

# Firms and the Decline of Earnings Inequality in Brazil\*

Jorge Alvarez<sup>†</sup>   Felipe Benguria<sup>‡</sup>   Niklas Engbom<sup>§</sup>   Christian Moser<sup>¶</sup>

April 15, 2016

## Abstract

We document a large decline in earnings inequality in Brazil between 1996 and 2012, with the variance of log earnings falling by 26 log points. Using administrative linked employer-employee data, we fit high-dimensional worker and firm fixed effects models within overlapping subperiods to identify the sources of this decline. Compression in firm effects accounts for 45 percent of the total decline and compression in worker effects accounts for 24 percent, with a fall in their covariance and the residual explaining the remainder. Half of the decrease in firm pay differences and a fifth of the decline in heterogeneity between workers are explained by observable characteristics. While firm and worker characteristics became more dispersed over the period, a more than commensurate decline in returns to these variables lead to an overall decrease in earnings inequality. We conclude that changes in pay policies, rather than changes in firm and worker fundamentals, played a significant role in Brazil's inequality decline.

**Keywords:** Earnings Inequality, Linked Employer-Employee Data, Firms, Productivity

**JEL classification:** D22, E24, J08, J31

---

\*We are grateful for the generous advice of Richard Rogerson since the inception of this project. We thank Elhanan Helpman and Marc Muendler for granting us access to part of the data. We appreciate helpful comments from Mark Aguiar, Adrien Auclert, Angus Deaton, Mike Golosov, Oleg Itskhoki, Gregor Jarosch, Greg Kaplan, Nobu Kiyotaki, Alan Krueger, Ilyana Kuziemko, Alex Mas, Ben Moll, Ezra Oberfield, Stephen O'Connell, Stephen Redding, Hannes Schwandt, Tom Vogl, Chris Woodruff, as well as participants at the PEDL conference in Warwick, the Princeton Macroeconomics lunch seminar, the Princeton Center for Health and Wellbeing lunch seminar, the Princeton Industrial Relations labor lunch seminar, the Princeton Public Finance Working Group, the Chicago Fed Rookie Conference, and the Peterson Institute for International Economics. Special thanks go to Carlos Lessa and Luis Pinto at IBGE as well as Carlos Corseuil, Glauca Ferreira, and Leandro Justino at IPEA for facilitating our data work. The authors gratefully acknowledge financial support from CEPR PEDL. Moser also benefitted from financial support from the Ewing Marion Kauffman Foundation and the Princeton Institute for International and Regional Studies. All errors are our own.

<sup>†</sup>Department of Economics, Princeton University.

<sup>‡</sup>Department of Economics, Gatton College of Business and Economics, University of Kentucky.

<sup>§</sup>Department of Economics, Princeton University.

<sup>¶</sup>Department of Economics, Princeton University.

# 1 Introduction

Brazil has experienced a large reduction in earnings inequality since the mid-1990s. This came after decades of Brazil being infamously known as the most unequal country in Latin America, which itself ranked among the most unequal regions in the world. While the decline of earnings inequality in Brazil resembles other Latin American economies' experience during this period,<sup>1</sup> it stands in stark contrast to that of the U.S. and many developed countries, which saw inequality steadily increasing over the past two decades.<sup>2</sup> To investigate the sources of this decline, in this paper we decompose Brazil's earnings inequality evolution into changes in firm and worker characteristics on the one hand, and the returns to such characteristics on the other.

Guided by recent research, which suggests that firms are an important determinant of earnings dispersion in a number of high-income countries,<sup>3</sup> we decompose the sources of Brazil's inequality decline by exploiting a large administrative linked employer-employee dataset containing information on over one billion job spells between 1988 and 2012. By linking individual workers to their employers and tracking both over time, we are able to separately identify the contributions of firm- and worker-specific factors towards worker's pay (Abowd, Kramarz and Margolis, 1999). Applying this econometric framework repeatedly within overlapping subperiods, we decompose the overall inequality decline into changes between firms and between workers. We investigate the link between firm performance and the firm component of pay by linking our first stage estimation results to another confidential dataset containing detailed information on the characteristics of hundreds of thousands of Brazilian firms between 1996 and 2012.

We uncover three main results. First, firms played an important role in the decline in earnings inequality in Brazil over this period, explaining 45 percent of the fall in the variance of log earnings between 1996 and 2012. Compression in worker fixed effects explains an additional 24 percent of the decline, with the remaining part being attributed a decline in the covariance between worker and firm fixed effects and the residual. As worker heterogeneity is most important in levels, the compression in firm-specific pay components contributed more than proportionately towards Brazil's inequality decline.

---

<sup>1</sup>See Lopez and Perry (2008) and Tsounta and Osueke (2014).

<sup>2</sup>Using administrative data, Kopczuk, Saez and Jae (2010) document in detail earnings inequality trends for the U.S., while chapter 8 of Atkinson and Bourguignon (2015) discusses inequality trends in middle- and high-income countries.

<sup>3</sup>See Abowd, Kramarz and Margolis (1999); Card, Heining and Kline (2013); Barth, Davis and Bryson (2014); Bloom et al. (2015), and Mueller and Ouimet (2015).

Second, changes in the link between firm performance and pay accounts for a significant fraction of the compression in the firm component of workers' earnings. We first show that a substantial share of the variation in the firm component of pay is explained by differences in observable firm characteristics, with more productive firms paying more. Moreover, more than half of the decline in the firm component is accounted for by observable firm characteristics. All of this decline is driven by a weakening pass-through from productivity to pay, and none is due to firms becoming more similar in observable characteristics. A weaker link between observable firm characteristics and worker pay thus explains 25 percent of the overall fall in the variance of log earnings over this period.

Third, a decline in the return to measures of ability such as experience and education explains a sizable share of the fall in the variance of the worker component of pay. In levels, age, education and occupation explain around 40 percent of the variance of the worker component of pay. However, we observe no compression in the underlying distributions of such characteristics over time. Instead, the inequality decline is driven by a rapid fall in the returns to observable measures of worker ability, especially the returns to education. We find that lower returns to education explain 13 percent and lower returns to age explain three percent of the overall fall in the variance of log earnings over this period.

This decomposition of the sources of Brazil's earnings inequality decline informs our understanding of various commonly proposed stories explaining the decline. On the worker side, our results do not support a widely-held belief that changes in educational attainment accounted for a significant share of Brazil's inequality evolution over the period.<sup>4</sup> While educational attainment increased over this period, we find that all else equal this resulted in a small net *increase* in inequality. We reach similar conclusions regarding the implications of changes in the age structure of the workforce. On the firm side, a reading of the existing literature would suggest that trade dynamics and the productivity evolution during this period could have been important drivers behind changes to the earnings distribution.<sup>5</sup> Yet, in line with U.S. trends, we find that also the Brazilian productivity distribution grew more dispersed over this period.

Our findings thus pose a challenge to candidate explanations for the decline in earnings inequality over this period. Our results suggest that a theory of the inequality decline needs to be

---

<sup>4</sup>See for example [Barros et al. \(2010\)](#).

<sup>5</sup>See for example [Dunne et al. \(2004\)](#) and [Faggio, Salvanes and Van Reenen \(2010\)](#), who conclude that parts of the increase in earnings inequality in the U.S. can be explained by widening dispersion in the firm productivity distribution.

consistent with the following three facts: (i) firm-level changes explain a significant share of initial inequality levels and an even larger share of its decline; (ii) a weaker pass-through from firm productivity to worker pay was a key driver behind the declining dispersion in the firm component of pay; and (iii) a lower return to worker ability explains a significant share of the decline in the worker component of pay. We conclude that changes in pay policies rather than in worker and firm fundamentals played an important role in the decline in inequality in Brazil during this period.

**Related literature.** With the current project we contribute to three broad strands of the literature. First, we provide a decomposition of earnings into worker and firm heterogeneity in a developing economy. With the increasing availability of large, administrative matched employer-employee datasets, a recent literature has started to examine the role of firms in wage determination. The first paper to make use of such large, linked employer-employee datasets to jointly study the role of worker unobservables and firms for pay is [Abowd, Kramarz, and Margolis \(1999, henceforth AKM\)](#), who study the role of firm and worker heterogeneity for wage inequality in France. They find an important role for firms in generating earnings inequality. Similar conclusions using the same methodology have been reached among others for the state of Washington in the U.S. ([Abowd, Creedy and Kramarz, 2002](#)), Denmark ([Bagger, Sorensen and Vejlin, 2013](#)), Austria ([Gruetter and Lalive, 2009](#)), and Germany ([Card, Heining and Kline, 2013](#)). The last paper is close to our methodology of applying the AKM framework in overlapping subperiods to study changes in wage determinants over time. They find that increasing dispersion in the firm-specific component of pay contributed significantly to rising earnings inequality in West Germany. [Bloom et al. \(2015\)](#) highlight firms as an important driver behind the increase in U.S. labor earnings inequality since 1980, but do not employ the AKM methodology to control for sorting of highly-paid workers into high-paying firms.

Second, we highlight the sources of changes in the worker and firm component of inequality by studying the role of changes in observable characteristics versus changes in the return to such characteristics. AKM study the link between firm observable characteristics and the estimated firm effects, but do not focus on changes over time to this relationship. [Menezes-Filho, Muendler and Ramey \(2008\)](#) study the link between firm characteristics and wages in Brazil in the cross-section of linked data on worker earnings and firm characteristics in Brazil's manufacturing and

mining sectors. [Bagger, Jesper and Holloway \(2014\)](#) investigate the role of labor misallocation in driving the positive correlation between labor productivity and wages at the firm using Danish data. [Card, Cardoso and Kline \(2015\)](#) study the degree of rent-sharing in Portugal with a particular emphasis on gender differences in profit participation and the allocation of workers across firms.

Third, we add to the literature on the sources of the decline in inequality in Brazil by providing a comprehensive decomposition of the decline. Many previous papers have studied the role of specific, isolated mechanisms in Brazil's inequality decline over the last two decades. For example, [Ulyssea \(2014\)](#) considers the role of worker flows between the informal and formal sectors. In related work, [de Araujo \(2014\)](#) studies the role of labor adjustment costs in propagating wage inequality in a frictional search framework. [Dix-Carneiro and Kovak \(2015\)](#) analyze the long-lasting impact of industry-specific tariff cuts in the presence of wage-equalizing migration. [Barros et al. \(2010\)](#) use Brazilian household data to study inequality trends since 1977 and decompose the decline in labor earnings inequality in Brazil since 1990. Given their data and method, those authors conclude that the inequality decline was in equal shares driven by education reform and labor market integration. [Medeiros, Souza and de Castro \(2014\)](#) use administrative tax return data to study the evolution of top income inequality in Brazil from 2006–2012, but they cannot distinguish between the role played by worker versus firm characteristics during that period. Using linked employer-employee data, [Lopes De Melo \(2013\)](#) conducts a static decomposition of earnings inequality levels in Brazil's formal sector into components due to firms and workers. [Helpman et al. \(2013\)](#) use the same worker-level data to show that a significant share of overall wage inequality is due to between-firm differences and that Brazil's trade liberalization starting in the late 1980s led to increasing between-firm earnings inequality.

**Outline.** The rest of the paper is structured as follows: Section 2 provides an overview of the main institutional changes and macroeconomic trends affecting Brazilian labor markets from 1988 to 2012. Section 3 summarizes the administrative datasets used in our empirical analysis and discusses sample selection and variable definitions. Section 4 provides descriptive statistics on trends in earnings inequality in Brazil during this time. Section 5 introduces the empirical framework we use to decompose the variance of log earnings into a worker and firm effect as well as the subsequent regressions we run to link these estimates to worker and firm fundamentals. Section 6 presents our main empirical results as well as checks on the validity of our empirical framework.

Finally, Section 7 summarizes our key findings and concludes.

## 2 Institutions and macroeconomic trends in Brazil

During our period of study, Brazil resumed democratic elections (1989), ended a decade of hyperinflation (1994), and inaugurated two decades of sustained economic growth—between 1996 and 2012 real gross domestic product grew by 2.3 percent per year on average. In this section, we discuss some of the institutional changes that could have affected inequality during this period, including labor regulation, trade liberalization and social policy.

Brazil had a highly regulated labor market before reforms started in the late 1980s. For instance, since 1965, a national Wage Adjustment Law mandated yearly wage increases for all workers in the economy and dismissal costs were high. After the transition to civil rule and the signing of a new constitution in 1988, flexibility in labor markets was further affected by firing penalties and an increased power of labor unions, which gather about a quarter of employed formal workers in Brazil.<sup>6</sup>The Wage Adjustment Law was finally abandoned in 1995, introducing a period of greater flexibility and less regulated wage-setting practices. Further legislation in 1997–1998 eased restrictions on temporary contracts and lowered dismissal barriers. Subsequently, formal employment increased by around five percent and unemployment fell from 10 percent in 2000 to around six percent in 2011 ([World Bank, 2015](#)). The overall labor participation rate has remained stable at 73–75 percent over this period.<sup>7</sup>

Hyperinflation also encouraged the adoption of automatic wage adjustment practices. From 1980 to 1989, yearly inflation averaged 355 percent, which was followed by a yearly average of 1,667 percent between 1990 and 1994 ([World Bank, 2015](#)). As a result, wage indexation to the minimum wage became the norm, with labor payments being adjusted first annually and then on a monthly basis proportionately to the previous period's realized inflation rate. In 1994, hyperinflation finally subsided with the introduction of the "Real Plan". This ambitious stabilization program introduced a gradual float of the local currency, tightened monetary and fiscal policy, and lowered inflation below two-digits.

In parallel to monetary stabilization, Brazil also undertook trade liberalization reforms during this period. Starting with initially high import tariffs that had substituted import bans from the

---

<sup>6</sup>Bioto and Marcelino (2011) argue that there has been an uptake in labor strike activity in Brazil since the year 2000.

<sup>7</sup>Labor force as a percentage of total population aged 15–64, from OECD Employment and Labor Market Statistics.

previous decade, a series of trade liberalization bills in the late 1980s eliminated selected tariffs and eradicated quantitative import controls. When social democrat Fernando Henrique Cardoso became president in 1995, he strengthened this agenda with a reduction of tariff and non-tariff trade barriers to one tenth of their levels in 1987 (Pavcnik et al., 2004). The opening up to trade over the last 25 years has been frequently cited as a major contributor to the country's growth in total factor productivity (TFP)(Ferreira and Rossi, 2003; Ferreira, Leite and Wai-Poi, 2007; Moreira, 2004; Muendler, 2004; Córdova and Moreira, 2003). In addition, Helpman et al. (2013) argue that trade reforms contributed to the rise in income inequality seen in the late 1980s and early 1990s, and later to the start of the decline in wage dispersion in 1995.

Health, education and other social programs began expanding during the late 1990s, a trend that strengthened once the left-wing Workers' Party ascended to power in 2003. It doubled social expenditure as a fraction of GDP and, although it remains less than one percent, it is often portrayed as an important contributor to the reduction in household income inequality.<sup>8</sup> The reach of the public cash transfer program, Bolsa Família, increased to cover 11 million families in 2006, which comprised nearly 25 percent of the total population (Barros et al., 2010). Education spending increased reaching 5.5 percent of GDP in 2009 (compared to 3.5 percent in 2000 and 5.7 percent among G20). As we discuss in Section 4 this is reflected in a rapidly rising share of the labor force with a high school degree. Moreover, the quality of education relative to other countries, as measured by the international PISA scores, has also improved, with Brazil having the greatest increase in mathematics among 65 countries since 2003 (OECD, 2012).

The Worker's Party complemented social policies with minimum wage increases above the previous upward trend. Within their first year in office, they established a 20 percent increase in 2003 and continued to implement yearly increases averaging over 10 percent during the next 10 years. As a result, the minimum to median wage in Brazil increased from around 34 percent in 1996—similar to U.S. levels—to over 50 percent, which is close to the level in France. Engbom and Moser (2015) argue that this large increase in the minimum wage can explain a significant fraction of the reduction in earnings inequality in Brazil over the 1996–2012 period, while being consistent with the other facts we document in the current paper.

With this brief overview of recent developments in Brazil, we turn to a discussion of the data

---

<sup>8</sup>Using household data, Barros et al. (2010) estimate that social programs accounted for about 20 percent of the decline in household income inequality.

we use to decompose the decline in inequality experienced in Brazil over the past two decades.

### 3 Data

Our analysis uses two confidential administrative datasets from Brazil: the *Relação Anual de Informações (RAIS)* contains earnings and demographic characteristics of workers as reported by employers, and the *Pesquisa Industrial Anual Empresa (PIA)* contains detailed information on revenues and costs of large firms in Brazil’s mining and manufacturing sectors. To make the reader familiar with these confidential data, we briefly discuss their collection, coverage, variable definitions, and sample selection.

#### 3.1 Description of linked employer-employee data (RAIS)

**Collection and coverage.** The RAIS data contains linked employer-employee records that are constructed from a mandatory survey filled annually by all registered firms in Brazil and administered by the Brazilian Ministry of Labor and Employment (*Ministério do Trabalho e Emprego*, or *MTE*). Data collection was initiated in 1986 within a broad set of regions, reaching complete coverage of all employees at formal establishments of the Brazilian economy in 1994.<sup>9</sup> Fines are levied on late, incomplete, or inaccurate reports, and as a result many businesses hire a specialized accountant to help with the completion of the survey. In addition, MTE conducts frequent checks on establishments across the country to verify the accuracy of information reported in RAIS, particularly with regards to earnings, which are checked to adhere to the minimum wage legislation.<sup>10</sup>

The RAIS contains an anonymized, time-invariant person identifier for each worker, which allows us to follow individuals over time. It also contains anonymized time-invariant establishment and firm IDs that we use to link multiple workers to their employers and follow those over time. Although it would be possible to conduct part of our analysis at the establishment instead of firm level, this paper focuses on firms for three reasons. First, to the extent that there is substan-

---

<sup>9</sup>Because registration with the central tax authorities is necessary and sufficient for a firm to be surveyed, the RAIS covers only workers in Brazil’s formal sector. Complementing our analysis with data from the Brazilian household survey *Pesquisa Nacional por Amostra de Domicílios (PNAD)*, we find that the formal sector employment share among male workers of age 18–64 grew from 64 to 74 percent between 1996 and 2012. Differential inequality trends between formal and informal sector workers are discussed at more length in Engbom and Moser (2015).

<sup>10</sup>In addition to being fined, non-compliant firms are added to a “Black List of Slave Work Employers,” made available publicly under law Decree No. 540/2004. A recent version of the list dated March 2015 is available from Brazilian television news channel Repórter Brasil at [http://reporterbrasil.org.br/documentos/lista\\_06\\_03\\_2015.pdf](http://reporterbrasil.org.br/documentos/lista_06_03_2015.pdf).

tial variation in pay across establishments within firms, our firm-level analysis provides a lower bound on the importance of workers' place of employment.<sup>11</sup> Second, we think that many of the factors that could give rise to employer-specific components of pay including corporate culture, company leadership, etc., act at the firm level. Additionally many regulations targeting pay policies differ as a function of firm-level employment, not establishment-level employment. Third, we will later use data on firm characteristics such as financial performance that are not available at the establishment-level.

**Variable definitions.** For each firm at which a worker was employed during the year, the RAIS contains information on the start and end date of the employment relationship, the amount the worker was paid and a broad set of worker and job characteristics. Reported earnings are gross and include regular salary payments, holiday bonuses, performance-based and commission bonuses, tips, and profit-sharing agreements. Although this is a broad measure of earnings, it does not contain other sources of income such as capital income or in-kind transfers. We divide total earnings from an employment relationship in a given year by the duration of the job spell.<sup>12</sup> This accounts to some extent for labor supply. As hours worked only exists for some years, we do not use this to construct a measure of per hour pay. Instead, to limit the impact of unmeasured labor supply differences, we focus on adult males.<sup>13</sup>

We define a consistent age variable by calculating the year of birth for any observation, and then setting an individual's year of birth as the modal implied value and finally reconstructing age in each year using this imputed year of birth.<sup>14</sup> Because age is only reported in bins prior to 2002, we code all subsequent years into the same age bins (18–24, 25–29, 30–39, 40–49, and 50–64 years old).

We define a consistent measure of years of schooling by first setting it to its modal value within a year in case of multiple job spells in a year and then ensuring that the years of schooling are non-decreasing across years. Subsequently, we define four education groups based on attained degree

---

<sup>11</sup>As we will show later, however, the explanatory power of our model incorporating firm and person effects is high, leaving little variation to be explained by separate establishment level effects.

<sup>12</sup> That is, if an employment relationship is reported as active for seven months during the year, we divide total earnings reported for that employment relationship for that year by seven.

<sup>13</sup>In the years for which we have data on hours, we find relatively little variation in hours, with most adult males reporting 44 hours of work a week.

<sup>14</sup>We use age instead of experience throughout our analysis; results are similar using age plus six minus years of education as a measure of experience.

implied by the reported number of years of schooling and the education system in Brazil (primary school, middle school, high school, and college).

The data also contain information on detailed occupation classification of the job and detailed sector classification of the employer establishment. Both the industry and occupation classification systems underwent a significant change during the period we study. For occupations, we use the pre-2003 classification (*Classificação Brasileira de Ocupações*, or *CBO*) at the one-digit level. We also use two-digit sectoral classifications (*Classificação Nacional de Atividades Econômicas*, or *CNAE*) according to the pre-2003 period. We make occupations and sectors reported for 2003–2012 consistent with the older CBO and CNAE classifications by using conversion tables provided by IBGE. In order to achieve a high level of consistency between the old and the new classification schemes, we cannot go less coarse than one digit occupation and two digit sector, but we believe that for the purpose of this paper this restriction is not of major importance.

Our firm size measure is the number of full-time equivalent workers during the reference year. Importantly, we calculate this prior to making any sample restrictions so that it reflects to the greatest extent possible the total amount of labor used by the firm during the year. We calculate it as the total number of worker-months employed by the firm during the year divided by 12.

**Sample selection.** We exclude observations with either firm IDs or worker IDs reported as invalid as well as data points with missing earnings, dates of employment, educational attainment or age. Together, these cleaning procedures drop less than one percent of the original population, indicative of the high quality of the administrative dataset. Subsequently, to limit the computational complexity associated with estimating our model, we restrict attention to one observation per worker-year. We impose this restriction by choosing the highest-paying among all longest employment spells in any given year. As the average number of jobs held during the year is 1.2 and there is not trend in this, we do not believe that loosening this restriction would meaningfully affect our results.

Finally, we restrict attention to adult male workers of age 18–64. We make this restriction as a trade-off between our results being comparable to a large part of the literature focusing on prime age males on the one hand and to obtain as complete as possible coverage of the changes in the Brazilian wage structure over the period on the other.<sup>15</sup> We have tried alternative sample

---

<sup>15</sup>The restriction to only male workers has the advantage of avoiding issues with changing patterns of female labor supply and labor market discrimination. In a separate ongoing research project, we investigate the degree to which

restrictions, including focusing on only prime age males of age 18–49, as well as both male and female adult workers.

**Descriptive statistics.** Table 1 provides key summary statistics for the RAIS data for six sub-periods of five years each with one year overlap between adjacent periods, namely 1988–1992, 1992–1996, 1996–2000, 2000–2004, 2004–2008, and 2008–2012. Since our analysis focuses on adult males and adult males working for large manufacturing and mining firms, we provide a brief comparison of these subpopulations to the overall population of formal sector employees. As we will be primarily concerned with the later four subperiods during which inequality declined markedly and for which we have firm level data, we focus our discussion on these periods.

Panel A shows statistics for the overall formal sector work force in Brazil and Panel B for the subpopulation of adult males. Adult males are consistently about 0.3–0.4 years older than the population average. They also have 0.78 years of schooling less than the overall sample in the 1996–2000 subperiod; this gradually drops to 0.65 years in the last subperiod. Finally, adult males earn about eight to nine log points more than the overall population, but the variance of log earnings is very similar to the overall population.

Panel C presents statistics on the subpopulation of adult males working at large mining and manufacturing firms. Adult males in the PIA subpopulation are about 0.8 years younger than all adult males in the 1996–2000 subperiod, which gradually increases to 1.3 years younger in the last subperiod. They are similar to all adult males in terms of education. The PIA sample of adult males earned on average 27 log points more than all adult males in the 1996–2000 subperiod; this declined to only a 19 log point premium in the last subperiod. Finally, they display a two log point higher standard deviation of log earnings in the 1996–2000 period, which increases to four log points in the last subperiod.

### 3.2 Description of firm characteristics data (PIA)

**Collection and coverage.** The PIA data contain information on firm financial characteristics from 1996–2012. The dataset is constructed by the Brazilian National Statistical Institute (*Instituto Brasileiro de Geografia e Estatística*, or IBGE) based on annual firm surveys in the manufacturing and mining sector. This survey is mandatory for all firms with either more than 30 employees or 

---

firm-level average pay and profit sharing differs between male and female employees and how the gender pay gap has evolved over time.

Table 1. RAIS summary statistics

	(1)	(2)	Earnings		Age		Schooling	
			(3)	(4)	(5)	(6)	(7)	(8)
	# Worker-years	# Unique workers	Mean	St.d.	Mean	St.d.	Mean	St.d.
PANEL A. ALL FORMAL SECTOR WORKERS (RAIS)								
1988–1992	165.5	41.9	1.10	0.86	31.91	11.47	7.65	4.45
1992–1996	162.1	43.4	1.18	0.86	33.20	11.32	8.08	4.41
1996–2000	174.6	47.0	1.19	0.84	33.68	11.27	8.60	4.27
2000–2004	202.7	52.7	1.00	0.80	34.02	11.33	9.49	4.05
2004–2008	254.2	62.7	0.81	0.74	34.26	11.48	10.25	3.78
2008–2012	326.5	76.2	0.71	0.71	34.55	11.66	10.78	3.52
PANEL B. ADULT MALE WORKERS								
1988–1992	86.5	25.5	1.24	0.87	33.26	10.82	7.04	4.30
1992–1996	87.3	26.4	1.29	0.87	33.88	10.79	7.37	4.27
1996–2000	92.7	28.8	1.27	0.85	33.97	10.78	7.82	4.16
2000–2004	105.3	32.5	1.07	0.80	34.14	10.92	8.70	4.00
2004–2008	126.9	37.3	0.88	0.75	34.44	11.11	9.50	3.79
2008–2012	154.2	43.9	0.80	0.72	34.91	11.34	10.13	3.59
PANEL C. ADULT MALE WORKERS AT LARGE MANUFACTURING AND MINING FIRMS (PIA)								
1996–2000	16.6	6.3	1.54	0.87	33.20	10.07	7.83	4.05
2000–2004	18.0	6.8	1.29	0.85	33.04	10.20	8.75	3.91
2004–2008	23.2	8.5	1.09	0.80	33.11	10.45	9.41	3.77
2008–2012	26.9	9.9	0.99	0.76	33.60	10.70	10.04	3.60

Notes: The number of worker-years and number of unique workers are reported in millions. Statistics on earnings are in log multiples of the current minimum wage, schooling is in years. Panel A includes all workers in the RAIS dataset. Panel B includes male workers that are between 18 and 64 years old. Panel C includes male workers age 18–64 working at large manufacturing and mining firms included in the PIA firm characteristics data. Means are computed by period. The standard deviation is calculated by first demeaning variables by year and then pooling the years within a subperiod.

above a revenue threshold as well as for an annual random sample of smaller firms.<sup>16</sup> As with RAIS, completion of the survey is mandatory and non-compliance is subject to a fine by national authorities. Each firm has a unique, anonymized identifier, which we use to link firm characteristics data from PIA data to worker-level outcomes in the RAIS data.

**Variable definitions.** The PIA dataset includes a breakdown of operational and non-operational revenues, costs, investment and capital sales, number of employees and payroll. All nominal values are converted to real values using the CPI index provided by the IBGE. Instead of the measure

<sup>16</sup>The revenue threshold for inclusion in the deterministic survey has grown over the years, standing at USD300,000 in 2012.

of firm size in the PIA, we prefer our measure of full-time-equivalent employees constructed from the RAIS as it accounts for workers only employed during part of the year. We define operational costs as the cost of raw materials, intermediate inputs, electricity and other utilities, and net revenues as the gross sales value due to operational and non-operational firm activities net of any returns, cancellations, and corrected for changes in inventory.<sup>17</sup> We finally construct value added as the difference between net revenues and intermediate inputs, and value added per worker as value added divided by full-time equivalent workers. This is our main measure of firm productivity. We have also constructed alternative measures of firm productivity by cleaning value added per worker off industry-year effects and some measures of worker skill. In our main analysis, we focus on “raw” value added per worker and present results containing these alternative measures in the Appendix.

Our productivity measure differs from the commonly used total factor productivity (TFP) (Bartelsman, Haltiwanger and Scarpetta, 2009, 2013) since it does not control for capital intensity. A major reason for this is that we do not have data on capital, only on investment. To construct a measure of the capital stock, we would need to assume a depreciation rate to be able to impute capital using reported investment. We would also need to impute capital in 1996 since we do not have data prior to that, as well as for any firm that enters the PIA population. We have constructed such a measure of the capital stock using an assumed annual depreciation rate of five percent and using data on the aggregate capital stock at the subsector level.<sup>18</sup> However, the multiple imputations required to obtain capital as well as the fact that the investment data is incomplete for many firms lead us to prefer value added per worker as our measure of firm productivity.<sup>19</sup>

**Sample selection.** The PIA firm survey spans the universe of large firms (as defined above) in Brazil’s manufacturing and mining sectors in addition to a random sample of smaller firms. Because parts of our analysis make use of the panel dimension on the firm side and to avoid issues

---

<sup>17</sup>We have explored alternative revenue definitions such as only restricting attention to operational revenues or excluding certain types of non-operational revenues. Such robustness checks yield very similar results to what we report below.

<sup>18</sup>Each new firm starts with an initial capital equal to its current net investment plus a share of total capital in its subsector. The shares are given by taking the share of capital at a firm to be proportional to the share of total net revenues assuming a firm-level production function of the form  $y = Ak^\alpha$  for  $\alpha = 1/3$ . Firms entering the PIA at a later year are initiated by applying the same method to get those firms’ capital stock proportional to scaled firm revenues relative to the subsector total.

<sup>19</sup>In addition, several bargaining models of the labor market have in common that workers and capital owners split the surplus from production, and value added per worker is arguably the best measure of that surplus. Thus, to the extent that such models well describe Brazilian labor markets, value added per worker is an important metric.

with excessive sample attrition related to our later estimation procedure, we focus our analysis on the deterministic set of relatively large firms.

**Descriptive statistics.** Table 2 shows key summary statistics on firms during the four periods for which we have firm financial data: 1996–2000, 2000–2004, 2004–2008, and 2008–2012. All results are weighted by the number of full-time equivalent workers employed by the firm. The number of firms in the PIA increased by 57 percent between the first and the last period. The average firm size increased by 32 percent and average real value added per worker grow by 27 percent. There is significant dispersion in both log firm size and log value added per worker across firms, with the standard deviation of the former being close to two and that of the latter exceeding one. Furthermore, there is no evidence of convergence in either measure. The standard deviation of firm size monotonically increases whereas the standard deviation of value added per worker first increases rapidly, then falls again in the last subperiod. To the extent that firm characteristics matter for employees’ labor remuneration, these results suggest that the decline in earnings inequality in Brazil cannot be explained by declining dispersion in these characteristics over time.

Table 2. PIA summary statistics

	(1)	(2)	Firm size		Value added p.w.	
			(3)	(4)	(5)	(6)
	# Firm-years	# Unique firms	Mean	St.d.	Mean	St.d.
1996–2000	142.4	51.1	6.34	1.80	11.15	1.13
2000–2004	168.7	59.9	6.26	1.85	11.19	1.32
2004–2008	202.4	73.0	6.49	1.96	11.22	1.34
2008–2012	230.2	80.2	6.62	2.05	11.30	1.31

Note: The number of firm-years and number of unique firms are reported in thousands. Firm size is the log number of full-time equivalent employees. Value added per worker is the log of real value added per worker. Means and standard deviations are weighted by the number of full-time employees. The standard deviation is calculated by first demeaning variables by year and then pooling the years within a subperiod.

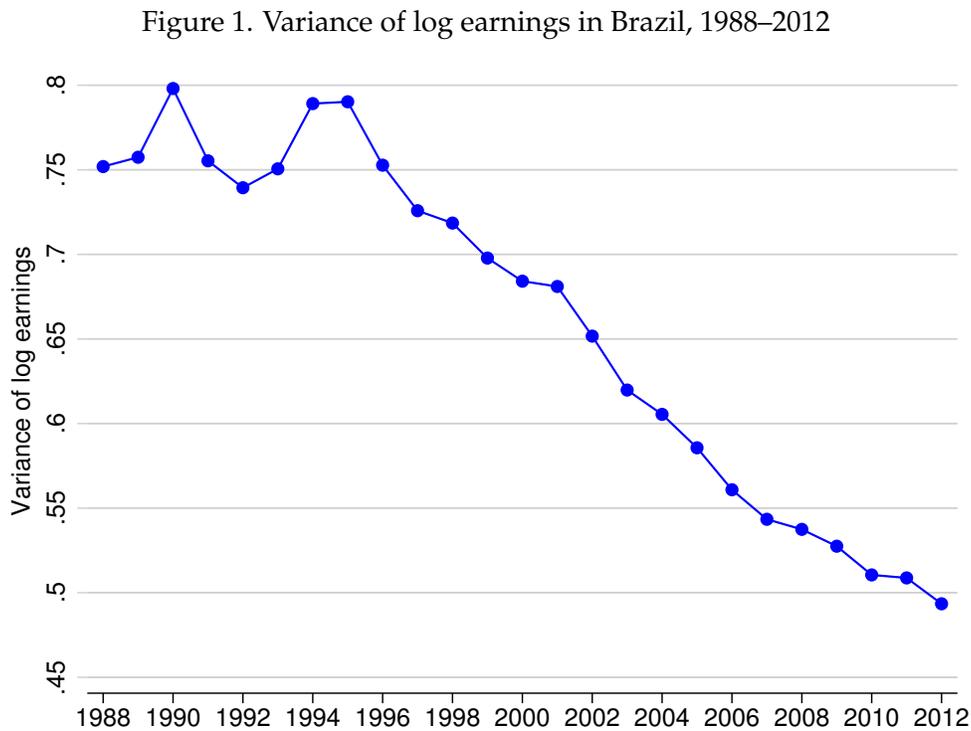
## 4 Inequality trends in Brazil from 1988–2012

In this section, we first document Brazil’s rapid decline in earnings inequality. We then demonstrate that the decline in inequality in Brazil occurred throughout a large part of the earnings distribution. Subsequently, we present results from a series of Mincer regressions, which provides a first look at possible factors behind the decline. Finally, we provide some suggestive evidence of

firms being an important source of inequality as well as a factor behind the decline in inequality in Brazil.

#### 4.1 The evolution of earnings inequality

Starting from high initial levels, Brazil experienced a rapid and steady decline in earnings inequality from 1995 onwards. The inequality decline followed years of stable or slightly increasing inequality between the late 1980s and the mid-1990s. While the inequality decline is apparent in many measures, Figure 1 plots the variance of log earnings, which is a common inequality measure in the labor literature. Between 1996 and 2012, the variance of log earnings in Brazil declined by 26 log points or 35 percent, from 0.75 to 0.49 over the period. To put this decline into context, the U.S. saw an increase in the same measure of 22 log points from 1968–2006.

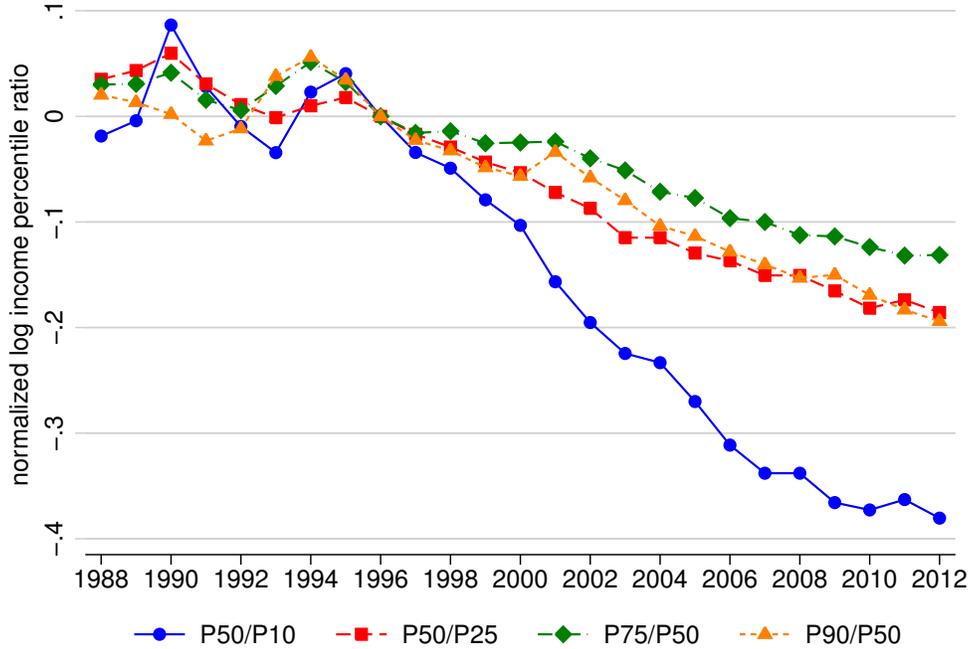


#### 4.2 Compression in different parts of the earnings distribution

Figure 2 plots the log percentile ratios of earnings, normalized to zero in 1996. Two striking facts emerge from the graph: First, there was widespread compression in the distribution of earnings—

inequality declined past the 75th percentile. Second, the amount of compression gradually declines as one moves further up in the distribution. For instance, whereas the log 90–50 percentile ratio falls by 20 log points, the log 50–10 ratio falls by a remarkable 35 log points. Similarly, compression in the log 50–25 percentile ratio exceeds compression in the log 75–50 ratio.<sup>20</sup>

Figure 2. Log percentile ratios of the earnings distribution in Brazil (1996 = 0)



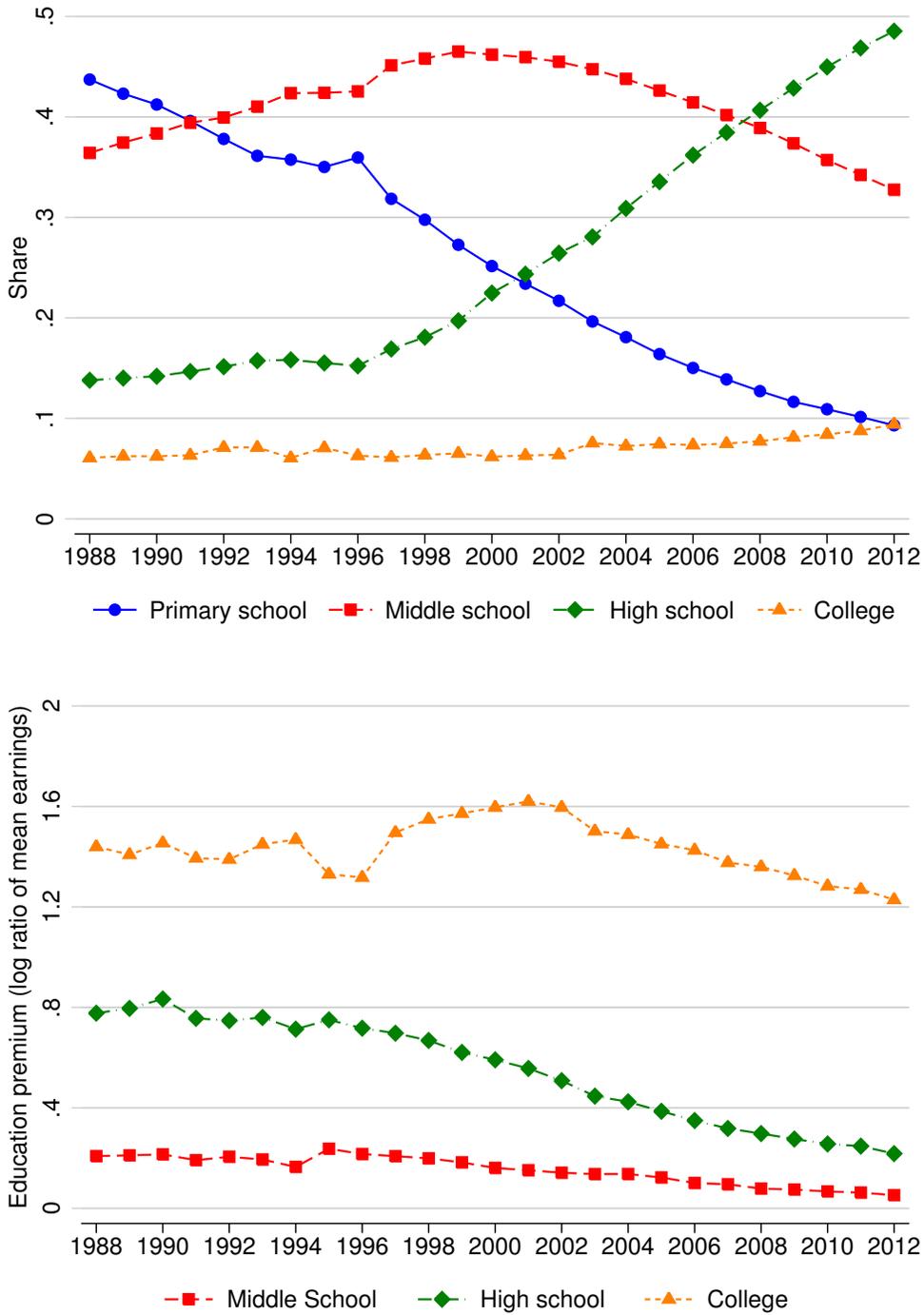
### 4.3 The (un)importance of worker observables

One candidate explanation for the decline in earnings inequality is increasing educational attainment. As can be seen in the left panel of Figure 3, the fraction of the Brazilian formal sector workforce with a high school degree rose rapidly during this time, while the fraction with primary school fell sharply (the fraction with a middle school degree and a college degree remained relatively flat). There were also important changes to the premia associated with a higher degree, as can be seen in the right pane of Figure 3. In particular, the premia associated with a middle

<sup>20</sup>For ease of interpretation and to show that there was significant real wage growth at all percentiles of the income distribution over the period for which such data exist, Appendix A.1 plots the real percentile earnings levels evolution from 1996–2012, normalized to zero in 1996.

school and high school degree relative to the lowest education group fell rapidly over the past 20 years.

Figure 3. Educational attainment (top) and education premia (bottom)



To provide a first look at whether changes in observable worker characteristics were an impor-

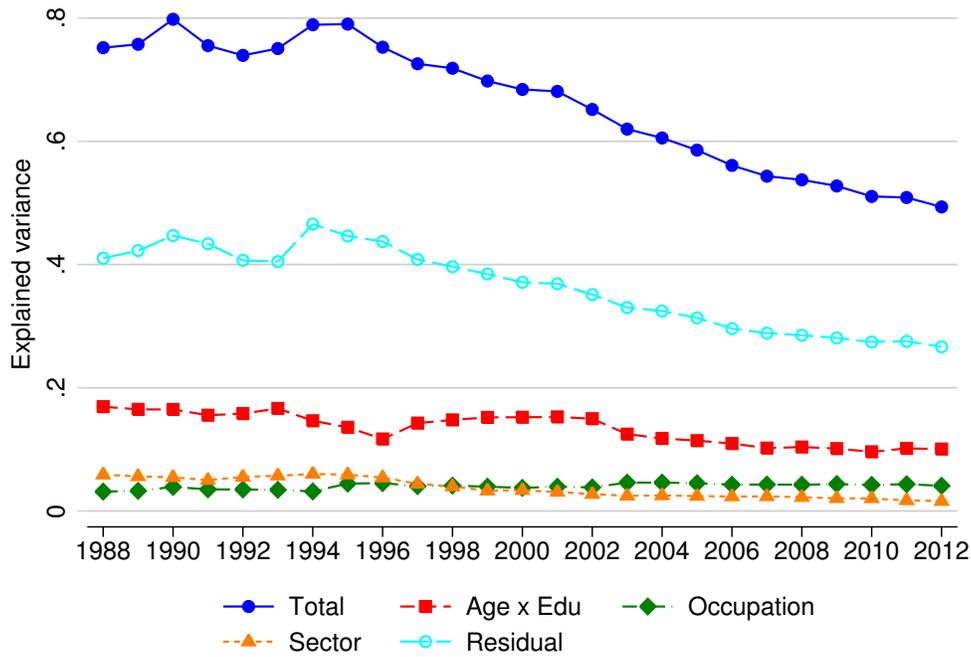
tant driver of earnings inequality in Brazil over this period, we run a series of Mincer regressions. In particular, we regress log earnings of individual  $i$  in year  $t$ , denoted  $y_{it}$ , on five age group dummies interacted with nine education group dummies, two digit occupation dummies and two digit sector dummies,

$$\log(y_{it}) = age_{it} \times edu_{it} + occ_{it} + sec_{it} + \varepsilon_{it}$$

Note that all explanatory variables are allowed to vary freely by year. Based on this regression, we calculate the predicted value due to each component and report the variance of these predicted values.

Figure 4 plots the results. In levels, worker observables jointly explain about 45 percent of the overall variance in log earnings. This does not change much over time, and hence worker observables explain close to 45 percent of also the fall in inequality. Decomposing this, age and education account for roughly 20 percent of the variance of log earnings and 27 percent of the decline. Inequality between occupations increases relative to overall inequality from four to eight percent of total variance (and in fact it also grows slightly in absolute terms). The fraction of total inequality explained by differences in means across sectors falls from eight to three percent—this accounts for 12 percent of the overall decline in variance. Finally, covariances between the explanatory variables increase slightly in importance from 11 percent to 14 percent of total variance and explain four percent of the overall fall in inequality. We conclude that even when controlling for detailed worker characteristics, more than half of the level of inequality as well as its decline is residual in nature.

Figure 4. Variance decomposition from Mincer regressions



#### 4.4 The evolution of earnings inequality between and within firms

As a first step towards understanding the role of firms for earnings inequality, we investigate the variance of earnings within and between firms. To this end, we define between-firm inequality as the variance of the average log earnings at the firm across firms (weighted by firm size) and within-firm inequality as the variance of the difference between workers' log earnings and the average log earnings at their firm. Based on these definitions, one could imagine two hypothetical polar extremes. First, average earnings could be identical across firms so that overall earnings inequality is completely due to variance in earnings within firms. In this case, a firm is just a microcosm of the overall economy. Second, all workers could earn the same wage within the firm so that inequality arises entirely due to differences in earnings across firms. In reality, the question is which channel is quantitatively most important.

Formally, we follow Bloom et al. (2015) in letting  $y_{ijt}$  denote earnings of worker  $i$  employed by

firm  $j$  in year  $t$ , then:

$$y_{ijt} = \underbrace{\bar{y}_t}_{\text{economy average}} + \underbrace{(\bar{y}_t^j - \bar{y}_t)}_{\text{employer deviation}} + \underbrace{(y_{ijt} - \bar{y}_t^j)}_{\text{worker deviation}}$$

Re-arranging and taking variances on both sides we get

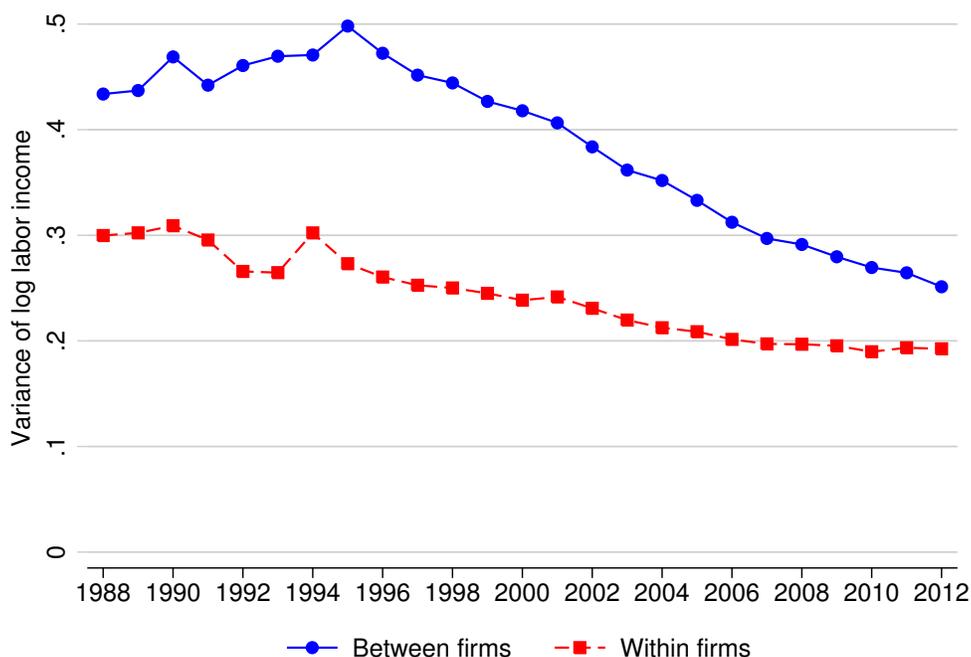
$$\text{Var}(y_{ijt} - \bar{y}_t) = \text{Var}(\bar{y}_t^j - \bar{y}_t) + \text{Var}(y_{ijt} - \bar{y}_t^j) + \underbrace{2\text{Cov}(\bar{y}_t^j - \bar{y}_t, y_{ijt} - \bar{y}_t^j)}_{=0}$$

where the last term is zero by construction. Simplifying, we have our decomposition of the overall variance of earnings into between- and within-firms inequality:

$$\underbrace{\text{Var}(y_{ijt})}_{\text{overall}} = \underbrace{\text{Var}(\bar{y}_t^j)}_{\text{between firms}} + \overbrace{\text{Var}(y_{ijt} | i \in j)}^{\text{within firms}}$$

Figure 5 plots this decomposition over time in Brazil. We note two insights: Firstly, there is significant variability in earnings within firms, but an even greater amount of earnings inequality across firms. Secondly, although both measures of inequality fell during this time, the decline was particularly pronounced across firms: inequality across firms declined by 25 log points or 45 percent between 1988 and 2012, whereas within-firm inequality dropped by 10 log points or 33 percent.

Figure 5. Variance of log earnings within and between firms



Another way of illustrating the importance of firm is to compare earnings growth of different workers with average earnings growth at their employers. Making use of the link between workers and firms in the matched employer-employee data (RAIS), one can ask how much of the earnings growth accruing to a certain income group was mediated by rises in average pay at firms employing workers from that income group.<sup>21</sup>

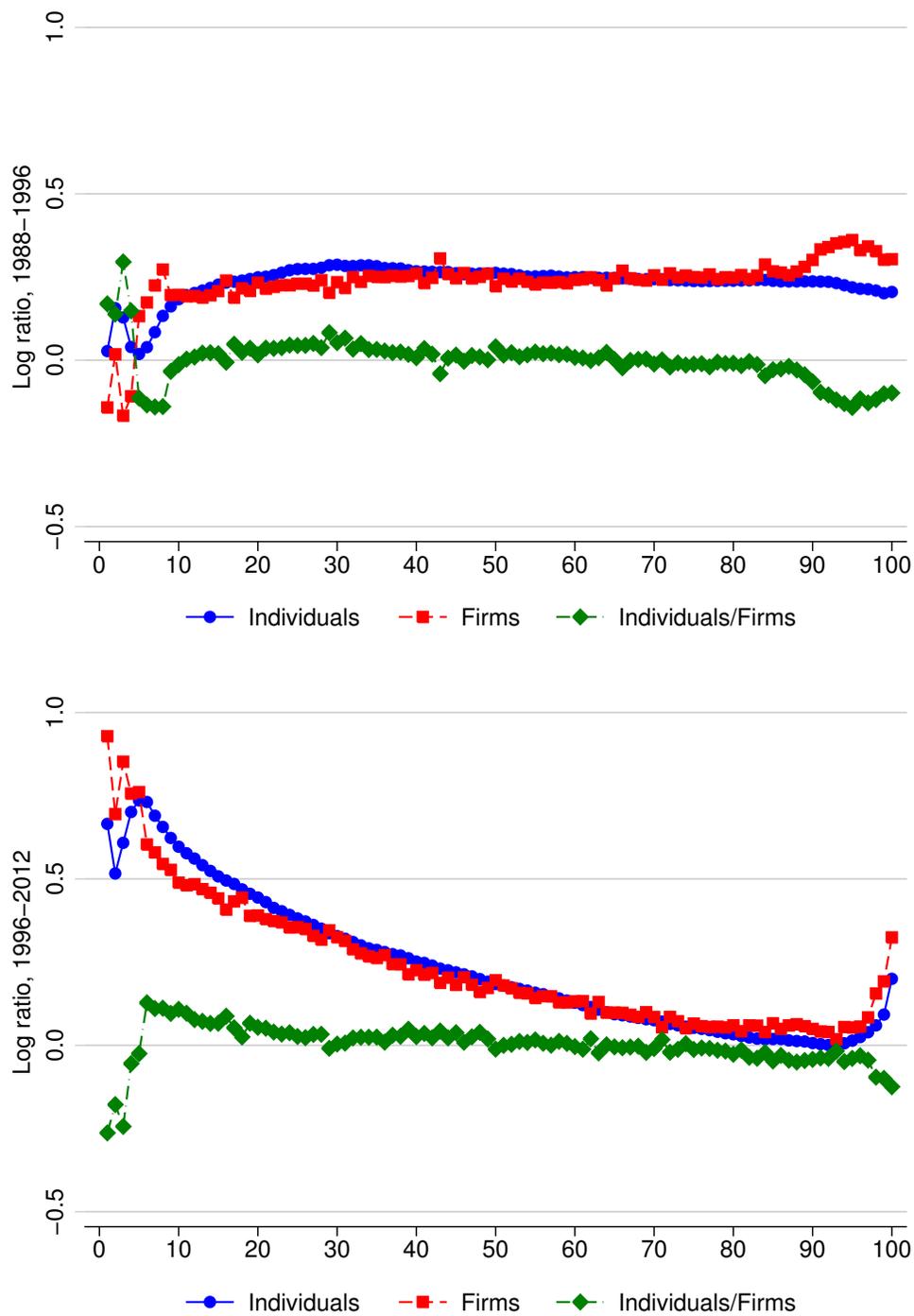
The results of this exercise are shown in Figure 6, first for the period 1988–1996, when earnings inequality remained roughly constant, and then for 1996–2012, when earnings inequality declined rapidly. Average earnings growth (solid blue line with circles) was relatively evenly distributed throughout the earnings distribution between 1988 and 1996, and firm average pay (solid red line with squares) grew equally in line with the growth rate of wages. The period from 1996–2012, on the other hand, was marked by a rapid catch-up of the lowest earnings groups, which in turn is almost entirely explained by the growth of firm average earnings among those groups. Throughout both periods, there were no significant changes in within-firm earnings inequality (solid green line with diamonds).

Although informative, however, these type of decompositions of raw earnings cannot neces-

<sup>21</sup>Similar calculations are presented for the U.S. in Barth, Davis and Bryson (2014) and Bloom et al. (2015).

sarily be interpreted as firms differing fundamentally in the way they compensate their workers. The reason is that some firms could hire workers who always get paid more regardless of where they work (maybe because they are more productive, have a higher bargaining power, etc). In this case differences in pay across firms would arise as a result of recruitment policies and not pay policies. The next section formalizes our approach to identifying the importance of firm pay policies for earnings inequality.

Figure 6. Individual labor earnings growth between and within firms, 1988–1996 (top) and 1996–2012 (bottom)



## 5 Empirical framework

For a long time, economists have recognized that worker observables fail to explain a large fraction of the variance of earnings. As we showed above, this is true also for Brazil—even with detailed occupation and sector controls, Mincer regressions explain less than half of the overall variation in earnings in Brazil. Furthermore, a recent literature has argued that there are important differences across firms in terms of pay (Abowd, Kramarz and Margolis, 1999), and we presented some evidence above suggesting that firms might be an important determinant of earnings also in Brazil. Motivated by these insights, we estimate two-way fixed effects econometric models controlling for both unobserved worker and firm heterogeneity. To be able to speak to changes over time in the components of inequality, we estimate our model separately in six sub-periods covering 1988–1992, 1992–1996, 1996–2000, 2000–2004, 2004–2008, and 2008–2012, respectively. Subsequently, we correlate the estimated firm and worker effects with observed characteristics of firms and workers in order to investigate what led to changes in the firm and worker component of pay over time.

### 5.1 Introducing the two-way fixed effects model

In order to identify the two-way fixed effects framework pioneered by Abowd, Kramarz and Margolis (1999), one needs to observe a panel of workers with the ability to link multiple workers to the same firm. Our data satisfy these requirements. Within each subperiod, we observe a large number  $I$  of workers for up to five years while working at  $J$  firms for a total of  $N$  worker-years. Let  $J(i, t)$  give the employer of worker  $i$  in year  $t$ . We assume that the log earnings of individual  $i$  in year  $t$  within a period time window  $p$ , denoted  $y_{it}$ , can be written as the sum of a worker effect,  $\alpha_i^p$ , a firm effect,  $\alpha_{J(i,t)}^p$ , a time dummy,  $\gamma_t$ , and an error term,  $\varepsilon_{it}$ . Although our current paper does not provide a microfoundation to this reduced form empirical model, in subsequent work we show how it can be rationalized in a frictional labor market with firm productive heterogeneity and worker ability differences (Engbom and Moser, 2015).

Our specification does not control in this first stage for observable worker and firm characteristics. Instead, we correlate the estimated fixed effects with observable characteristics of workers and firms in a second stage of our analysis. We prefer this specification to avoid identifying these effects off changes within workers and firms during the limited time frame of each subperiod. With regards to age, we additionally notice that unrestricted age controls would be perfectly

collinear with person effects and the time dummies.<sup>22</sup> Similarly, we also correlate worker observables, including age, education, and occupation, with the estimated worker effect in a second stage. We argue that any error due to growing earnings with age is likely to be of second order importance within our five year subperiods. We thus estimate

$$\log y_{it} = \alpha_i^p + \alpha_{J(i,t)}^p + \gamma_t + \varepsilon_{it}$$

for  $t \in p = \{t_1, \dots, t_5\}$  and where  $\alpha_i^p$  denotes the individual fixed effect of worker  $i$  in period  $p$ ,  $\alpha_{J(i,t)}^p$  denotes the firm effect of the employer of worker  $i$  at year  $t$ ,  $\gamma_t$  is a year dummy, and  $\varepsilon_{it}$  is an error term that satisfies the strict exogeneity condition  $\mathbb{E}[\varepsilon_{it} | i, t, J(i, t)] = 0$ .<sup>23</sup>

As shown by [Abowd, Kramarz and Margolis \(1999\)](#), worker and firm effects can only be separately identified within a set of firms and workers connected through the mobility of workers. Table 3 presents summary statistics on the largest set of workers in each subperiod—this covers 97–98 percent of all workers in each subperiod. Given that it covers such a large fraction of all adult males, it is not surprising that this subpopulation looks very similar to the overall population in all observable dimensions. Thus the restriction to the largest connected set imposed in the rest of our analysis appears to be innocuous.

Table 3. Summary statistics on workers in largest connected set (as a fraction of adult males)

	(1)	(2)	Earnings		Age		Schooling	
			(3)	(4)	(5)	(6)	(7)	(8)
	# Worker-years	# Unique workers	Mean	St.d.	Mean	St.d.	Mean	St.d.
1988–1992	85.4 (99%)	25.1 (98%)	1.3	0.9	33.3	10.8	7.0	4.3
1992–1996	85.6 (98%)	25.8 (98%)	1.3	0.9	33.9	10.8	7.4	4.3
1996–2000	90.2 (97%)	27.9 (97%)	1.3	0.8	34.0	10.8	7.8	4.2
2000–2004	102.1 (97%)	31.4 (97%)	1.1	0.8	34.2	10.9	8.7	4.0
2004–2008	123.7 (97%)	36.2 (97%)	0.9	0.8	34.4	11.1	9.5	3.8
2008–2012	151.0 (98%)	42.8 (98%)	0.8	0.7	34.9	11.3	10.1	3.6

Notes: We report in parentheses the proportion of the reported statistics relative to the group of adult males listed in Table 1. Earnings are in log multiples of the minimum wage, schooling is years of education.

As identification critically derives from workers switching between firms, Table 4 presents statistics on the fraction of switchers in each subperiod. The degree of labor mobility is high in Brazil with more than 30 percent of the population switching firm at some point in the subperiod.

<sup>22</sup>Although restrictions could be imposed to address this collinearity problem, we note that for instance the popular restriction advocated by [Deaton \(1997\)](#) requires many years to be well identified.

<sup>23</sup>We later discuss in greater detail the empirical validity of the assumption on the error term.

The average number of firms worked at during the five years in each subperiod is about 1.5. There is no strong trend in either statistic.

Table 4. Frequency of switches, by period

	(1)	(2)	(3)
	# Unique workers	Average # of jobs	% switchers
1988–1992	25.1	1.55	0.36
1992–1996	25.8	1.47	0.32
1996–2000	27.9	1.44	0.31
2000–2004	31.4	1.45	0.31
2004–2008	36.2	1.53	0.36
2008–2012	42.8	1.64	0.40

Note: Number of unique workers in millions. A switcher is defined as a worker who is associated with two or more employers during the period.

The assumption we impose on the error term is often referred to in the literature as that of requiring *exogenous mobility*. As explained by AKM, this rules out dependency of the error term on the worker effect, the firm effect or the time controls. In particular, a worker is not allowed to switch between firms based on the unobserved error term, because if say matches with a particularly poor match effect are more likely to break up, the residual of remaining matches does not have mean zero. Moreover, this assumption rules out assortative matching of the type found in Roy models, since these models emphasize the complementarity in matches between workers and firms, whereas the AKM framework imposes log-additivity between the two.

We investigate whether this assumption is violated in two ways. First, we follow [Card, Heining and Kline \(2013\)](#) in dividing estimated firm effects into quartiles and investigate whether the gain in the firm component of those switching between for instance the first and fourth quartile is similar to the loss of those making the reverse switch. To the extent that the labor market is better characterized by assortative matching as in Roy models, we would expect these to be very different. Second, we examine the distribution of error terms across worker and firm effects quantiles to check for systematic variation, which could be an indication that our log additive model is misspecified.

Based on our estimated equation, we decompose the variance of log earnings within any subperiod into the variance of the worker component, the firm component, the year trend and the residual, as well as the covariance between the worker and the firm component, the worker and

year component, and the firm and year component:

$$\begin{aligned} \text{Var}(\log y_{it}) = & \text{Var}(\alpha_i) + \text{Var}(\alpha_{J(i,t)}) + \text{Var}(\gamma_t) + 2\text{Cov}(\alpha_i, \alpha_{J(i,t)}) \\ & + 2\text{Cov}(\alpha_i, \gamma_t) + 2\text{Cov}(\alpha_{J(i,t)}, \gamma_t) + \text{Var}(\varepsilon_{it}) \end{aligned} \quad (1)$$

We note that sampling error in the estimated person and firm effects will cause us to overestimate the variance of worker and firm effects and induce a negative bias in the covariance between worker and firm effects. We do not attempt to correct for this. Instead, we assume that this error is constant over time, so that even if the level of our estimated variances is slightly overstated, the changes we document over time are still valid.<sup>24</sup>

## 5.2 Determinants of the estimated firm effects

In the second stage of our empirical investigation, we study how the estimated firm effects relate to observable measures of firm performance available in the PIA survey. In particular, we are interested in understanding what firm characteristics are related to pay, and whether any changes in the distribution of firm effects over time can be explained by underlying changes in firm characteristics or the way the labor market translates those into pay. Since the PIA only covers the set of large manufacturing and mining firms,<sup>25</sup> we are forced to restrict attention to only these firms and workers when linking firm effects to firm characteristics. We implement this by first estimating the AKM model for the universe of firms and then subsequently restricting attention to only large firms.

Consider a given subperiod and let  $a_j$  be the estimated firm component of pay,  $VA_j$  average log value added per worker during the subperiod and  $S_j$  a set of 26 two digit subsector controls. For each subperiod, we regress

$$a_j = \gamma_0 + \gamma_1 VA_j + S_j + \varepsilon_j$$

Notice that all regressions are run with sub-period averages of all variables. Additionally, we

---

<sup>24</sup> To better estimate firm effects, [Bonhomme, Lamadon and Manresa \(2015\)](#) suggest restricting attention to firms whose fixed effect is “well-identified” due to a high number of switchers. In practice, this procedure boils down to restricting attention to workers at firms with at least 10 switchers during the estimation period. With this restriction, we find a slightly more pronounced role for worker effects in explaining both the initial levels and the decline of earnings inequality between 1996–2000 and 2008–2012.

<sup>25</sup> As described in Section 3, we restrict attention to the deterministic stratum of PIA containing only large firms. We drop small firms contained in the random stratum to ensure that firms stay in the sample for multiple years for our estimation procedure below.

have considered versions including a range of other firm characteristics, but as these are only marginally important we do not show them here. Based on the above regression, we compute the variance due to value added per worker as

$$\text{Var}(\hat{a}_j) = (\hat{\gamma}_1)^2 \text{Var}(\text{VA}_j)$$

In a similar vein, we can compute the variance of firm effects due to sub-sector differences.

In order to isolate the importance of a compression in firm fundamentals versus a compression in the pass-through from such fundamentals to pay, we consider two counterfactuals. First, we keep the pass-through from value added per worker to pay,  $\hat{\gamma}_1$  constant at the estimated level in 1996–2000 and let the variance of value added per worker change as in the data. Second, we keep the variance of value added per worker at its 1996–2000 level and letting the estimated coefficient  $\hat{\gamma}_1$  change as in the data. Similarly, we decompose the change in the variance due to sub-sectors into changes in sub-sector pay premia versus a change in the sectoral composition of the economy. A comparison of the two counterfactuals allows us to address whether a change in the variance of firm pay is explained by changes in underlying firm characteristics or due to a change in the degree of pass-through from firm value added to worker pay.

### 5.3 Determinants of the estimated worker effects

In the second stage of our analysis, we also investigate what factors influence the worker component of pay. To this end, we regress the estimated worker effects,  $a_i$ , on five age bin dummies, four education dummies and ten occupation dummies:

$$a_i = \text{age}_i + \text{edu}_i + \text{occ}_i + \varepsilon_i$$

We report results both with and without occupation controls. We have also examined versions of this regression with age and education interacted as well as including sector controls, but neither of these alternatives changes the estimated results significantly. Based on the above regression estimates, we predict the variance due to each of age, education and occupation.

Subsequently, we decompose the evolution of the variance of the age, education and occupation components into that due to a change in the underlying distribution of such characteristics

versus a change in the return to them. That is, we first keep the estimated return to the characteristic of interest constant and change only the underlying distribution to match its evolution in the data. This shows how important changes in the distribution of worker characteristics were for the overall decline. Subsequently, we instead change the return to the characteristic of interest as in the data, while holding the underlying distribution constant. This allows us to evaluate how important changes in the return to age, education and occupation were for the overall decline.

## 6 Results

In this section, we first present the results from our two-way fixed effects model decomposing earnings inequality into a firm and a worker component. Subsequently, we investigate the sources of the firm and worker component of pay. Finally we provide additional results evaluating the assumptions imposed by our econometric model.

### 6.1 AKM decomposition

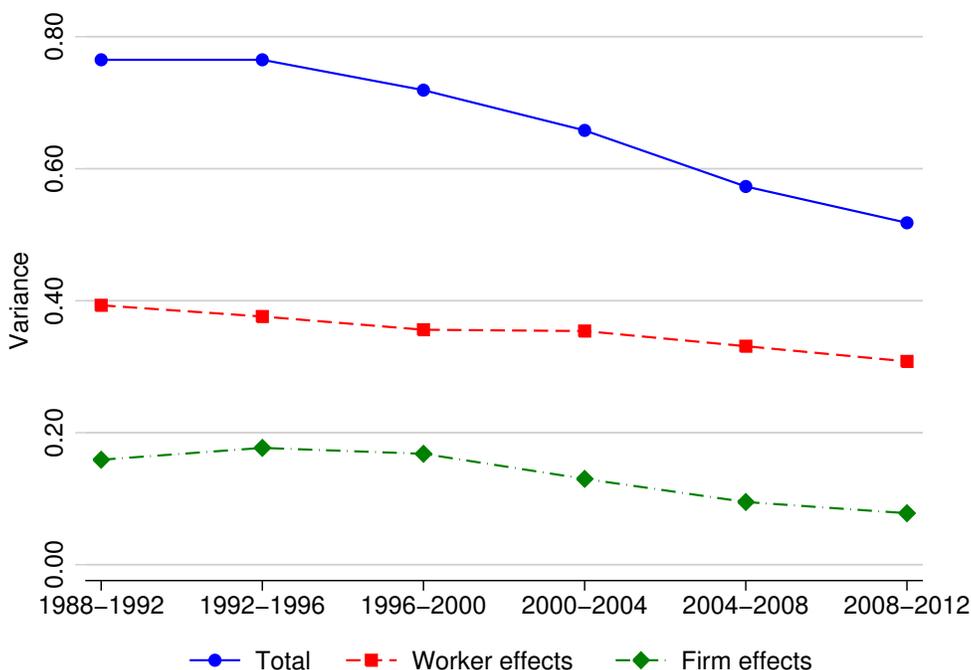
Table 5 presents the variance decomposition based on the estimation results of the AKM model in equation (1) for each of the six five-year subperiods between 1988 and 2012. To illustrate the relative importance of the various components of this decomposition over time, Figure 7 plots the variance of raw earnings (solid blue line with circles), the variance of estimated worker effects (dashed red line with squares), and the employee-weighted variance of firm effects (dash-dotted green line with diamonds) in each subperiod.

Two important results emerge from our analysis. First, although firm heterogeneity is a non-negligible source of earnings inequality *levels*, worker heterogeneity is the single most important determinant. In the 1996–2000 subperiod, the variance of worker fixed effects makes up 50 percent of the total variance of log earnings. This share increases monotonically to 60 percent by the last subperiod. The variance of firm effects, on the other hand, makes up 23 percent of the variance of log earnings in the 1996–2000 subperiod, decreasing to 15 percent by the last subperiod.

Second, in terms of explaining *time trends*, we observe a more than proportionate fall in the variance of firm effects. Between 1996–2000 and 2008–2012, the variance of firm effects falls from 17 to eight log points whereas the variance of person effects falls from 36 to 31 log points. Additionally, the declining variance of each component is reflected in a lower covariance between the

two, with the correlation between worker and firm effects staying fairly constant at around 0.3 throughout the period.<sup>26</sup> Given the large role played by firms in the decline, understanding the drivers of more equal pay across employers over time is an important question which we address in the following section.

Figure 7. Variance decomposition from AKM model with firm and worker fixed effects



When studying the link between firm effects and firm performance, we limit attention to the manufacturing and mining sector, for which we have data on firm performance and characteristics.<sup>27</sup> Table 6 compares AKM estimates for this subpopulation with the overall population. The overall variance of log earnings is four log points higher in the PIA subpopulation in the 1996–2000 subperiod and falls by 19 log points over the subsequent three subperiods (versus 20 log points in the overall population). The variance of worker effects is three log points higher in the 1996–2000 subperiod and four log points greater in the 2008–2012 subperiod. The variance of firm effects is two log points less in 1996–2000 and falls by eight instead of nine log points. We conclude from this that trends in inequality are similar in the PIA subpopulation as in the overall population.

<sup>26</sup>Given that sorting patterns—measured by the correlation between worker and firm effects—remains fairly constant over time, the true reduction in earnings inequality due to firms would be even greater if the covariance term were distributed proportionately between workers and firms.

<sup>27</sup>As noted earlier, we impose the restriction to the PIA subpopulation after estimating the AKM model on the entire

Table 5. AKM variance decomposition results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1988–1992	1992–1996	1996–2000	2000–2004	2004–2008	2008–2012	Cumulative change from 1988 to 2012	Cumulative change from 1996 to 2012
Variance of log earnings	0.77 (100%)	0.77 (100%)	0.72 (100%)	0.66 (100%)	0.57 (100%)	0.52 (100%)	-0.25 (100%)	-0.20 (100%)
Variance of worker effects	0.39 (51%)	0.38 (49.2%)	0.36 (50%)	0.35 (54%)	0.33 (58%)	0.31 (60%)	-0.09 (34%)	-0.05 (24%)
Variance of firm effects	0.16 (21%)	0.18 (23.1%)	0.17 (23%)	0.13 (20%)	0.10 (17%)	0.08 (15%)	-0.08 (33%)	-0.09 (45%)
Variance of year effects	0.02 (2%)	0.01 (2%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	-0.02 (6%)	0.00 (0%)
2×Cov. worker and firm effects	0.14 (18%)	0.14 (18%)	0.14 (20%)	0.12 (18%)	0.10 (18%)	0.10 (20%)	-0.04 (14%)	-0.04 (22%)
2×Cov. worker and year effects	-0.01 (-2%)	-0.02 (-2%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.01 (-2%)	0.00 (0%)
2×Cov. firm and year effects	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (-1%)	0.00 (-1%)
Variance of residual	0.08 (10%)	0.07 (10%)	0.06 (8%)	0.05 (7%)	0.04 (7%)	0.04 (7%)	-0.04 (17%)	-0.02 (10%)
# worker years	85.4	85.6	90.2	102.0	123.7	151.0		
# firm years	0.98	1.04	1.23	1.44	1.75	2.22		
$R^2$	0.90	0.91	0.92	0.93	0.93	0.93		

Note: Variance decomposition is  $Var(y_{it}) = Var(a_i) + Var(\hat{\gamma}_t) + Var(e_{it}) + 2Cov(a_i, \hat{\gamma}_t) + 2Cov(a_i, \hat{\gamma}_t) + 2Cov(a_i, \hat{\gamma}_t)$ . Cells contain variance explained by each decomposition element. The share of the total variance explained by each decomposition element is given in parentheses.

Table 6. AKM decomposition results for workers at large manufacturing&mining firms (as fraction of largest connected set)

	(1)	(2)	(3)	(4)	(5)
	1996–2000	2000–2004	2004–2008	2008–2012	Cumulative change from 1996 to 2012
Variance of log earnings (% of pop. estimate)	0.76 (100%)	0.74 (100%)	0.65 (100%)	0.58 (100%)	-0.19 (100%)
Variance of worker effects (% of pop. estimate)	0.39 (51%)	0.39 (53%)	0.38 (58%)	0.35 (61%)	-0.04 (20%)
Variance of firm effects (% of pop. estimate)	0.15 (19%)	0.12 (16%)	0.08 (13%)	0.07 (11%)	-0.08 (43%)
Variance of year effects (% of pop. estimate)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)
2×Cov. worker and firm effects (% of pop. estimate)	0.18 (23%)	0.18 (24%)	0.15 (23%)	0.12 (21%)	-0.06 (30%)
2×Cov. worker and year effects (% of pop. estimate)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)
2×Cov. firm and year effects (% of pop. estimate)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (0%)	0.00 (-1%)
Variance of residual (% of pop. estimate)	0.05 (7%)	0.05 (6%)	0.04 (6%)	0.04 (6%)	-0.02 (10%)
# Worker years (% of pop. estimate)	16.60 (18%)	17.90 (18%)	22.80 (18%)	26.30 (17%)	
# Firm years (% of pop. estimate)	0.11 (3%)	0.13 (3%)	0.16 (3%)	0.18 (2%)	
R <sup>2</sup> (% of pop. estimate)	0.93 (101%)	0.94 (101%)	0.94 (101%)	0.94 (101%)	

Note: Variance decomposition of AKM model estimated using manufacturing firms covered by PIA. The ratio between estimates using manufacturing firms relative to AKM estimates using all sectors is given in parentheses. Worker-years, unique workers, firm-years, and unique firms are in millions.

## 6.2 The link between firm effects and firm characteristics

Given that a reduction in dispersion of firm effects has been an important element in the decline, a natural question is whether differences in pay at the firm level are related to observable characteristics of the firm and whether changes in such observable characteristics explain the changes we observe in the firm-specific component of pay over time. Using measures of firm performance from the PIA data, Table 7 reports results from regressing estimated firm effects on log value added per worker both controlling for and not controlling for two-digit subsectors.<sup>28</sup>

Several features are worth highlighting. First, greater value added per worker is associated with a higher firm component of pay. Second, the amount of dispersion in firm effects explained by firm performance is notable, with an  $R^2$  above 0.6 (around 0.5 without subsector controls). Third, the pass-through from value added per worker to the firm component of pay declines substantially over time. In 1996–2000, a one log point increase in value added per worker was associated with a 0.24 log points increase in estimated firm effects (0.26 without subsector controls); in 2008–2012, the same increase in value added per worker was only associated with a 0.13 log point higher firm effect (0.15 without subsector controls).

Table 7. Regression of estimated firm effects on firm productivity

	1996–2000		2000–2004		2004–2008		2008–2012	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Value added p.w.	0.258	0.236	0.201	0.182	0.163	0.143	0.147	0.126
Sector controls	No	Yes	No	Yes	No	Yes	No	Yes
Worker-years	16.6	16.6	17.9	17.9	22.8	22.8	26.3	26.3
$R^2$	0.583	0.655	0.562	0.650	0.537	0.639	0.483	0.624

Note: Dependent variable is the estimated firm effect  $a_j$ . Independent variable is log value added per worker. Number of worker-years in millions.

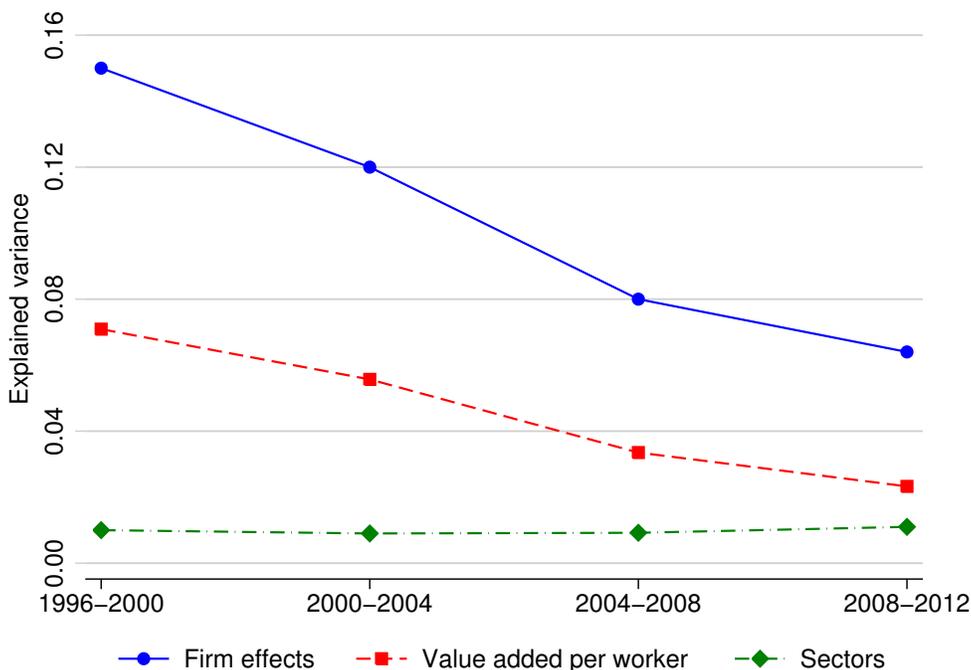
Figure 8 evaluates the implication of the estimation results in Table 7 for the compression in firm pay in Brazil. It plots the overall variance of firm effects (solid blue line with circles), the predicted variance only due to value added per worker (dashed red line with squares), and the predicted variance only due to sectors (dash-dotted green line with diamonds) across periods. Value added per worker explains almost half of the initial variance in firm effects, while sectors initially

population of adult males.

<sup>28</sup>We find that other firm characteristics only marginally improve the explanatory power of the model. Moreover, the variance of the predicted values due to other firm characteristics is essentially zero, and hence we do not report them here.

explain seven percent. Between 1996–2000 and 2008–2012, the variance explained by value added per worker declines by five log points or 25 percent of the overall decline in the variance of log earnings in Brazil over this period. The variance due to sectoral variation, on the other hand, is small and remains constant in levels.<sup>29</sup>

Figure 8. Regression of firm effects on productivity and sector controls



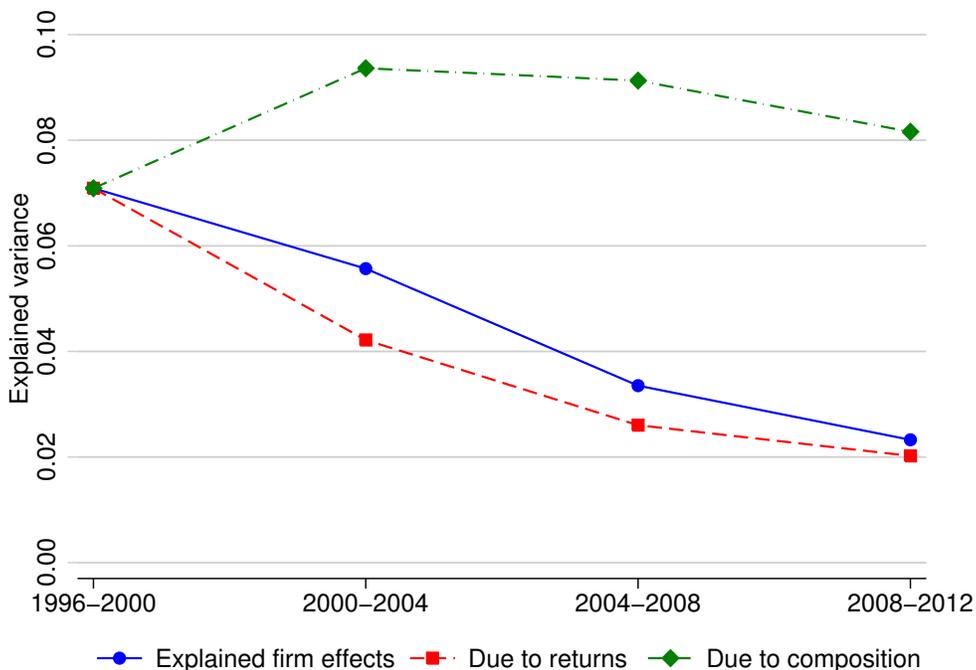
To quantify the importance of changes in the firm productivity distribution versus the pass-through from productivity to pay, we consider the two counterfactual exercises outlined in Section 5. Thus, we first hold the pass-through from value added per worker to the firm component of pay constant at its estimated 1996–2000 value and allow only the variance of value added per worker to change as in the data. Second, we hold the variance of value added per worker fixed and change only the pass-through to match the data.

Figure 9 plots the result from this exercise. The predicted variance of firm effects due to value added per worker (solid blue line with circles) declines from seven to two log points, the predicted variance from value added per worker holding pass-through constant (dashed red line with squares) increases from seven to eight log points, and the predicted variance holding the dis-

<sup>29</sup>We separately investigated the role of sectors in the overall economy, not just the manufacturing and mining sectors, and find a significant role for variation in firm-level pay within sectors; results are available upon request.

person in value added per worker constant (dash-dotted green line with diamonds) falls slightly more than the total predicted variance. We conclude that, *ceteris paribus*, a declining pass-through from firm performance to pay contributed significantly to reduced earnings inequality in Brazil during this period.

Figure 9. Variance of predicted firm effects, variance predicted by fixed pass-through and variance predicted by fixed productivity distribution



### 6.3 The link between worker effects and worker characteristics

We finally turn to our results regarding the link between observable worker characteristics and the worker component of pay. Table 8 shows the result from a regression of the estimated worker effects on age and education with and without occupation controls. We note the following: First, more able workers, as measured by labor market experience or education level, are paid more. For instance, the oldest age group earns on average 50 log point more than the youngest age group conditional on place of work. Secondly, such worker characteristics explain roughly 40 percent of the variance in worker effects, leaving substantial room for unobserved worker heterogeneity. Third, both the age gradient and the return to education declines over time, with a particularly pronounced fall in the latter. For instance, a college degree was associated with a 116 log point

higher worker component of pay in 1996–2000 but only a 78 log point higher worker effect in 2008–2012.

Table 8. Regression of estimated worker effects on worker characteristics

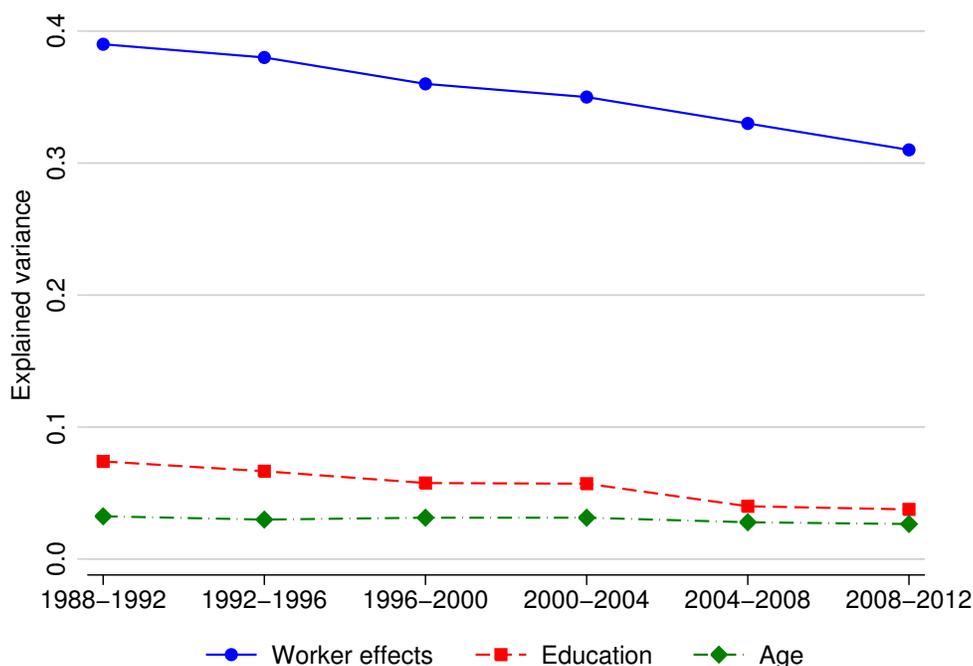
	1988–1992		1992–1996		1996–2000		2000–2004		2004–2008		2008–2012	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Age groups</i>												
25–29	0.20	0.18	0.18	0.18	0.19	0.18	0.18	0.18	0.16	0.15	0.16	0.15
30–39	0.39	0.35	0.35	0.32	0.35	0.32	0.34	0.32	0.31	0.29	0.30	0.28
40–49	0.52	0.47	0.50	0.46	0.51	0.47	0.49	0.45	0.45	0.41	0.42	0.39
50–64	0.48	0.44	0.47	0.44	0.50	0.46	0.52	0.48	0.52	0.47	0.51	0.47
<i>Education groups</i>												
Middle school	0.21	0.16	0.19	0.15	0.16	0.12	0.14	0.10	0.13	0.08	0.11	0.07
High school	0.61	0.51	0.56	0.46	0.45	0.37	0.37	0.29	0.32	0.21	0.27	0.19
College or more	1.21	1.10	1.19	1.04	1.16	1.00	1.19	1.00	1.15	0.82	1.10	0.78
Occupation controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
# Workers	85.4	85.4	85.6	85.6	90.2	90.2	102.0	102.0	123.7	123.7	151.0	151.0
$R^2$	0.35	0.40	0.34	0.37	0.34	0.38	0.35	0.39	0.36	0.41	0.37	0.44

Note: Dependent variable is the estimated worker effect  $a_i$ . Number of workers in millions. Education estimates are relative to “less than middle school (<7 years)” category. Age estimates are relative to “age 18–24” category.

To quantify the contribution of the estimation results in Table 8 towards the declining dispersion in worker heterogeneity in Brazil, Figure 10 plots the overall variance of worker effects (solid blue line with circles), the predicted variance only due to age structure (dash-dotted green line with diamonds), and the predicted variance only due to variation in educational attainment (dashed red line with squares). We do not plot the variance due to occupation as this remains small and steady at around one, two log points over the entire period (in particular, it increases slightly from 0.019 in 1988–1992 to 0.023 in 2008–2012).<sup>30</sup> We find that while education and age explain only around 20–25 percent of total worker heterogeneity, they account for almost half of the decline in the variance of worker effects.

<sup>30</sup> Also not shown, the sum of covariances increases modestly from 0.032 in 1988–1992 to 0.040 in 2008–2012.

Figure 10. Regression of worker effects on age and education



We further decompose the decline in the dispersion due to age and education into one due to changes in returns and one due to changes in the composition of the workforce.<sup>31</sup> Figure 11 plots in the top (bottom) panel in solid blue with circles the overall compression explained by education (age), in dash-dotted green with diamonds the compression resulting from holding the returns to education (age) fixed, and in dashed red with squares that when holding the underlying education (age) distribution fixed.

Holding returns fixed, changes over this period in the age structure as well as the educational attainment of the workforce resulted in a small net increase in inequality. Given the large increase in educational attainment documented earlier, it might come as a surprise that this did not induce a larger change in the underlying distribution of educational attainment. This is the result of two offsetting forces: on the one hand, the share of the lowest education group shrinks, but on the other more workers obtain high-school or college degrees, which earn more than the average. Altogether, these two effects have small positive net effects on inequality.

At the same time the rapid decline in the return to education—and to a lesser extent age—

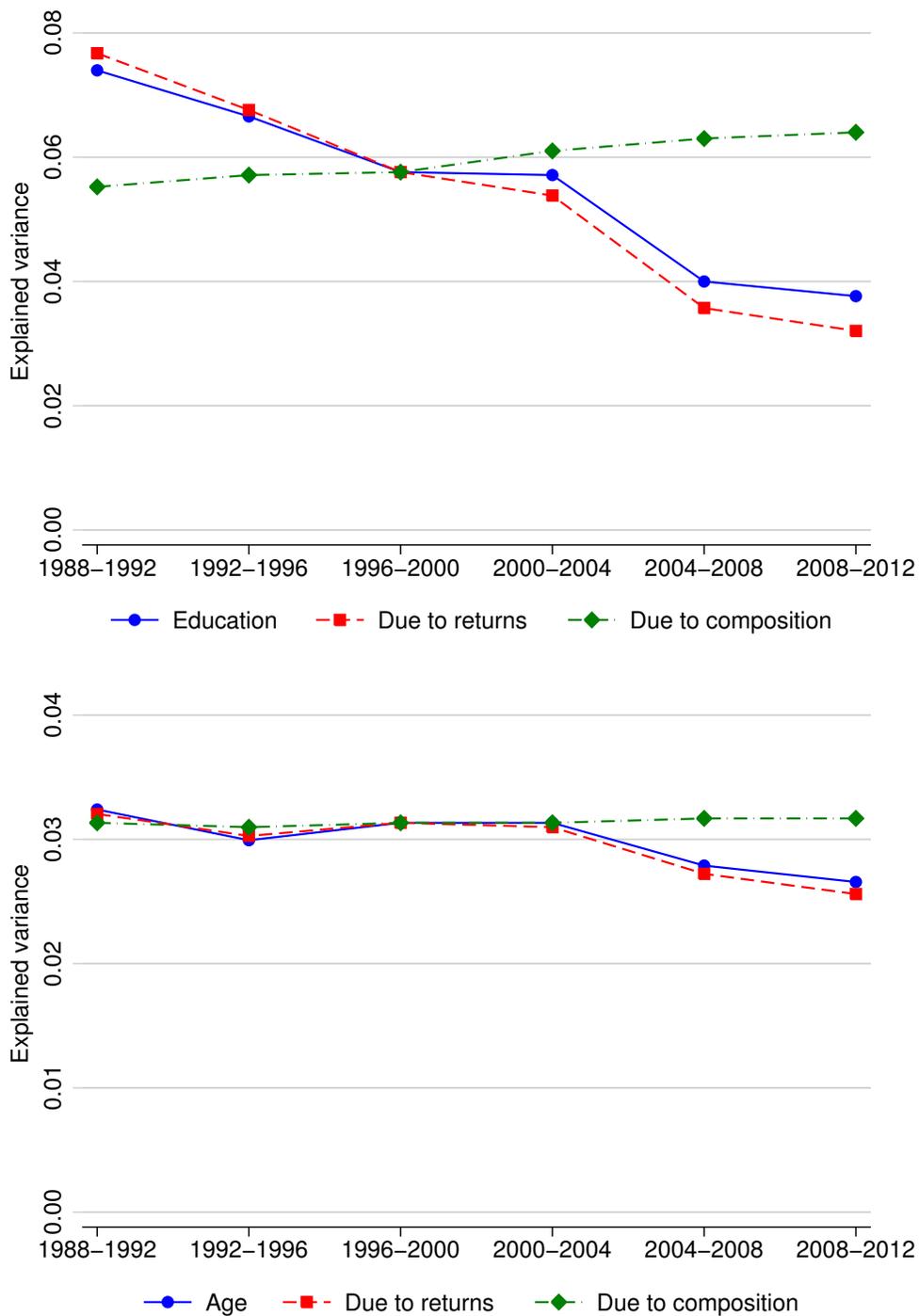
<sup>31</sup>Since the variance due to occupation is small and does not change much over time we do not present a decomposition of this.

resulted in significant earnings compression. In fact as can be seen by comparing the solid blue lines with circles to the dashed red line with squares, the compression in returns resulted in an even larger fall in inequality than the overall fall. Holding everything else constant, the fall in the return to education (age) explains 13 (three) percent of the overall decline in the variance of log earnings between 1996–2000 and 2008–2012.<sup>32</sup>

---

<sup>32</sup>Of course, the compression in the return to education could have partly resulted from changes in educational attainment. The counterfactual of holding the distribution constant while changing returns should be interpreted with this qualification in mind.

Figure 11. Decomposition of variance of worker effects predicted by age (top) and education (bottom)



## 6.4 Empirical support for the AKM model

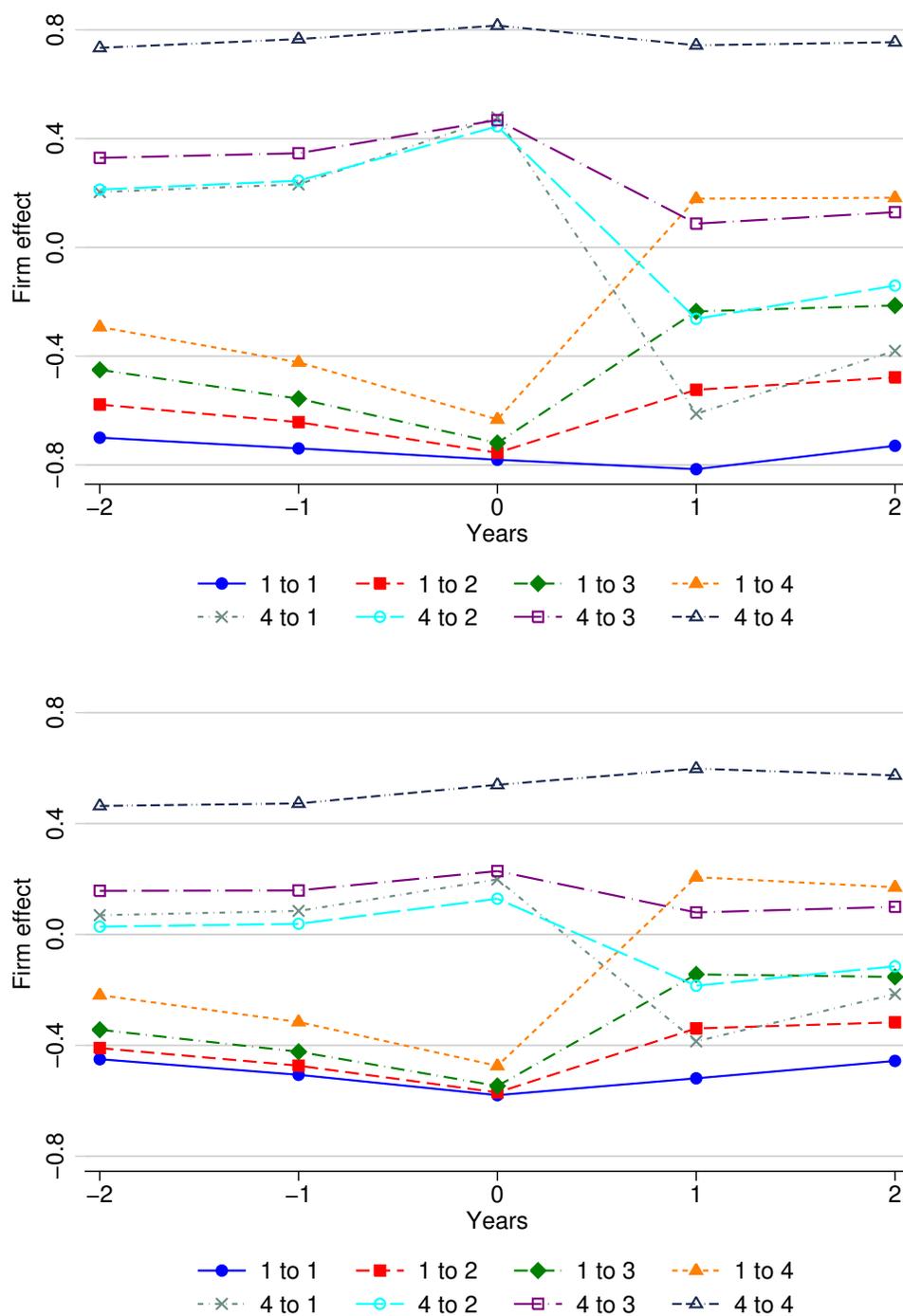
A large part of our analysis has been based on estimation results from repeated applications of the AKM framework. Although the consistently high  $R^2$  coefficient of our model suggests that the model provides a good fit to the data, our estimates can be biased if the residual is correlated with either the firm or worker component of earnings. To investigate this possibility further, we replicate using our Brazilian data two diagnostics exercises proposed by [Card, Heining and Kline \(2013\)](#) for the case of Germany.

Figure 12 shows the average firm effect of workers who switch firms up to two years prior to the switch and two years after the switch for the first and last period of our sample. Switchers are classified by the firm effect quartile of the pre and post transition firms. Consistent with the AKM specification, workers that switch from the lowest quartile experience gains in firm effect and workers that switch from the highest quartile experience losses. Additionally, the gains of those switching up are similar to the losses of those making the reverse switch.<sup>33</sup> This suggests that our additive model is consistent with the pattern observed among workers who transition between firms.

---

<sup>33</sup>For expositional ease we only show switches out of the first and fourth quartile, but other quartiles display similar pattern and are available on request.

Figure 12. Average changes in firm effects of workers switching employers, 1996–2000 (top) and 2008–2012 (bottom)



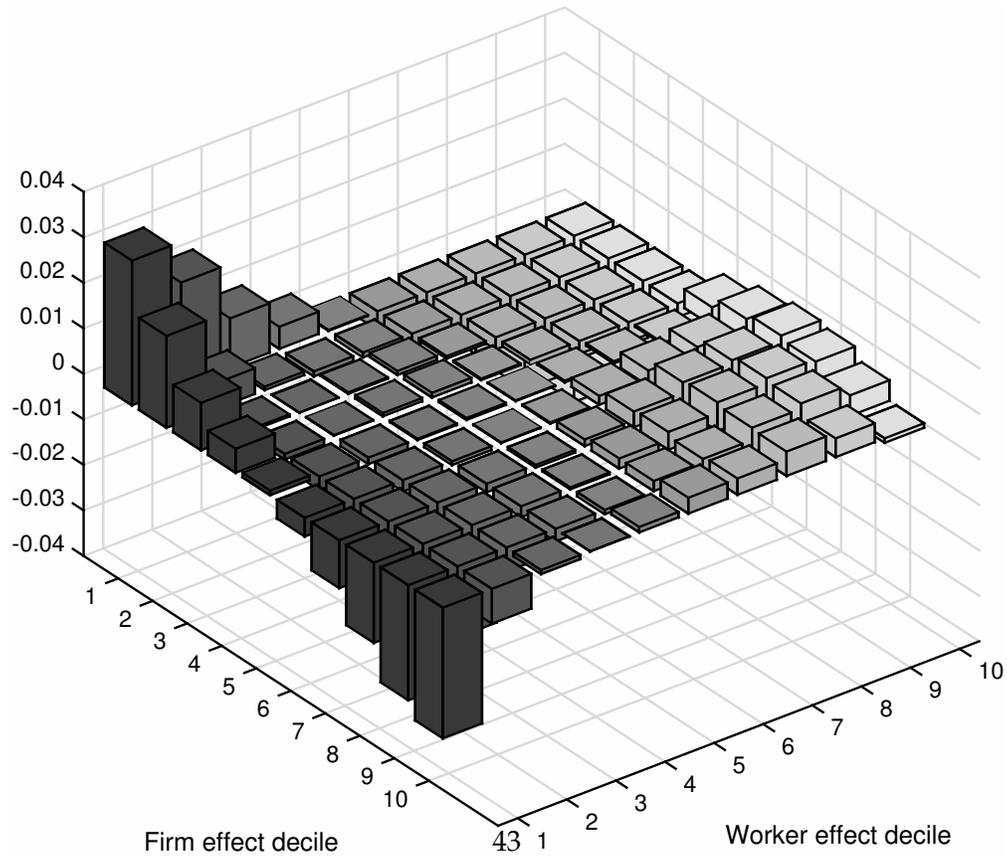
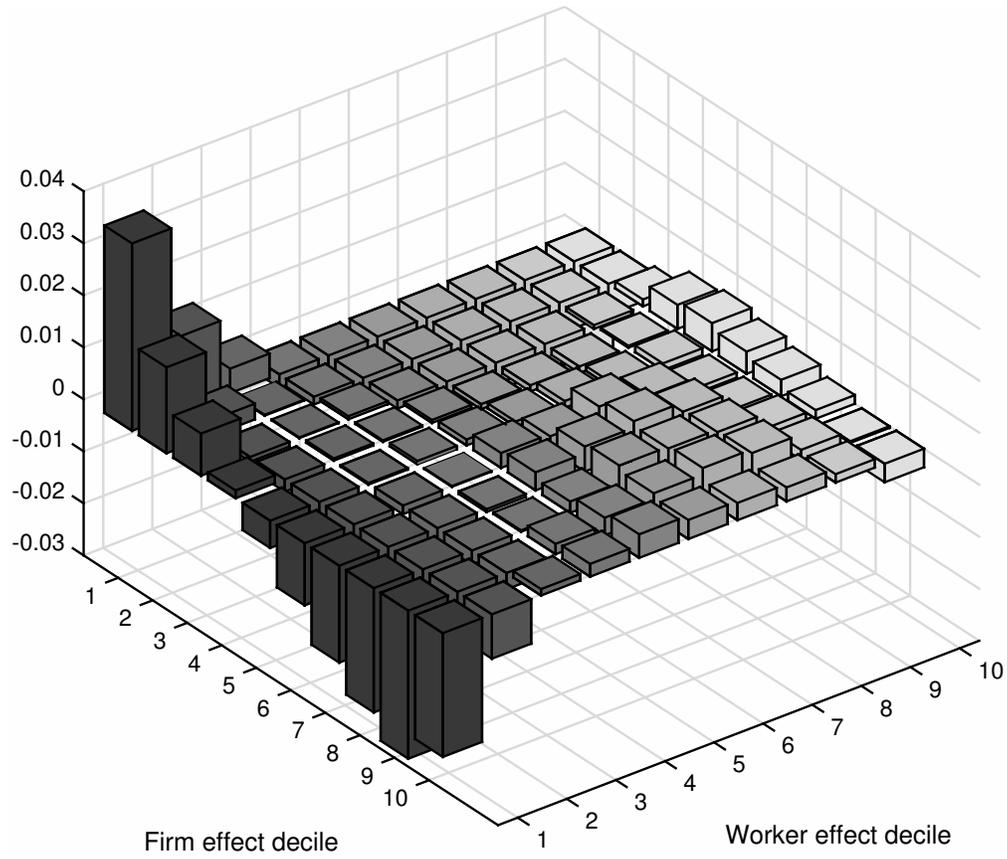
Second, we find little evidence in the data for match effects that are systematically correlated

with either person or match effects. Figure 13 shows the average estimated residual by decile of worker and firm effect. There is some evidence of misspecification for the lowest decile of workers in the sense that they display a systematically positive residual while working at the lowest paying firms. However, the magnitude of the error is modest and beyond the lowest decile of worker effects, errors do not exhibit any systematic relationship with firm and worker effects.<sup>34</sup> This reassures us that the log additive assumption is an appropriate characterization of Brazilian wage setting policies.

---

<sup>34</sup>The systematic variation among the lowest decile of worker effects is due to those individuals having a strongly positive error term when working at the lowest-paying firms, and strongly negative error term when working at the highest-paying firms. In other words, for lowest worker effects individuals the pay gradient across firms is very flat. We argue in Engbom and Moser (2015) that this pattern is consistent with a frictional labor market with a binding minimum wage.

Figure 13. AKM residual by firm and worker fixed effect deciles, 1988–1992 (top) and 2008–2012 (bottom)



## 7 Conclusion

In this paper, we document and dissect a large decline in earnings inequality in Brazil between 1996 and 2012. To understand the sources of this decline, we estimate a high-dimensional fixed effects model of earnings controlling for both worker and firm heterogeneity. Our results suggest that firms contributed more than proportionately to the decline. While more productive firms pay more, compression in firm productivity was not a factor behind declining inequality. Instead, we find that a declining pass-through from firm productivity to pay played an important role behind the decline. On the worker side, higher ability workers as measured by educational attainment or experience are paid more, yet also such measures of worker ability show no compression over time. Rather, a significant share of the explained decline in worker heterogeneity is due to rapidly falling returns to measures of worker ability. We conclude that changes in pay policies, rather than changes in firm and worker fundamentals, played a significant role in Brazil's inequality decline.

Our paper suggests a set of facts that a potential theory of the inequality decline in Brazil should be consistent with. Such a theory must generate pay differences between firms for identical workers, and they have to be strongly positively correlated with firm productivity. Moreover, any promising explanation needs to generate a compression in firm pay differences over time, not through compression in firm productivity but through a weaker link between productivity and pay. On the worker side, it should generate compression primarily through declining returns to ability, and not through a compression in underlying ability.

We think that promising candidates behind the decline in inequality in Brazil are changes in the nature of wage setting. In follow-up work [Engbom and Moser \(2015\)](#), we argue that the rapid rise in the minimum wage in Brazil during this period can explain a significant fraction of the decline in inequality, while being consistent with the stylized facts presented in this paper. Other institutional reforms, including changes in union bargaining structure and labor compensation laws, represent interesting avenues for future research.

## References

**Abowd, John M, Francis Kramarz, and David N Margolis.** 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251–333.

- Abowd, John M., Robert H. Creecy, and Francis Kramarz.** 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data."
- Atkinson, Anthony B., and François Bourguignon,** ed. 2015. *Handbook of Income Distribution*. . Vol. 2 ed., Elsevier.
- Bagger, Jesper, Bent Jesper, and Royal Holloway.** 2014. "Productivity and Wage Dispersion: Heterogeneity or Misallocation." 1–40.
- Bagger, Jesper, Kenneth L Sorensen, and Rune Vejlin.** 2013. "Wage sorting trends." *Economics Letters*, 118(1): 63–67.
- Barros, Ricardo, Mrela de Carvalho, Samuel Franco, and Rosane Mendonça.** 2010. "Markets, the State and the Dynamics of Inequality: The Case of Brazil." *UNDP Discussion Paper*.
- Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta.** 2009. "Measuring and Analyzing Cross-country Differences in Firm Dynamics." In *NBER Chapters*. 15–76. National Bureau of Economic Research, Inc.
- Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta.** 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review*, 103(1): 305–334.
- Barth, Erling, James C Davis, and Alex Bryson.** 2014. "It's Where You Work: Increases in Earnings Dispersion across Establishments and Individuals in the U.S." *NBER Working Paper 20447*.
- Bloom, Nicholas, Fatih Guvenen, David Price, and Jae Song.** 2015. "Evolution of Inequality within Firms (SED Abstract)."
- Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa.** 2015. "A Distributional Framework for Matched Employer Employee Data."
- Card, David, Ana Rute Cardoso, and Patrick Kline.** 2015. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women."
- Card, David, Joerg Heining, and Patrick Kline.** 2013. "Workplace heterogeneity and the rise of West German wage inequality." *Quarterly Journal of Economics*, 128: 967–1015.
- Córdova, José Ernesto López, and Mauricio Mesquita Moreira.** 2003. "Regional Integration and Productivity: The Experiences of Brazil and Mexico." *IDB Publications (Working Papers)*.
- de Araujo, Ana Luisa Pessoa.** 2014. "Wage Inequality and Job Stability." 1–53.
- Deaton, Angus.** 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*.
- Dix-Carneiro, Rafael, and Brian K. Kovak.** 2015. "Trade Reform and Regional Dynamics: Evidence From 25 Years of Brazilian Matched Employer-Employee."
- Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth R. Troske.** 2004. "Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment." *Journal of Labor Economics*, 22(2): 397–430.
- Engbom, Niklas, and Christian Moser.** 2015. "Earnings Inequality and the Minimum Wage: Evidence from Brazil."

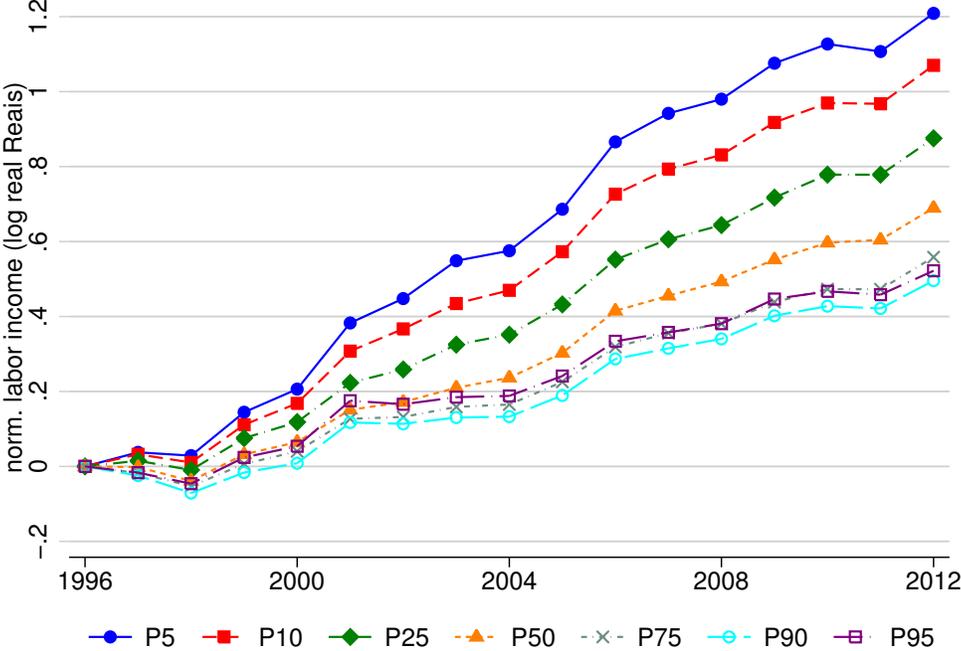
- Faggio, G., K. G. Salvanes, and J. Van Reenen.** 2010. "The evolution of inequality in productivity and wages: panel data evidence." *Industrial and Corporate Change*, 19(6): 1919–1951.
- Ferreira, Francisco H. G., Phillippe G. Leite, and Matthew Wai-Poi.** 2007. *Trade Liberalization, Employment Flows, And Wage Inequality In Brazil. Policy Research Working Papers*, The World Bank.
- Ferreira, Pedro Cavalcanti, and Jose Luiz Rossi.** 2003. "New Evidence from Brazil on Trade Liberalization and Productivity Growth." *International Economic Review*, 44(4): 1383–1405.
- Gruetter, Max, and Rafael Lalive.** 2009. "The importance of firms in wage determination." *Labour Economics*, 16(2): 149–160.
- Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen Redding.** 2013. "Trade and Inequality: From Theory to Estimation."
- Kopczuk, Wojciech, Emmanuel Saez, and Song Jae.** 2010. "Earnings inequality and mobility in the United States: evidence from social security data since 1937." *Quarterly Journal of Economics*, 125(1): 91–128.
- Lopes De Melo, Rafael.** 2013. "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence." 1–39.
- Lopez, J. Humberto, and Guillermo Perry.** 2008. "Inequality in Latin America: Determinants and Consequences."
- Medeiros, Marcelo, Pedro H G Ferreira De Souza, and Fábio Avila de Castro.** 2014. "O topo da distribuição de renda no Brasil: primeiras estimativas com dados tributários e comparação com pesquisas domiciliares , 2006 - 2012." 2006–2012.
- Menezes-Filho, Naércio Aquino, Marc-Andreas Muendler, and Garey Ramey.** 2008. "The Structure of Worker Compensation in Brazil, with a Comparison to France and the United States." *Review of Economics and Statistics*, 90(2): 324–346.
- Moreira, Mauricio Mesquita.** 2004. "Brazil's Trade Liberalization and Growth: Has it Failed?" *International Trade*.
- Mueller, Holger M, and Paige P Ouimet.** 2015. "Wage Inequality and Firm Growth."
- Muendler, Marc-Andreas.** 2004. "Trade, Technology, and Productivity: A Study of Brazilian Manufacturers, 1986-1998." *CESifo Working Paper Series*.
- Pavcnik, Nina, Andreas Blom, Pinelopi Goldberg, and Norbert Schady.** 2004. "Trade Liberalization and Industry Wage Structure: Evidence from Brazil." *World Bank Economic Review*, 18(3): 319–344.
- Tsounta, Evridiki, and Anayochukwu I. Osueke.** 2014. "What is behind Latin America's Declining Income Inequality?"
- Ulyssea, Gabriel.** 2014. "Firms, Informality and Development: Theory and evidence from Brazil."
- World Bank.** 2015. "World Development Indicators."

# Appendix

## A Additional figures

### A.1 Evolution of real earnings levels

Figure 14. Earnings levels evolution



## A.2 Cross-section and time series comparison of alternative productivity measures

Figure 15. Cross-sectional comparison of alternative productivity measures

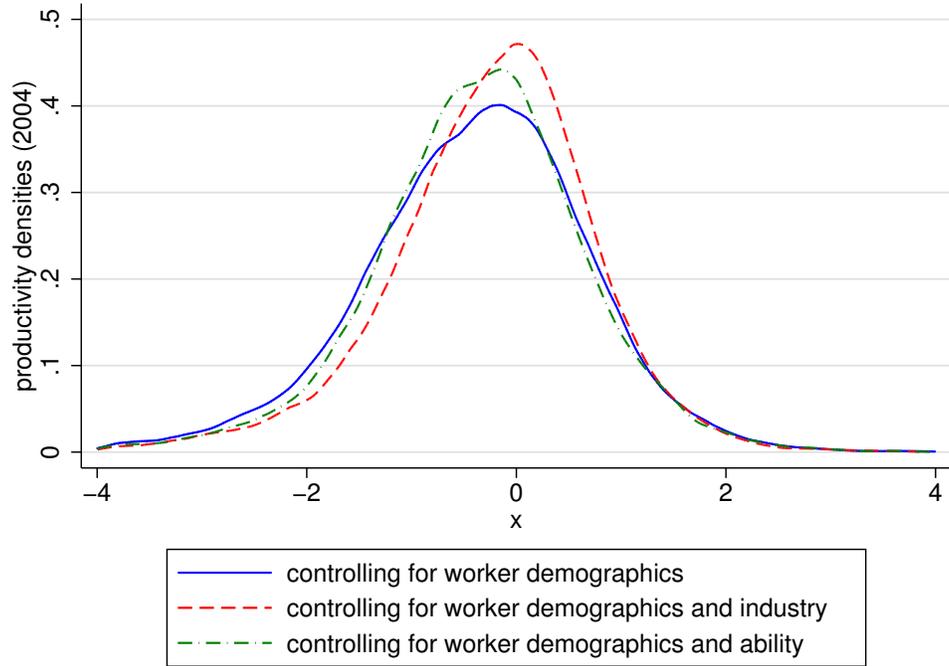
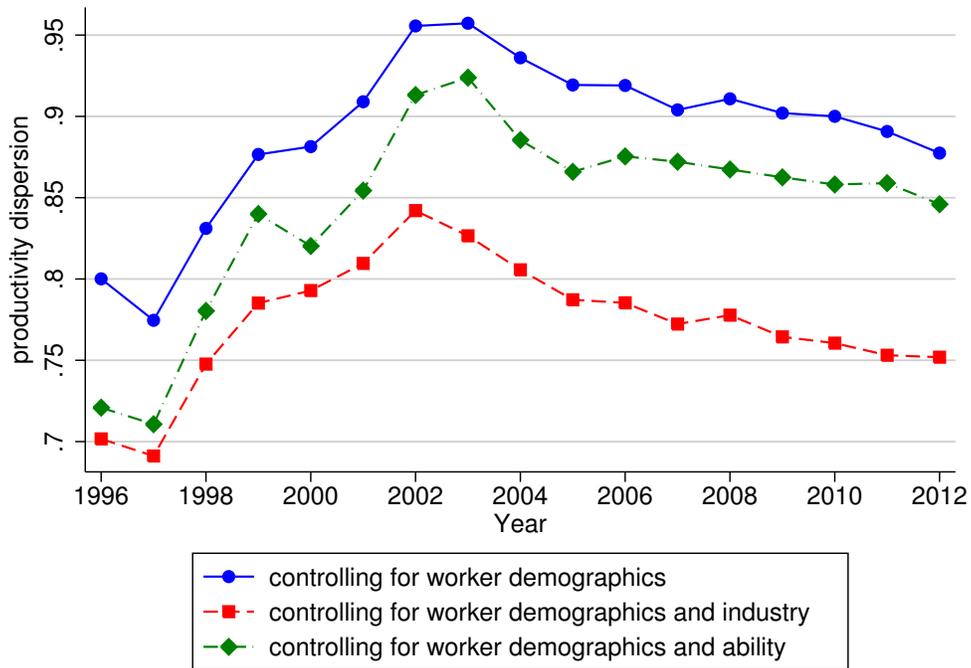


Figure 16. Evolution of dispersion of alternative productivity measures



### A.3 Earnings inequality by education groups

Figure 17. Within education group inequality

