012	0.0017	-0.015	0.0034
$\gamma_{2004}$			0.440	0.0011	0.381	0.0018	0.772	0.0032
$\gamma_{2005}$			0.343	0.0009	0.265	0.0017	0.535	0.0033
$\gamma_{2006}$			0.365	0.0010	0.352	0.0014	0.708	0.0028
$\gamma_{2007}$			0.394	0.0011	0.269	0.0014	0.518	0.0031
$\gamma_{2008}$			0.452	0.0011	0.130	0.0019	0.210	0.0043
$\gamma_{2009}$			0.079	0.0022	0.515	0.0013	0.978	0.0016
$\gamma_{2010}$			0.113	0.0017	0.534	0.0013	1.029	0.0011
$\gamma_{2011}$			0.653	0.0003	0.277	0.0000	0.523	0.0020
$\gamma_{2012}$			0.314	0.0009	0.531	0.0012	1.008	0.0013
$\gamma_{2013}$			0.415	0.0008	0.564	0.0015	1.072	0.0013
$\gamma_{2014}$			0.507	0.0005	0.626	0.0013	1.221	0.0011
$\gamma_{2015}$			0.548	0.0006	0.633	0.0013	1.220	0.0011
$\gamma_{2016}$			0.585	0.0006	0.659	0.0016	1.271	0.0014
$\kappa_1$			1.000	–	1.000	0.0011	1.000	–
$\kappa_2$			1.000	–	1.000	0.0012	1.000	–
$\kappa_3$			1.000	–	1.000	0.0013	1.000	–
$\kappa_4$			1.463	0.0010	0.918	0.0017	1.008	0.0017
$\kappa_5$			1.412	0.0007	0.867	0.0013	0.958	0.0014
$\kappa_6$			1.319	0.0007	0.878	0.0011	0.944	0.0011
$\kappa_7$			1.261	0.0007	0.927	0.0010	0.962	0.0010
$\kappa_8$			1.217	0.0007	0.955	0.0010	0.975	0.0009
$\kappa_9$			1.124	0.0011	0.900	0.0010	0.880	0.0011
$\kappa_{10}$			1.134	0.0013	0.820	0.0008	0.839	0.0008
$\kappa_{11}$			1.511	0.0021	0.966	0.0013	1.077	0.0015
$\kappa_{12}$			1.413	0.0009	0.007	0.1512	-0.013	0.0810
$\kappa_{13}$			0.570	0.0018	-0.774	0.0029	-0.771	0.0020
$\chi^2$ statistics	9,154,325		16,132,060		5,902,870		7,741,826	
Degrees of freedom	2,017		1,953		1,951		1,951	

Source: Elaborated by the authors. Data: RAIS.

Table 14  
Residuals – estimation results.

Parameter	Canonical		Canonical with factor loadings		Factor loadings with AR(1)		Factor loadings and ARMA(1,1)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
<i>Permanentcomp.</i>								
$\sigma_u^2$	0.393	0.0002	0.370	0.0001	0.148	0.0006	0.040	0.0004
$\rho_{1994}$			1.000	–	1.000	–	1.000	–
$\rho_{1995}$			1.036	0.0005	1.038	0.0010	1.042	0.0044
$\rho_{1996}$			1.059	0.0005	0.974	0.0010	0.746	0.0059
$\rho_{1997}$			1.113	0.0005	1.026	0.0011	0.872	0.0053
$\rho_{1998}$			1.232	0.0005	0.965	0.0014	0.742	0.0057
$\rho_{1999}$			1.070	0.0004	1.050	0.0011	0.872	0.0052
$\rho_{2000}$			1.091	0.0004	0.988	0.0015	0.748	0.0058
$\rho_{2001}$			1.126	0.0005	0.921	0.0021	0.511	0.0073
$\rho_{2002}$			1.120	0.0005	1.298	0.0013	1.650	0.0035

Table 14 (Continued)

Parameter	Canonical		Canonical with factor loadings		Factor loadings with AR(1)		Factor loadings and ARMA(1,1)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\rho_{2003}$			1.041	0.0006	1.733	0.0019	2.285	0.0046
$\rho_{2004}$			1.205	0.0006	1.374	0.0009	2.046	0.0037
$\rho_{2005}$			1.206	0.0005	1.333	0.0010	1.998	0.0042
$\rho_{2006}$			1.165	0.0005	1.257	0.0006	1.918	0.0047
$\rho_{2007}$			1.070	0.0005	1.217	0.0009	1.781	0.0049
$\rho_{2008}$			0.957	0.0004	1.235	0.0009	1.877	0.0050
$\rho_{2009}$			1.484	0.0006	1.293	0.0010	1.417	0.0018
$\rho_{2010}$			1.475	0.0007	1.222	0.0011	1.324	0.0021
$\rho_{2011}$			1.008	0.0008	0.616	0.0017	0.369	0.0042
$\rho_{2012}$			1.446	0.0007	1.289	0.0016	1.361	0.0024
$\rho_{2013}$			1.295	0.0007	1.109	0.0014	1.206	0.0029
$\rho_{2014}$			1.211	0.0004	1.124	0.0012	1.281	0.0023
$\rho_{2015}$			1.160	0.0005	1.114	0.0013	1.291	0.0026
$\rho_{2016}$			1.113	0.0005	1.089	0.0014	1.280	0.0029
$\lambda_1$			1.000	–	1.000	–	1.000	–
$\lambda_2$			1.000	–	1.000	–	1.000	–
$\lambda_3$			1.000	–	1.000	–	1.000	–
$\lambda_4$			0.805	0.0005	0.988	0.0009	1.006	0.0017
$\lambda_5$			0.828	0.0005	1.001	0.0010	1.046	0.0018
$\lambda_6$			0.839	0.0004	1.036	0.0011	1.113	0.0020
$\lambda_7$			0.812	0.0004	0.990	0.0010	1.074	0.0020
$\lambda_8$			0.762	0.0004	0.907	0.0009	0.983	0.0019
$\lambda_9$			0.680	0.0004	0.795	0.0011	0.814	0.0022
$\lambda_{10}$			0.624	0.0004	0.683	0.0010	0.666	0.0014
$\lambda_{11}$			0.573	0.0004	0.356	0.0032	–0.001	3.3450
$\lambda_{12}$			0.547	0.0004	0.011	0.0944	0.018	0.0996
$\lambda_{13}$			0.661	0.0005	0.000	0.0000	0.000	0.0000
<i>Transitory comp.</i>								
$\sigma_v^2$	0.204	0.0002	0.346	0.0001				
$\sigma_{v-1}^2$					0.536	0.0018		
$\sigma_\varepsilon^2$					0.074	0.0001	0.131	0.0002
$\phi$					0.915	0.0003	0.950	0.0001
$\theta$							–0.327	0.0007
$\gamma_{1994}$			1.000	–	1.000	–	1.000	–
$\gamma_{1995}$			0.897	0.0005	0.981	0.0011	0.990	0.0004
$\gamma_{1996}$			0.859	0.0005	1.010	0.0010	1.016	0.0004
$\gamma_{1997}$			0.743	0.0006	0.974	0.0006	0.986	0.0003
$\gamma_{1998}$			0.520	0.0011	1.030	0.0006	0.999	0.0004
$\gamma_{1999}$			0.770	0.0004	0.959	0.0004	0.971	0.0003
$\gamma_{2000}$			0.723	0.0004	0.986	0.0005	0.969	0.0005
$\gamma_{2001}$			0.694	0.0005	1.041	0.0006	1.003	0.0006
$\gamma_{2002}$			0.540	0.0007	0.748	0.0011	0.809	0.0007
$\gamma_{2003}$			0.702	0.0007	0.499	0.0018	0.739	0.0012
$\gamma_{2004}$			0.519	0.0005	0.744	0.0007	0.777	0.0008
$\gamma_{2005}$			0.488	0.0005	0.776	0.0008	0.791	0.0007
$\gamma_{2006}$			0.536	0.0006	0.820	0.0008	0.803	0.0005
$\gamma_{2007}$			0.686	0.0004	0.853	0.0007	0.826	0.0005
$\gamma_{2008}$			0.809	0.0004	0.836	0.0009	0.806	0.0006
$\gamma_{2009}$			0.287	0.0006	0.838	0.0009	0.894	0.0004
$\gamma_{2010}$			0.294	0.0005	0.843	0.0010	0.885	0.0004
$\gamma_{2011}$			0.555	0.0003	0.764	0.0010	0.709	0.0004
$\gamma_{2012}$			0.386	0.0004	0.847	0.0009	0.916	0.0001
$\gamma_{2013}$			0.467	0.0005	0.874	0.0010	0.875	0.0002
$\gamma_{2014}$			0.528	0.0003	0.865	0.0010	0.866	0.0004
$\gamma_{2015}$			0.564	0.0004	0.871	0.0011	0.864	0.0004

Table 14 (Continued)

Parameter	Canonical		Canonical with factor loadings		Factor loadings with AR(1)		Factor loadings and ARMA(1,1)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\gamma_{2016}$			0.599	0.0004	0.895	0.0012	0.867	0.0005
$\kappa_1$			1.000	–	1.000	–	1.000	–
$\kappa_2$			1.000	–	1.000	–	1.000	–
$\kappa_3$			1.000	–	1.000	–	1.000	–
$\kappa_4$			1.401	0.0006	1.046	0.0006	1.018	0.0006
$\kappa_5$			1.357	0.0005	1.023	0.0006	1.005	0.0006
$\kappa_6$			1.305	0.0005	0.987	0.0005	0.990	0.0005
$\kappa_7$			1.260	0.0004	0.951	0.0004	0.948	0.0005
$\kappa_8$			1.199	0.0004	0.904	0.0004	0.889	0.0005
$\kappa_9$			1.110	0.0007	0.827	0.0006	0.811	0.0005
$\kappa_{10}$			1.149	0.0008	0.799	0.0005	0.759	0.0005
$\kappa_{11}$			1.202	0.0010	0.921	0.0008	0.790	0.0008
$\kappa_{12}$			1.238	0.0008	0.910	0.0010	0.732	0.0008
$\kappa_{13}$			–0.572	0.0012	0.822	0.0009	0.709	0.0007
$\chi^2$ statistics	13,615,028		25,300,111		6,908,308		32,476,101	
Degrees of freedom	2,017		1,953		1,951		1,951	

Source: Elaborated by the authors. Data: RAIS.

## References

- Abowd, J.M., Card, D., 1989. On the covariance structure of earnings and hours changes. *Econometrica* 57 (2), 411–445.
- Altonji, J.G., Segal, L.M., 1996. Small-sample bias in GMM estimation of covariance structures. *J. Business Econ. Stat.* 14 (3), 353–366.
- Arabage, A.C., 2013. Os determinantes da mudança da desigualdade de salários no setor formal do Brasil.
- Arabage, A.C., Souza, A.P.F.D., 2015. Labor Earnings Dynamics in Post-Stabilization Brazil. Working Paper.
- Azevedo, J.P., Inchauste, G., Sanfelice, V., 2013. Decomposing the Recent Inequality Decline in Latin America. World Bank Policy Research Working Paper (6715).
- Baker, M., 1997. Growth-rate heterogeneity and the covariance structure of life-cycle earnings. *J. Labor Econ.*, 338–375.
- Baker, M., Solon, G., et al., 2003. Earnings dynamics and inequality among Canadian men, 1976–1992: evidence from longitudinal income tax records. *J. Labor Econ.* 21 (2), 267–288.
- Barros, R., de Carvalho, M., Franco, S., Mendonça, R., 2010. Markets, the state and the dynamics of inequality: the case of Brazil. In: Lopez-Calva, L.F., Lustig, N. (Eds.), *Declining Inequality in Latin America: A Decade of Progress?*
- Bartels, C., Bönke, T., 2010. German Male Income Volatility 1984 to 2008: Trends in Permanent and Transitory Income Components and the Role of the Welfare State. Working Paper.
- Beach, C.M., Finnie, R., Gray, D., 2010. Long-run inequality and short-run instability of men's and women's earnings in Canada. *Rev. Income Wealth* 56 (3), 572–596.
- Blundell, R., Graber, M., Mogstad, M., 2014. Labor income dynamics and the insurance from taxes, transfers, and the family. *J. Publ. Econ.*
- Bönke, T., Giesecke, M.N., Lüthen, H., 2015. The Dynamics of Earnings in Germany: Evidence from Social Security Records. Working Paper.
- Cappellari, L., 2004. The dynamics and inequality of Italian men's earnings – long-term changes or transitory fluctuations? *J. Human Resour.* 39 (2), 475–499.
- Cervini-Plá, M., Ramos, X., 2012. Long-term earnings inequality, earnings instability and temporary employment in Spain: 1993–2000. *Br. J. Ind. Relat.* 50 (4), 714–736.
- Chamberlain, G., 1984. Panel data. *Handb. Econometr.* 2, 1247–1318.
- Corseuil, C.H., Foguel, M.N., 2002. Uma sugestão de deflatores para rendas obtidas a partir de algumas pesquisas domiciliares do IBGE.
- de Souza, P.H., 2012. Poverty, Inequality and Social Policies in Brazil, 1995–2009. Technical Report.
- DeBacker, J., Heim, B., Panousi, V., Ramnath, S., Vidangos, I., 2013. Rising inequality: transitory or persistent? New evidence from a panel of US tax returns. *Brook. Papers Econ. Activ.* 2013 (1), 67–142.
- Dickens, R., 2000. The evolution of individual male earnings in Great Britain: 1975–95. *Econ. J.* 110 (460), 27–49.
- Edwards, S., 1996. Crisis and reform in Latin America: from despair to hope. *Am. Econ.* 40 (2), 98.
- Ferreira, F., Firpo, S.P., Messina, J., 2016. Understanding recent dynamics of earnings inequality in Brazil. In: *New Order and Progress: Development and Democracy in Brazil.*, pp. 187.
- Ferreira, F.H., Leite, P.G., Wai-Poi, M., 2007. Trade Liberalization, Employment Flows, and Wage Inequality in Brazil. Research Paper, UNUWIDER, United Nations University (UNU) (2007/58).
- Freije, S., Souza, A., et al., 2002. Earnings Dynamics and Inequality in Venezuela: 1995–1997. Working Paper (02-W11).

- Gasparini, L., Galiani, S., Cruces, G., Acosta, P., 2011. Educational Upgrading and Returns to Skills in Latin America: Evidence from a Supply–Demand Framework, 1990–2010. *World Bank Policy Research Working Paper* (5921).
- Gonzaga, G., Menezes Filho, N., Terra, C., 2006. Trade liberalization and the evolution of skill earnings differentials in Brazil. *J. Int. Econ.* 68 (2), 345–367.
- Gottschalk, P., Moffitt, R., Katz, L.F., Dickens, W.T., 1994. The growth of earnings instability in the US labor market. *Brook. Papers Econ. Activ.*, 217–272.
- Gustavsson, M., 2008. A new picture of Swedish earnings inequality: persistent and transitory components, 1960–1990. *Rev. Income Wealth* 54 (3), 324–349.
- Haider, S.J., 2001. Earnings instability and earnings inequality of males in the United States: 1967–1991. *J. Labor Econ.* 19 (4), 799–836.
- Kalwij, A.S., Alessie, R., 2007. Permanent and transitory wages of British men, 1975–2001: year, age and cohort effects. *J. Appl. Econometr.* 22 (6), 1063–1093.
- Kässi, O., 2014. Earnings dynamics of men and women in Finland: permanent inequality versus earnings instability. *Emp. Econ.* 46 (2), 451–477.
- Li, Y., 2012. The Covariance Structure of male Wages in the United Kingdom. Working Paper.
- Lillard, L.A., Willis, R.J., 1978. Dynamic aspects of earning mobility. *Econometrica*, 985–1012.
- Lochner, L., Park, Y., Shin, Y., 2018. Wage Dynamics and Returns to Unobserved Skill. Technical Report. National Bureau of Economic Research.
- López-Calva, L.F., Lustig, N., 2010. Declining Inequality in Latin America: A Decade of Progress? *Brookings Institution Press*.
- Lora, E.A., 2012. Structural Reforms in Latin America: What Has Been Reformed and How To Measure It. IDB Working Paper No. IDB-WP-346.
- Lustig, N., Lopez-Calva, L.F., Ortiz-Juarez, E., 2013. Declining inequality in Latin America in the 2000s: the cases of Argentina, Brazil, and Mexico. *World Dev.* 44, 129–141.
- MaCurdy, T.E., 1982. The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *J. Econometr.* 18 (1), 83–114.
- Manacorda, M., Sánchez-Páramo, C., Schady, N., 2010. Changes in returns to education in latin America: the role of demand and supply of skills. *Ind. Labor Relat. Rev.* 63 (2), 307–326.
- Moffitt, R.A., Gottschalk, P., 2011. Trends in the covariance structure of earnings in the US: 1969–1987. *J. Econ. Inequ.* 9 (3), 439–459.
- Moffitt, R.A., Gottschalk, P., 2012. Trends in the transitory variance of male earnings: methods and evidence. *J. Human Resour.* 47 (1), 204–236.
- Myck, M., Ochmann, R.R., Qari, S., 2009. Dynamics of Earnings and Hourly Wages in Germany. Working Paper.
- Newey, W.K., 1985. Generalized method of moments specification testing. *J. Econometr.* 29 (3), 229–256.
- Ramos, X., 2003. The covariance structure of earnings in Great Britain\* 1991–1999. *Economica* 70 (278), 353–374.
- Rao, C.R., 1962. A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. *J. R. Stat. Soc. Ser. B Methodol.*, 152–158.
- Reis, J.G.A.D., Barros, R.P.D., 1991. Wage inequality and the distribution of education: a study of the evolution of regional differences in inequality in metropolitan Brazil. *J. Dev. Econ.* 36 (1), 117–143.
- Santos, A.L., Souza, A.P., 2007. Earnings inequality in Brazil: is it permanent or transitory? *Braz. Rev. Econometr.* 27 (2), 259–284.
- Sologon, D.M., O’Donoghue, C., 2009. Earnings Dynamics and Inequality in EU, 1994–2001. Working Paper.
- Sologon, D.M., O’Donoghue, C., 2010. Earnings Dynamics and Inequality Among Men in Luxembourg, 1988–2004: Evidence from Administrative Data. Working Paper.
- Sologon, D.M., O’Donoghue, C., 2012. Cross-national trends in permanent earnings inequality and earnings instability in Europe 1994–2001. Working Paper.
- Souza, A.P., Lima, L., Berlinger, M.M., 2018. Understanding Brazilian Income Inequality: Trends, Determinants and policies. Working Paper.
- Zhao, Z., 2007. Earnings Instability and Earnings Inequality in Urban China: 1989–2006. IZA Discussion Paper.