Firms and the Decline of Earnings Inequality in Brazil*

Jorge Alvarez† Niklas Engbom‡ Christian Moser§

December 2, 2015

Abstract

Brazil experienced a large decline in earnings inequality between 1996 and 2012, with the variance of log earnings falling by 26 log points. Using administrative linked employer-employee data, we fit models with log-additive worker and firm fixed effects within overlapping subperiods in order to identify the sources of this decline. We find that compression in firm fixed effects accounts for 45 percent of the decline in the variance of log earnings over the period and compression in worker fixed effects accounts for 28 percent, with a fall in their covariance and the residual explaining the remainder. The drop in firm pay differences is not driven by convergence in firm productivity. Instead, a significant fraction of the decline is due to a weaker productivity-pay gradient across firms. Our results suggest that changes in pay policies, rather than changes in firm fundamentals, played a significant role in Brazil's inequality decline.

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

JEL classification: D22, E24, 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, Nobu Kiyotaki, Alan Krueger, Ilyana Kuziemko, Alex Mas, Ben Moll, Ezra Oberfield, Stephen Redding, 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, and the Princeton Public Finance Working Group. Special thanks go to Carlos Lessa and Luis Pinto at IBGE as well as Carlos Corseuil, Glaucia 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.

†Department of Economics, Princeton University, 001 Fisher Hall, Princeton, NJ 08544, USA. Web: http://scholar.princeton.edu/jalvarez. Email: jaalvare@princeton.edu

‡Department of Economics, Princeton University.

§Department of Economics, Princeton University, 001 Fisher Hall, Princeton, NJ 08544, USA. Web: http://www.economoser.com. Email: cmoser@princeton.edu

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.[1] While the decline of earnings inequality in Brazil resembles the experience of other Latin American economies during this period, it stands in stark contrast to that of the United States, which like several other developed countries saw inequality steadily increasing at the same time.[2] In this paper, we study the sources of Brazil's decline in earnings inequality.

Guided by recent research, which suggests that firms are an important determinant of earnings dispersion, 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 separate the contributions of firm- and worker-specific factors towards the overall fall in inequality. Subsequently, we investigate the link between firm performance and the firm component of pay by using another confidential dataset containing information on firm characteristics of hundreds of thousands of firms between 1996 and 2012.

We uncover two main results. First, firms played an important role in the decline in earnings inequality in Brazil over this period, explaining about 45 percent of the fall in the variance of log earnings between 1996 and 2012. Compression in worker fixed effects explains an additional 28 percent of the decline, with the remaining part due to a decline in the covariance between worker and firm fixed effects and the residual. As worker heterogeneity accounts for 48–56 percent of the level of inequality whereas firm effects account for only 15–23 percent, the compression in the firm-specific pay component contributed more than proportionately to Brazil's inequality decline.

Second, changes in the link between firm performance and pay accounts for a significant fraction of the compression in the firm component of pay. We first show that a significant share of the variation in the firm component of pay can be explained by firm productivity differences, with more productive firms paying more. Subsequently, we show that the dispersion in firm productivity did not decline over this period. Rather, we identify a weakening pass-through from

[1] See Lopez and Perry (2008) and Tsounta and Osueke (2014).

[2] Using administrative data, Kopczuk et al. (2010) document in detail earnings inequality trends for the U.S., while chapter 8 of Atkinson and Bourguignon, eds (2015) discusses inequality trends in middle- and high-income countries.

productivity to pay as an important driver of Brazil's inequality decline, accounting for 25 percent of the overall decline in earnings inequality.

Our findings suggest that changes in pay policies, rather than changes in firm fundamentals, played an important role in Brazil's inequality decline. These findings shed new light on a lively debate around the drivers earnings inequality in many developed countries. For example, **?** and Faggio et al. (2010) argue that a significant share of the increase in earnings inequality in the U.S. can be explained by widening dispersion of the firm productivity distribution. By showing that over the period from 1996–2012 the underlying firm productivity distribution remained constant while the link between firm productivity and worker pay weakened in Brazil, we highlight the importance of a complementary determinant of inequality dynamics.

Related literature. Our paper is closely related to three broad strands of the literature. The first studies the role of specific 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 et al. (2014) use administrative tax return data to study the evolution of top income inequality in Brazil 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) decomposes the cross-sectional inequality levels in Brazil into components due to firms and workers. Helpman et al. (2013) use the same dataset 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. We add to this literature by studying changes on the worker and firm side towards this decline in a joint framework.

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 et al., 2002), Denmark (Bagger et al., 2013), Austria (Gruetter and Lalive, 2009), and Germany (Card et al., 2013). The last paper is closest 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. In this paper, we go one step further than previous decomposition exercises by linking a dynamic set of AKM decomposition results to changes in firm productivity and other firm characteristics.

Third, a growing literature studies the link between firm characteristics and worker outcomes. Menezes-Filho et al. (2008) investigate 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 et al. (2014) investigate the role of labor misallocation in driving the positive correlation between labor productivity and wages at the firm using Danish data. Barth et al. (2014) obtain a similar conclusion about the importance of firms using matched employer-employee data for a select number of U.S. states, but not controlling for sorting of workers to firms. Card et al. (2015) study the degree of rent-sharing in Portugal with a particularly emphasis on gender difference in profit participation and the allocation of workers across firms. Our contribution to this literature is to examine an economy in which firms were an important driver of a *decline* in earnings inequality, and to investigate the drivers behind these trends.

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

As part of a wave of democratization in Latin America, Brazil transitioned from military to civilian rule in 1985 and held its first democratic election in almost three decades in 1989. During the following two and a half decades, Brazil cycled through six elected presidents from four political parties. Simultaneously, the country experienced a sustained period of economic growth—between 1996 and 2012 real gross domestic product grew by on average 2.3 percent per year. In this section, we discuss some of the institutional changes that could have have affected inequality, including monetary policy reform, trade liberalization and social policy.

From 1980–1989, yearly inflation averaged 355 percent, which was followed by a yearly average of 1,667 percent between 1990 and 1994 (World Bank, 2015). Several monetary stabilization plans implemented during this period failed.[3] 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. By the early 2000s, exchange rates and inflation had stabilized.

Brazil's trade liberalization over the last 25 years has been frequently cited as a major contributor to the country's growth in total factor productivity (TFP) by opening up the economy to foreign investment (Ferreira and Rossi, 2003; Ferreira et al., 2007; Moreira, 2004; Muendler, 2004; Córdova and Moreira, 2003). Starting with initially high import tariffs that had substituted import bans from the 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). In addition to

[3] Garcia et al. (2014) provides a comprehensive overview of the nine stabilization plans, 15 wage policies, 19 changes to the exchange regime, 22 proposals for the renegotiation of the foreign debt and 20 fiscal adjustment programs that Brazil implemented during this period.

its potential effect on TFP growth, Helpman et al. (2013) argue that the opening up to trade 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.[4] 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 up to 70 percent 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.

Apart from the rapid increase in the minimum wage, several other important changes in labor regulation took place during this period. Before reforms started in the late 1980s, Brazil had a highly regulated labor market. 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. The latter gather about a quarter of employed formal workers in Brazil.[5]

[4]Using household data, Barros et al. (2010) estimate that social programs accounted for about 20 percent of the decline in household income inequality.

[5]Bioto and Marcelino (2011) argue that there has been an uptake in labor strike activity in Brazil since the year 2000.

Carvalho Filho and Estevão (2012) find evidence that these reforms shielded wage-setting conditions from firm performance and reduced wage flexibility. Lagging other liberalization reforms, it was not until the abolition of the Wage Adjustment Law in 1995 that a period of greater flexibility and less regulated wage-setting practices started. 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.[6]

With this brief overview of recent developments in Brazil, we turn to a discussion of the data 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.[7] Fines are levied on late, incomplete, or inaccurate reports, and as a result many businesses hire a specialized ac-

[6]Labor force as a percentage of total population aged 15–64, from OECD Employment and Labor Market Statistics.

[7]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 18–49 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).

countant 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.[8]

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 substantial variation in pay across establishments within firms, our firm-level analysis provides a lower bound on the importance of workers' place of employment.[9] 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. [10] 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

[8]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`.

[9]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.

[10] 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.

differences, we focus on prime age males.[11]

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.[12] Similarly, 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 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 two-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 two digit 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

[11] In the years for which we have data on hours, we find relatively little variation in hours, with most prime age males reporting 44 hours of work a week.

[12] 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.

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 male workers of age 18–49. We make this restriction partly to provide results comparable to a large part of the literature, which tends to focus on prime age males, partly to avoid issues related to intensive margin labor supply, since we lack a complete measure of hours worked.

Descriptive statistics. Table 1 provides key summary statistics for the RAIS data for six subperiods 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 prime age males and prime age 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 prime age males. Prime age 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, prime age 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 prime age males working at large mining and manufacturing firms. Prime age males in the PIA subpopulation are about 0.8 years younger than all prime age males in the 1996–2000 subperiod, which gradually increases to 1.3 years younger in the last subperiod. They are similar to all prime age males in terms of education. The PIA sample of prime age males earned on average 27 log points more than all prime age 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.

Table 1. RAIS summary statistics

	(1)	(2)	Earnings (3)	(4)	Age (5)	(6)	Schooling (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 49 years old. Panel C includes male workers age 18–49 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.

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 above a revenue threshold as well as for an annual random sample of smaller firms.[13] 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 characteris-

[13] The revenue threshold for inclusion in the deterministic survey has grown over the years, standing at USD300,000 in 2012.

tics 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 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 rerevenues as the gross sales value due to operational and non-operational firm activities net of any returns, cancellations, and corrected for changes in inventory.[14]. 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 et al., 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.[15] 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.[16]

[14]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.

[15]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.

[16]In addition, several bargaining models of the labor market have in common that workers and capital owners split

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 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

			Firm size		Value added	
	(1)	(2)	(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 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.

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.

4 Inequality trends in Brazil from 1988–2012

In the following section, we first demonstrate that the decline in inequality in Brazil was broad-based in the sense that it affected 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. Although we document significant changes in educational attainment and education premia in Brazil over the last two decades, we find that such changes cannot explain a majority of the decline in inequality observed in Brazil over this period. 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 Compression in different parts of the earnings distribution

Figure 1 plots log percentile ratios of earnings, from which two facts emerge. 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.

4.2 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 2, 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 2 . In particular, the premia associated with a middle school and high school degree relative to the lowest education group fell rapidly over the past 20 years.

Figure 1. Log percentile ratios of the earnings distribution in Brazil

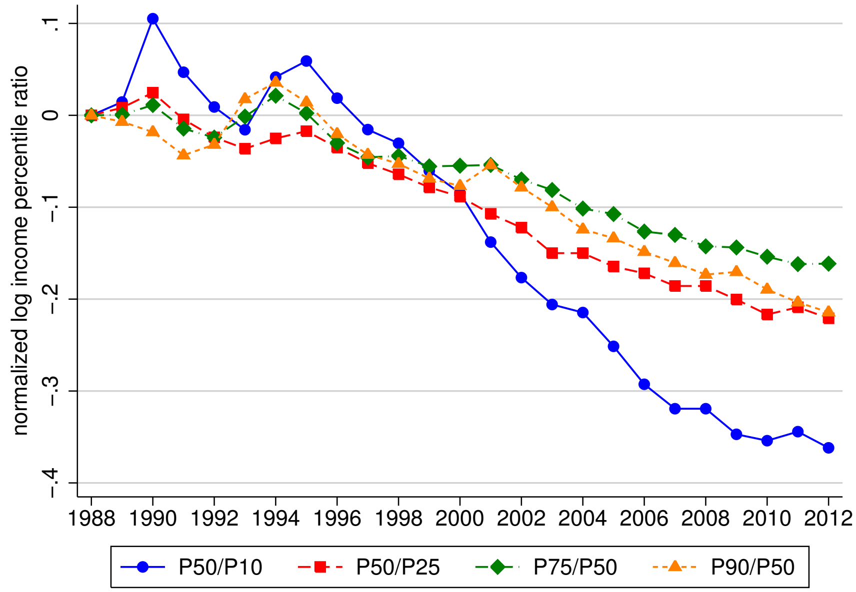
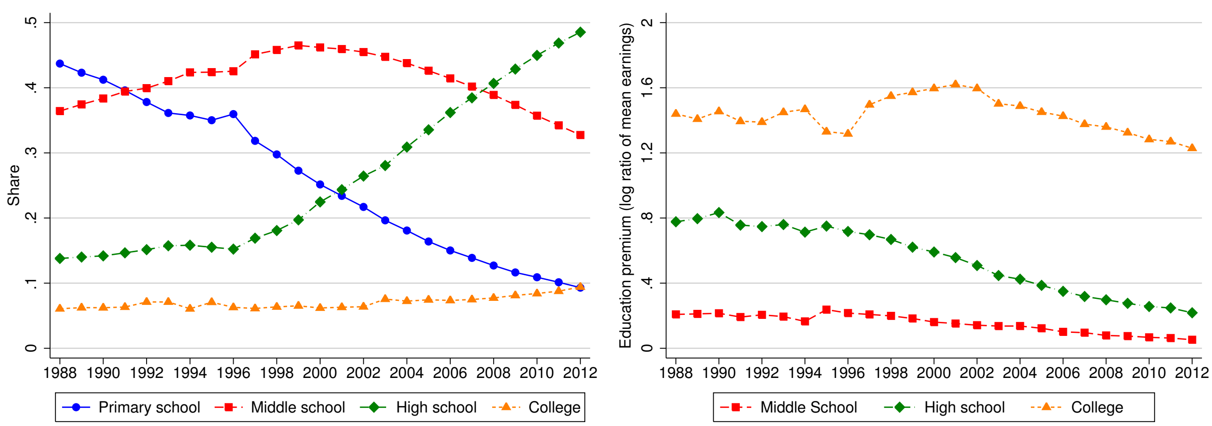


Figure 2. Educational attainment and education premia

(a) Education group shares

(b) Education premia relative to primary school



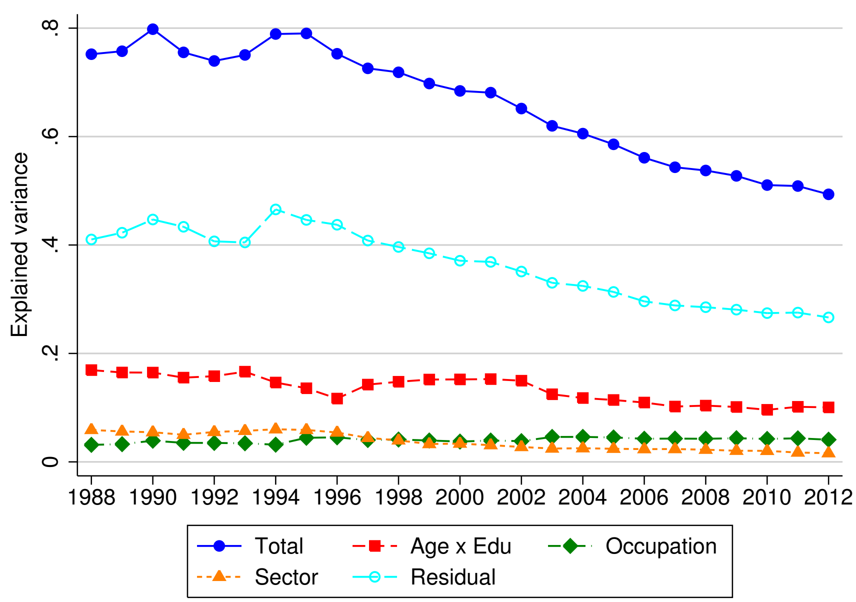
To provide a first look at whether changes in observable worker characteristics were an important driver of earnings inequality in Brazil over this period, we run a series of Mincer regressions. In particular, we regress log earnings of individual i in year t on age dummies interacted with four education dummies, two digit occupation dummies and two digit sector dummies,

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

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 3 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 (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 3. Variance decomposition from Mincer regressions, by year

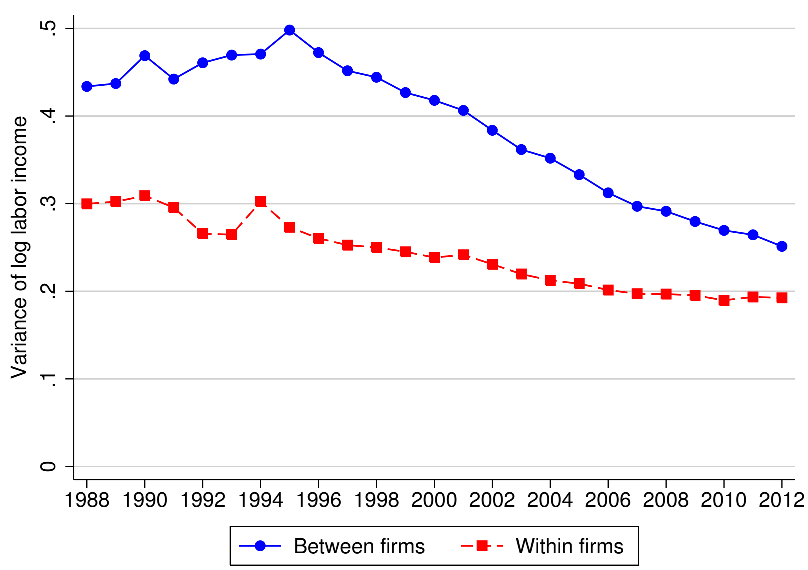


4.3 Earnings dispersion within and across firms

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 microcosmos 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.

Figure 4 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 4. Variance of log earnings within and between firms



Although informative, however, this decomposition cannot necessarily 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.

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.[17]

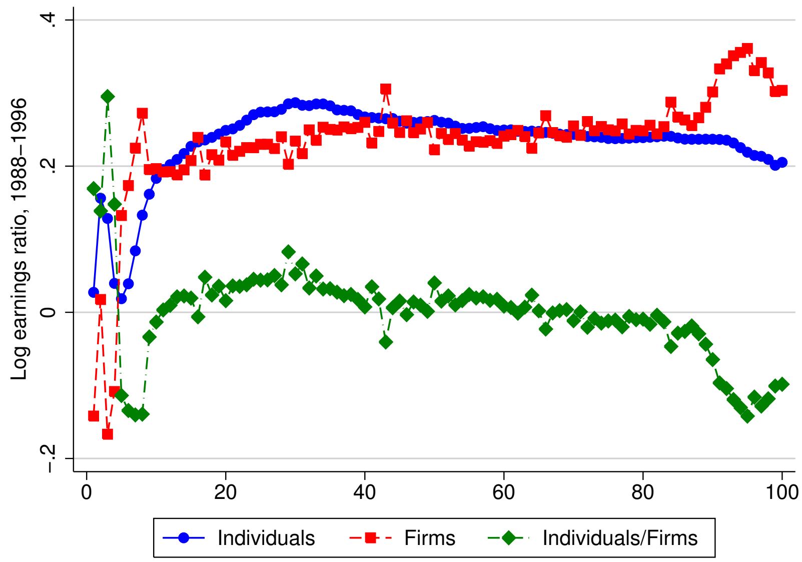
The results of this exercise are shown in Figure 5, first for the period 1988–1996, when earnings inequality remained roughly constant, and then for 1996–2012, when earnings inequality declined rapidly. Average earnings growth (solid blue line with circles) was relatively evenly distributed throughout 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).

Together, the above results suggest that in order to understand Brazil's inequality decline, we need to look beyond the standard wage determinants in the Mincerian tradition. Rather, we provided suggestive evidence that changes in pay across firms could be an important factor behind the overall fall in inequality during this time in Brazil. The next section formalizes our approach to identifying the importance of firm pay policies for earnings inequality.

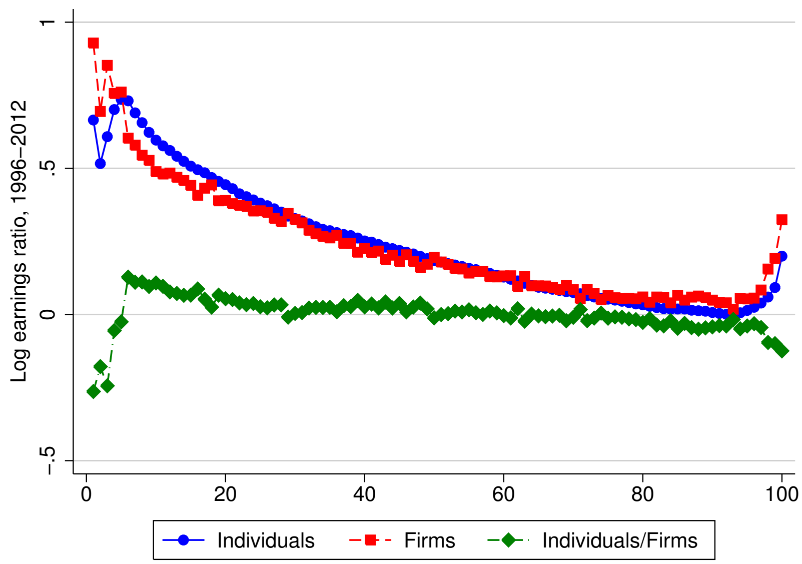
[17]Similar calculations are presented for the U.S. in Barth et al. (2014) and Bloom et al. (2015).

Figure 5. Labor earnings growth of individuals, between and within firms

(a) 1988–1996



(b) 1996–2012



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 et al., 1999). 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 effects with observed characteristics of firms in order to investigate what led to changes in the firm component of pay over time.[18]

5.1 The AKM framework

In order to identify the two-way fixed effects framework of Abowd et al. (1999), one needs to observe a panel of workers with the ability to link multiple workers to the same firm. Our data satisfy these requirments. 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 earnings of individual i in year t, y_{it}, in logs can be written as the sum of a worker effect, α_i, a firm effect, $\alpha_{J(i,t)}$, a time trend, Y_t, and an error, ε_{it}. Although our current paper does not model the underlying, fundamental sources of this specification, 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 thus 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 so as not to identify 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. 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. Thus, also for age we prefer to correlate it 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

[18]In ongoing work we also investigate what drives changes in the worker component of pay.

thus estimate

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

where

$$\mathbb{E}\left[\varepsilon_{it} | i, t, J(i,t), \right] = 0$$

We discuss in greater detail below the assumption on the error term.

As shown by Abowd et al. (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 prime age 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 inoccuous.

Table 3. Summary statistics on workers in largest connected set, relative to prime-age males

			Earnings		Age		Schooling	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	# Worker-years	# Unique workers	Mean	St.d.	Mean	St.d.	Mean	St.d.
1988–1992	85.4 (98.7%)	25.1 (98.4%)	1.3	0.9	33.3	10.8	7.0	4.3
1992–1996	85.6 (98.1%)	25.8 (97.7%)	1.3	0.9	33.9	10.8	7.4	4.3
1996–2000	90.2 (97.3%)	27.9 (96.8%)	1.3	0.8	34.0	10.8	7.8	4.2
2000–2004	102.1 (97.0%)	31.4 (96.6%)	1.1	0.8	34.2	10.9	8.7	4.0
2004–2008	123.7 (97.4%)	36.2 (97.1%)	0.9	0.8	34.4	11.1	9.5	3.8
2008–2012	151.0 (98.0%)	42.8 (97.5%)	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 described 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. 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.

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 partic-

Table 4. Frequency of switches, by period

	(1) # Unique workers	(2) Average # of jobs	(3) % 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.

ularly 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 et al. (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:

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

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. To better estimate firm effects, Bonhomme et al. (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. In ongoing work, we are working on further improving our algorithm and sample selection.

5.2 Factors influencing the estimated firm and worker 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,[19] 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.

Since the firm component of earnings is fixed at the subperiod level, all our results below are for subperiods. 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, REV_j average revenues per worker, CAP_j a firm's capital stock, FTE_j the number of full-time equivalent workers, EXPORT_j the ratio of exports to revenues at the firm, and ENTRY_j and EXIT_j dummies denoting whether the firm entered or exited during the subperiod.[20] For each subperiod, we regress versions of

$$a_j = \gamma_0 + \gamma_1 \text{VA}_j + \gamma_2 \text{REV}_j + \gamma_3 \text{FTE}_j + \gamma_4 \text{CAP}_j + \gamma_5 \text{EXPORT}_j + \gamma_6 \text{ENTRY}_j + \gamma_7 \text{EXIT}_j + \varepsilon_j$$

All our regressions are run at the firm-subperiod level and weighted by employee-years. We report results both with and without subsector controls.

As we will show, value added per worker is by far the most important determinant of the firm

[19] 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.

[20] Importantly, the exit and entry indicators denote whether a firm completely exited or entered the formal sector, and not whether it entered or exited the PIA subpopulation of firms.

component of pay. Thus, to investigate further the role of value added per worker, we set all other coefficients to zero and regress in each subperiod the firm effect on a constant and a linear term in average log value added per worker,

$$a_j = \gamma_0 + \gamma_1 \mathrm{VA}_j + \varepsilon_j$$

Based on the estimated coefficient, we calculate the variance of the predicted firm effects as

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

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. In the first, we keep $\hat{\gamma}_1$ constant at the estimated level in 1996–2000 and let the variance of value added per worker change as in the data. The second results from keeping 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. A comparison of the two counterfactuals allows us to address whether a change in the variance of firm pay is explained by changes in the underlying dispersion in value added per worker across firms or due to a change in the degree of pass-through from firm value added to worker pay.

6 Results

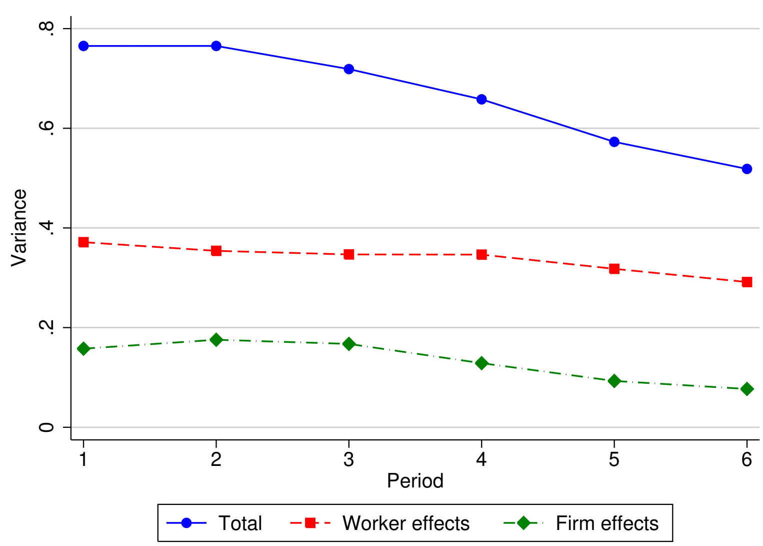
In this section, we first present the results from our two-way fixed effects model decomposing earnings inequality into a firm and a worker component. Second we discuss the importance of sectors for changes in the firm component of pay. Thirdly, we investigate the role of a reallocation of workers across firms versus an intrinsic change to the way firms compensate their workers. Fourth, we relate the firm components of pay to underlying characteristics of firms, and finally we investigate the assumptions imposed by our econometric model.

6.1 AKM decomposition

Table 5 presents the variance decomposition from equation (1) based on the results of the AKM estimation for each five-year subperiod from 1988 to 2012. To illustrate this decomposition over time, Figure 6 plots the variance of raw earnings (blue circles), the variance of estimated worker effects (red squares), and the employee-weighted variance of firm effects (green diamonds) in each subperiod. At least two important conclusions can be drawn from our estimation results. First, although firm heterogeneity is a non-negligible source of earnings inequality, worker heterogeneity is the single most important factor. In the 1996–2000 subperiod the variance of worker fixed effects is 48 percent of the variance of total earnings. This increases monotonically to 56 percent in the last subperiod. The variance of firm fixed effects is 23 percent of the variance of earnings in the 1996–2000 subperiod and decreases to 15 percent in the last subperiod.

Second, in terms of explaining changes over time, we observe a disproportionate 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 35 to 29 log points. Additionally, the declining variance of each component is reflected in a lower covariance between the two (the correlation between worker and firm effects remains fairly constant at about 0.25). Given the large role played by firms in the decline, it is important to understand what led firms to pay more equally over time.

Figure 6. Variance decomposition from AKM model with firm and worker fixed effects



When studying the link between firm effects and firm performance, we are limited to the manufacturing and mining sector for which we have data on firm performance and characteristics.[21] Table 6 compares AKM estimates for this subpopulation with the overall population. The over-

[21] As noted earlier, we impose the restriction to the PIA subpopulation after estimating the AKM model on the entire population of prime age males.

Table 5. Summary statistics from AKM model, by period

	(1) 1988–1992	(2) 1992–1996	(3) 1996–2000	(4) 2000–2004	(5) 2004–2008	(6) 2008–2012	(7) Cumulative change from 1988 to 2012	(8) Cumulative change from 1996 to 2012
Variance of log earnings	0.77 (100.0%)	0.77 (100.0%)	0.72 (100.0%)	0.66 (100.0%)	0.57 (100.0%)	0.52 (100.0%)	-0.25 (100.0%)	-0.20 (100.0%)
Variance of worker effects	0.37 (48.5%)	0.35 (46.3%)	0.35 (48.3%)	0.35 (52.6%)	0.32 (55.5%)	0.29 (56.2%)	-0.08 (32.4%)	-0.06 (27.6%)
Variance of firm effects	0.16 (20.6%)	0.18 (23.0%)	0.17 (23.3%)	0.13 (19.6%)	0.09 (16.2%)	0.08 (14.8%)	-0.08 (32.8%)	-0.09 (45.1%)
Variance of Y_t	0.02 (2.4%)	0.02 (2.0%)	0.00 (0.2%)	0.00 (0.2%)	0.00 (0.3%)	0.00 (0.5%)	-0.02 (6.3%)	0.00 (-0.6%)
2×Cov. worker and firm effects	0.13 (16.5%)	0.13 (17.6%)	0.14 (19.2%)	0.12 (18.2%)	0.10 (17.8%)	0.09 (17.8%)	-0.03 (13.7%)	-0.05 (23.0%)
2×Cov. worker effects and Y_t	0.01 (1.8%)	0.01 (1.6%)	0.01 (1.1%)	0.01 (1.5%)	0.02 (2.6%)	0.02 (3.1%)	0.00 (-1.1%)	0.01 (-4.2%)
2×Cov. firm effects and Y_t	0.00 (0.4%)	0.00 (0.0%)	0.00 (0.3%)	0.00 (0.5%)	0.00 (0.7%)	0.00 (0.8%)	0.00 (-0.4%)	0.00 (-0.8%)
Residual	0.08 (9.9%)	0.07 (9.5%)	0.05 (7.6%)	0.05 (7.4%)	0.04 (6.8%)	0.04 (6.8%)	-0.04 (16.3%)	-0.02 (9.9%)
# workers-years	85.4	85.6	90.2	102.0	124.0	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 $Var\left(y_{it}\right) = Var\left(a_i\right) + Var\left(a_j\right) + Var\left(\hat{Y}_t\right) + Var\left(e_{it}\right) + 2Cov\left(a_i, \hat{Y}_t\right) + 2Cov\left(a_j, \hat{Y}_t\right) + 2Cov\left(a_j, \hat{Y}_t\right)$.

Cells contain variance explained by each decomposition element. The share of the total variance explained by each decomposition element is given in parentheses.

all variance of log earnings is five log points higher in the PIA subpopulation in the 1996–2000 subperiod. It first falls at a slightly slower pace than in the overall population, then at a slightly faster pace, so that in the last subperiod the overall variance is again five log points greater. The variance of worker effects is two 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 both the 1996–2000 and 2008–2012 subperiods. We conclude from this that trends in inequality are similar in the PIA subpopulation as in the overall population.

6.2 The importance of sectors

We start our investigation of the firm component by investigating the importance of sectoral differences in the firm component of pay. These decompositions are done for the entire subpopulation of prime age males. To this cause, we follow the strategy used in Section 4 to decompose a trend into a between and within component. In each subperiod, we calculate the average firm effect and the variance in firm effects within 26 sectors. Subsequently, we compute the variance of the average as our measure of between-sector variance and the average of the variance as the within-sector variance. All calculations are weighted by worker-years.

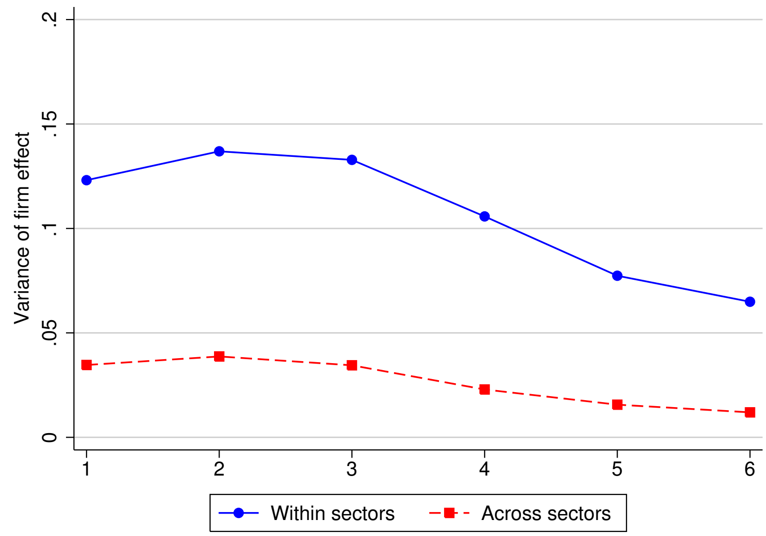
Figure 7 plots the results. Most inequality arises within sectors, but there are also differences in means across sectors. In 1996–2000 (2008–2012), about 20 (16) percent of the overall variance of firm effects arises across sectors. Over this period, the within-sector variance of firm effects rougly halves whereas the across-sector variance falls by 65 percent. However, as within-sector inequality is most important in levels, the fall in this accounts for 76 percent of the overall fall in the variance of firm effects.

Table 6. Comparison of AKM results between workers at large manufacturing&mining firms versus population, by period

	(1) 1996–2000	(2) 2000–2004	(3) 2004–2008	(4) 2008–2012
Total variance of log earnings (% of pop. estimate)	0.77 (106.6%)	0.73 (111.5%)	0.64 (112.2%)	0.57 (110.8%)
Variance of individual effects (% of pop. estimate)	0.37 (107.8%)	0.38 (108.5%)	0.36 (111.8%)	0.33 (112.3%)
Variance of firm effects (% of pop. estimate)	0.15 (88.7%)	0.12 (94.5%)	0.08 (86.5%)	0.06 (83.7%)
Variance of Y_t (% of pop. estimate)	0.00 (101.1%)	0.00 (99.6%)	0.00 (96.2%)	0.00 (98.9%)
2×Covariance between ind. and firm effects (% of pop. estimate)	0.18 (126.6%)	0.17 (141.2%)	0.14 (138.0%)	0.12 (126.5%)
2×Covariance between ind. and Y_t(% of pop. estimate)	0.01 (118.9%)	0.01 (124.5%)	0.02 (117.7%)	0.02 (120.0%)
2×Covariance between firm and Y_t (% of pop. estimate)	0.00 (132.8%)	0.00 (115.8%)	0.00 (113.7%)	0.01 (133.4%)
Variance of the residual (% of pop. estimate)	0.05 (94.9%)	0.05 (95.8%)	0.04 (97.5%)	0.03 (98.5%)
# Workers-years (% of pop. estimate)	16.60 (18.4%)	18.00 (17.6%)	23.20 (18.7%)	26.90 (17.8%)
# Unique workers (% of pop. estimate)	5.84 (21.0%)	6.27 (20.0%)	7.80 (21.5%)	8.98 (21.0%)
# Firm-years (% of pop. estimate)	0.11 (2.7%)	0.13 (2.6%)	0.16 (2.5%)	0.18 (2.3%)
# Unique firms (% of pop. estimate)	0.05 (4.0%)	0.06 (4.0%)	0.07 (4.0%)	0.08 (3.5%)
R^2 (% of pop. estimate)	0.93 (100.9%)	0.94 (101.1%)	0.94 (101.0%)	0.94 (100.8%)

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.

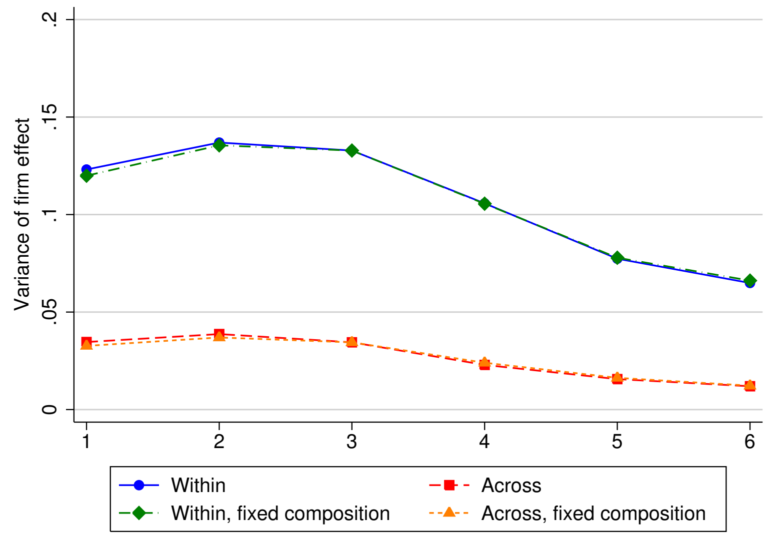
Figure 7. Firm effects within and across sectors



The precence of differences in means across sectors implies that some of the between-sector compression could be because of a reallocation of workers towards more equally paying sectors. Hence it need not be that sectors became fundamentally more equal in terms of pay. Furthermore, although we did not report it above, there are differences across sectors in the within-sector variance of the firm component of pay. Thus a decline in the within-sector variance of firm effects could have been driven by a reallocation of workers towards less unequal sectors, and not because sectors became fundamentally more equal.

To investigate the importance of reallocation versus intrinsic compression, we hold the distribution of workers across sectors constant at its 1996–2000 level and compute the within- and between-sector variance using these constant weights. As can be seen from Figure 8, holding the distribution of workers across sectors constant has essentially no impact on either the within- or the across-sector variance of firm effects. We conclude that most of both the level and the fall in the variance of firm effects happens within sectors, and almost all of the fall is due to firms within sectors becoming fundamentally more similar in terms of pay.

Figure 8. Firm effects within and across sectors, the role of composition



6.3 Underlying distribution of firms vs. allocation of workers

The reduction in the dispersion of the firm-specific component of pay could be decomposed into two sources. Firstly, firms could have intrinsically become more equal over time. Secondly, firms could have remained just as unequal, but a reallocation of workers across firms could have resulted in a more equal distribution of firm-specific pay. A way to assess which of these forces is more prominent is looking at the unweighted distribution of firm effects, since by construction this holds the weight placed on each firm constant.[22] This thus investigates any change in the underlying distribution of firm effects. The left pane of Figure 9 shows a significant and monotonic compression in the unweighted distribution of firm effects over time, indicating that firms fundamentally became more similar over time in terms of pay.

Conversely, in order to investigate the importance of worker movement between firms in explaining the inequality decline, we rank firms in each subperiod based on their estimated firm effects. Subsequently, we consider the distribution of workers across firm ranks. If workers have reallocated across firms so as to produce a more equal distribution of firm effects, we would expect the distribution of workers across firm ranks to compress. As can be seen in the right pane of

[22]Subject to the caveat that we do not observe firm effects for firms after they have exited and before they have entered, and thus they get a weight of zero both in the weighted and unweighted distribution. We discuss in detail the role of entry and exit below.

Figure 9, there is little evidence of such reallocation of workers. We conclude that most of the compression in in the firm component of pay appears to be driven by firms fundamentally becoming more similar in terms of pay.

Figure 9. Allocation of workers vs. underlying firm effects distribution

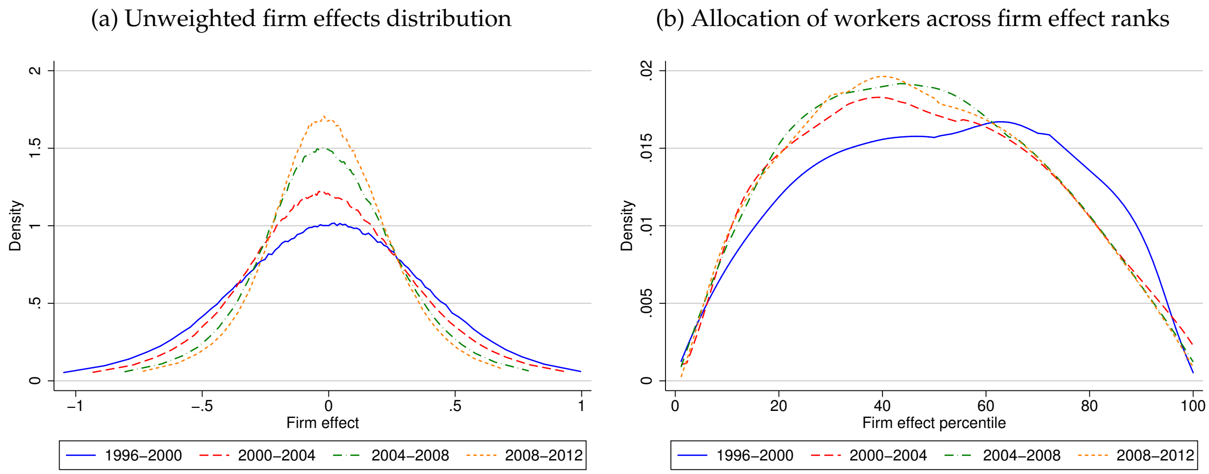


Table 7 investigates the importance of entry and exit of firms for the level and trend in the variance of firm effects by comparing the overall variance of firm effects to that among only incumbent and only non-exiting firms in each period. To the extent that new firms meaningfully affect the variance of firm effects, we would expect the total variance to be significantly different from the variance among only incumbent firms. Similarly, if the exit of firms had an important impact on inequality, we would expect the variance of non-exiting firms to be significantly different from the total variance.

Comparing row one and two in Table 7, the variance of incumbent firms is very similar to the total variance. Thus, the entering of firms does not meaningfully affect the overall variance of firm effects. Similarly, non-exiting firms in row three are not much different from the overall population of firms. We conclude from this that exit and entry of firms does not significantly affect the overall variance of firm effects. As a corrollary we find little scope for the churning of firms to be an important driver behind the decreasing dispersion of firm effects over time.

Table 7. Variance of firm effects with and without entrant and exiting firms

	(1) 1988–1992	(2) 1992–1996	(3) 1996–2000	(4) 2000–2004	(5) 2004–2008	(6) 2008–2012
Total variance	0.16	0.18	0.17	0.13	0.09	0.08
Variance of incumbents	0.15	0.17	0.16	0.13	0.09	0.08
Variance of non-exiters	0.15	0.18	0.17	0.13	0.09	0.08

Note: A firm is an entrant if in any year during the subperiod it existed but it did not exist the year before. A firm is an exiter if it existed in any year during the subperiod but not in the subsequent year.

6.4 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 can explain the changes we observe in the firm-specific component of pay over time. Using measures of firm performance from the PIA data, Table 8 reports results from regressing estimated firm effects on firm characteristics, controlling for two-digit subsectors.

Several features are worth highlighting. First, both value added per worker and firm size are associated with higher firm components of pay (the same is true for revenues per worker and capital per worker). Exiters also appear to pay more, whereas there is no consistent, statistically significant difference between entrants and non-entrants. In the early periods, a greater export intensity was associated with a lower firm effect, but this appears to have vanished over time. Overall, from a cross-sectional standpoint, larger and better performing firms have a higher firm component of pay. Second, the amount of dispersion in firm effects explained by firm performance is notable, with an R^2 around 0.7. In fact, a linear regression of firm effects on a constant and log value added per worker alone explains 47–58 percent of the variance in firm effects. As adding additional measures of firm performance only boosts the explanatory power of the model marginally, we focus our discussion below on the relationship between value added per worker and the firm component of pay.

As can be seen in Table 8, 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 0.18 log point increase in the estimated firm effect; in 2008–2012, the same increase in value added per worker was only associated with a .10 log point higher firm effect. Translating this into inequality, we find that a change in the variance of value added per worker as well as the pass-through from it to pay explains 25 percent of the overall decline in the variance of log earnings in Brazil over this period. The remaining four log point decline in the variance of firm effects is due to factors orthogonal to value added per worker.

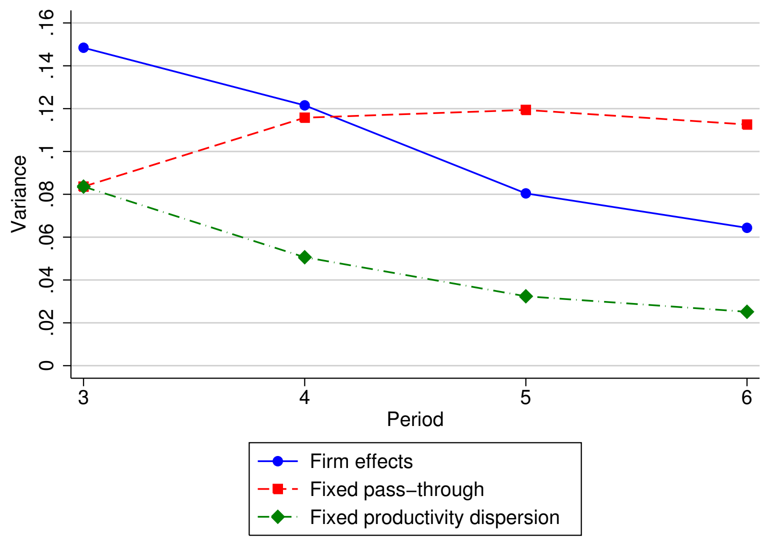
Table 8. Regression of estimated firm effects on firm characteristics

	1996–2000		2000–2004		2004–2008		2008–2012	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
V.a. p.w.	0.211***	0.181***	0.166***	0.152***	0.129***	0.117***	0.112***	0.099***
	(0.004)	(0.004)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Firm size		0.047***		0.033***		0.021***		0.019***
		(0.003)		(0.003)		(0.002)		(0.002)
Export intensity		-0.067***		-0.127***		-0.026*		-0.013
		(0.019)		(0.016)		(0.014)		(0.014)
Exit indicator		0.093***		0.079***		0.069***		0.029**
		(0.017)		(0.011)		(0.011)		(0.012)
Entry indicator		0.018		0.035**		0.017		0.044***
		(0.016)		(0.014)		(0.014)		(0.009)
Sector effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker-years	15.5	15.5	16.7	16.7	21.0	21.0	23.9	23.9
R^2	0.706	0.726	0.700	0.714	0.675	0.684	0.667	0.677

Note: Dependent variable is the estimated firm effect a_j. Number of worker-years in millions. Standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

To further 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 10 plots the result from this exercise. The total variance of firm effects (solid blue line with circles) declines from 15 to 6.5 log points, the predicted variance from value added per worker holding pass-through constant (dashed red line with squares) increases from eight to 11 log points, and the predicted variance holding the dispersion in value added per worker constant (dash-dotted green line with diamonds) falls from eight to two log points. 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 10. Variance of firm effects, variance predicted by fixed pass-through and variance predicted by fixed productivity dispersion



6.5 Empirical support of the AKM model

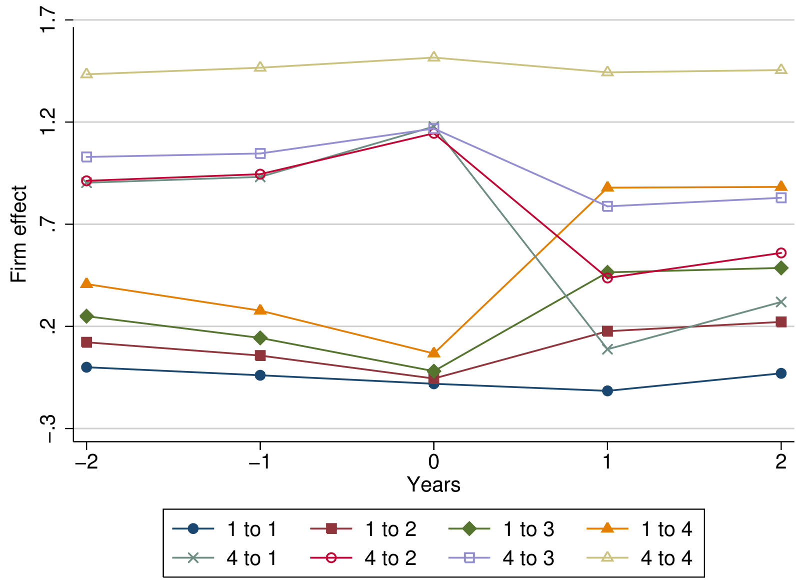
Although the consistently high R^2 of our model suggests that the model fits the data well, our estimates can be biased if the residual is correlated with either the firm or worker component of earnings. To investigate this further, we replicate two exercises conducted by Card et al. (2013) in the case of Germany in our Brazilian data.

Figure 11 shows the average firm effect of workers who switch firms up to two years prior to the swtich 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.[23] This suggests that our additive model is consistent with the pattern observed among workers who transition between firms.

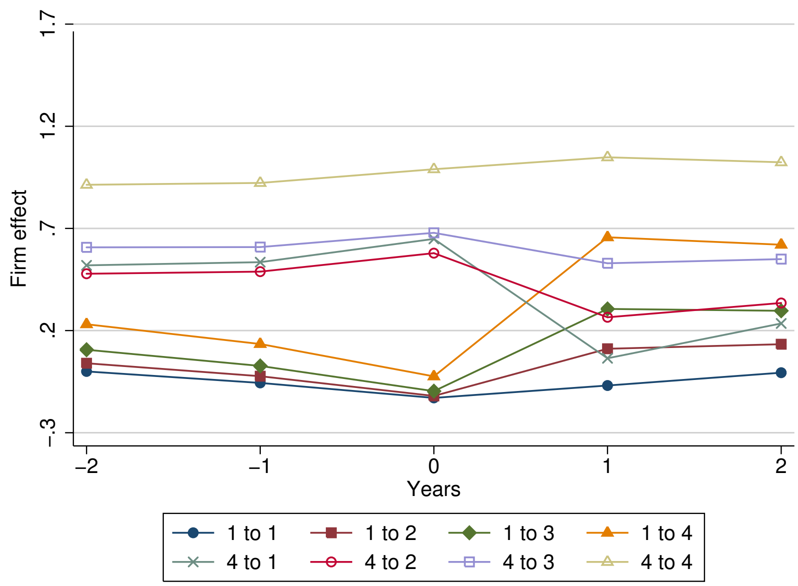
[23]For dispositional ease we only show switches out of the first and fourth quartile, but other quartiles display similar pattern and are available on request.

Figure 11. Average changes in firm effects of workers that switch firms, classified by firm effect deciles of pre and post transition firms

(a) 1996–2000

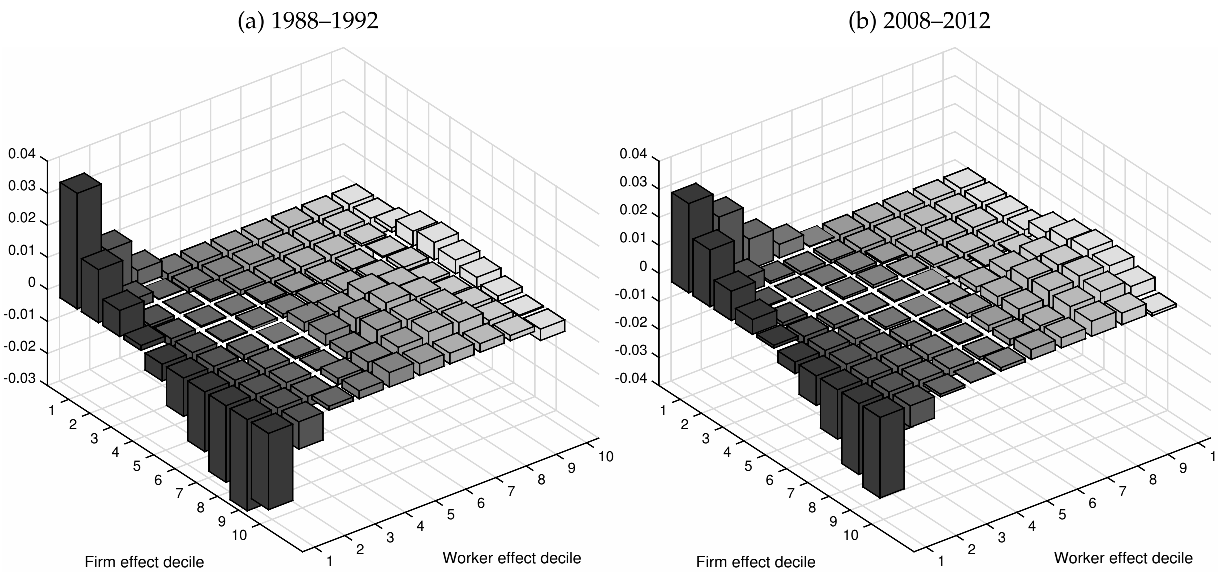


(b) 2008–2012



Second, we find little evidence in the data for match effects that are systematically correlated with either person of match effects. Figure 12 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.[24] However, the magnitude of the error is modest, and beyond the lowest two deciles of workers errors do not exhibit any systematic relationship with firm and worker effects. This boosts our confidence that the log additive assumption is a good description of the Brazilian labor market.

Figure 12. AKM residual by firm and worker fixed effect deciles



7 Conclusion

In this paper, we estimate two-way fixed effects models controlling for unobserved worker and firm heterogeneity in order to understand the sources of a substantial decline in earnings inequality in Brazil between 1996 and 2012. We find that while the firm-specific components of pay only explain 15–23 percent of the variance of log earnings, a compression in firm effects explains 45 percent of the decline in earnings inequality. Worker effects, on the other hand, explain 48–56 percent of the level of inequality, but only 28 percent of the decline. Thus compression in pay across firms played an outsized role behind the decline.

Furthermore, although measures of firm performance are strongly positively correlated with the firm component of pay, compression in such measures was not a factor behind declining in-

[24]The fact that they have a positive residual while working at high paying firms is mechanical since the residuals have to sum to zero within each worker type. Although we do not investigate this further, it is consistent with a frictional labor market with a binding minimum wage as in Engbom and Moser (2015).

equality. Instead, we show that more than half of the compression in firm effects is due to a declining pass-through from firm productivity to pay. In terms of overall changes in the distribution of earnings, a declining pass-through from firm productivity to pay explains 25 percent of the compression between 1996 and 2012. In ongoing work, we also decompose the compression in worker effects into compression in observable worker characteristics, compression in the return to such characteristics, and residual compression.

Our paper suggests a set of stylized facts that a potential theory of the inequality decline in Brazil would have to match. Such a theory must generate pay differences between firms for identical workers, and such pay differences must be strongly positively correlated with firm productivity. Moreover, it needs to generate a compression in such pay differences over time, but not through compression in firm productivity. Instead, it has to produce a significantly weaker link between firm productivity and pay. We think that promising candidates behind the decline in inequality in Brazil are changes in wage setting induced by for instance changes in the minimum wage or labour contract regulation.

References

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

Abowd, John M., Robert H. Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," 2002.

Atkinson, Anthony B. and François Bourguignon, eds, *Handbook of Income Distribution*, vol. 2 ed., Elsevier, 2015.

Bagger, Jesper, Bent Jesper, and Royal Holloway, "Productivity and Wage Dispersion: Heterogeneity or Misallocation," 2014, pp. 1–40.

_ , Kenneth L Sorensen, and Rune Vejlin, "Wage sorting trends," *Economics Letters*, 2013, *118* (1), 63–67.

Barros, Ricardo, Mrela de Carvalho, Samuel Franco, and Rosane Mendonça, "Markets, the State and the Dynamics of Inequality: The Case of Brazil," *UNDP Discussion Paper*, jun 2010.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Measuring and Analyzing Cross-country Differences in Firm Dynamics," in "NBER Chapters," National Bureau of Economic Research, Inc, 2009, pp. 15–76.

_ , _ , and _ , "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 2013, *103* (1), 305–334.

Barth, Erling, James C Davis, and Alex Bryson, "It's Where You Work: Increases in Earnings Dispersion across Establishments and Individuals in the U.S.," *NBER Working Paper 20447*, 2014.

Bloom, Nicholas, Fatih Guvenen, David Price, and Jae Song, "Evolution of Inequality within Firms (SED Abstract)," 2015.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," 2015.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," 2015.

_ , Joerg Heining, and Patrick Kline, "Workplace heterogeneity and the rise of West German wage inequality," *Quarterly Journal of Economics*, 2013, *128*, 967–1015.

Carvalho Filho, Irineu E. and Marcello M. Estevão, "Institutions, Informality, and Wage Flexibility: Evidence From Brazil," *IMF Working Papers*, 2012, *12* (84), 1.

Córdova, José Ernesto López and Mauricio Mesquita Moreira, "Regional Integration and Productivity: The Experiences of Brazil and Mexico," *IDB Publications (Working Papers)*, 2003.

de Araujo, Ana Luisa Pessoa, "Wage Inequality and Job Stability," 2014, pp. 1–53.

Deaton, Angus, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy* 1997.

Dix-Carneiro, Rafael and Brian K. Kovak, "Trade Reform and Regional Dynamics: Evidence From 25 Years of Brazilian Matched Employer-Employee," 2015.

Engbom, Niklas and Christian Moser, "Earnings Inequality and the Minimum Wage," 2015.

Faggio, G., K. G. Salvanes, and J. Van Reenen, "The evolution of inequality in productivity and wages: panel data evidence," *Industrial and Corporate Change*, oct 2010, *19* (6), 1919–1951.

Ferreira, Francisco H. G., Phillippe G. Leite, and Matthew Wai-Poi, *Trade Liberalization, Employment Flows, And Wage Inequality In Brazil* Policy Research Working Papers, The World Bank, jan 2007.

Ferreira, Pedro Cavalcanti and Jose Luiz Rossi, "New Evidence from Brazil on Trade Liberalization and Productivity Growth," *International Economic Review*, 2003, *44* (4), 1383–1405.

Garcia, Márcio, Diogo Guillén, and Patrick Kehoe, "The Monetary and Fiscal History of Latin America: Brazil," *Working Paper*, 2014.

Gruetter, Max and Rafael Lalive, "The importance of firms in wage determination," *Labour Economics*, 2009, *16* (2), 149–160.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen Redding, "Trade and Inequality: From Theory to Estimation," 2013.

Kopczuk, Wojciech, Emmanuel Saez, and Song Jae, "Earnings inequality and mobility in the United States: evidence from social security data since 1937.," *Quarterly Journal of Economics*, 2010, *125* (1), 91–128.

Lopes De Melo, Rafael, "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," 2013, pp. 1–39.

Lopez, J. Humberto and Guillermo Perry, "Inequality in Latin America: Determinants and Consequences," 2008.

Medeiros, Marcelo, Pedro H G Ferreira De Souza, and Fábio Avila de Castro, "O topo da distribuição de renda no Brasil: primeiras estimativas com dados tributários e comparação com pesquisas domiciliares , 2006 - 2012," 2014, pp. 2006–2012.

Menezes-Filho, Naércio Aquino, Marc-Andreas Muendler, and Garey Ramey, "The Structure of Worker Compensation in Brazil, with a Comparison to France and the United States," *Review of Economics and Statistics*, may 2008, *90* (2), 324–346.

Moreira, Mauricio Mesquita, "Brazil's Trade Liberalization and Growth: Has it Failed?," *International Trade*, dec 2004.

Muendler, Marc-Andreas, "Trade, Technology, and Productivity: A Study of Brazilian Manufacturers, 1986-1998," *CESifo Working Paper Series*, 2004.

Pavcnik, Nina, Andreas Blom, Pinelopi Goldberg, and Norbert Schady, "Trade Liberalization and Industry Wage Structure: Evidence from Brazil," *World Bank Economic Review*, 2004, *18* (3), 319–344.

Tsounta, Evridiki and Anayochukwu I. Osueke, "What is behind Latin America's Declining Income Inequality?," 2014.

Ulyssea, Gabriel, "Firms, Informality and Development: Theory and evidence from Brazil," 2014.

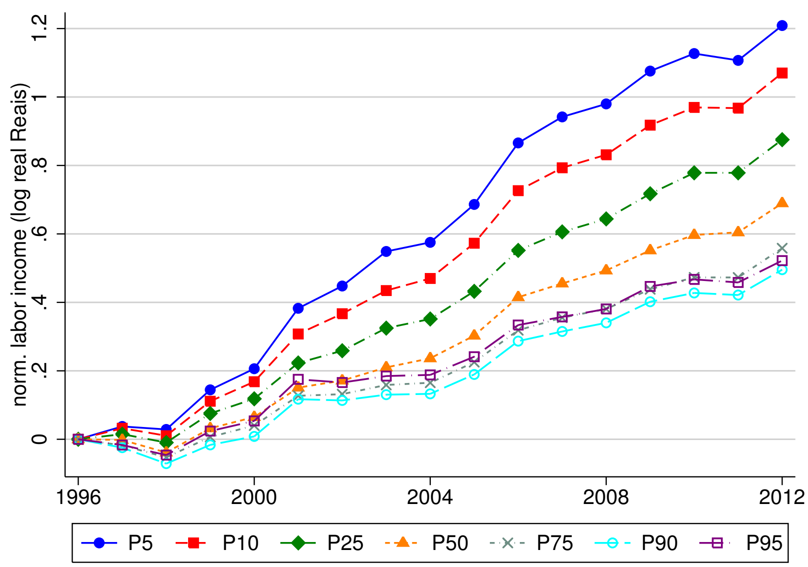
World Bank, "World Development Indicators," 2015.

Appendix

A Additional figures

A.1 Earnings levels evolution

Figure 13. Earnings levels evolution



A.2 Alternative productivity measures

Figure 14. Cross-sectional comparison of alternative productivity measures

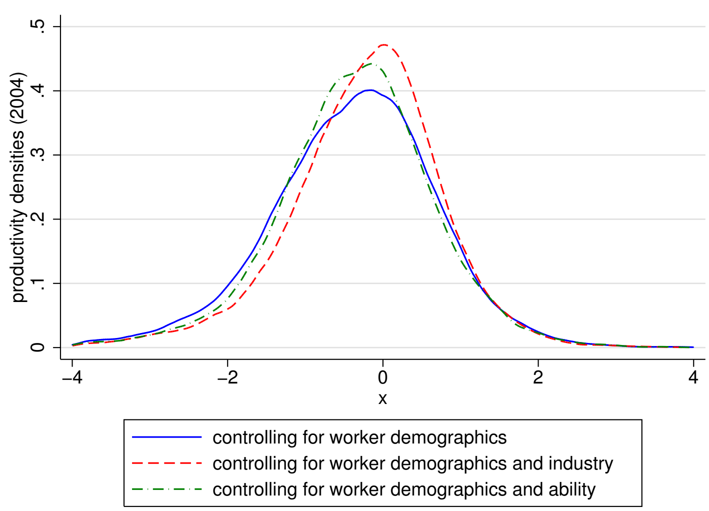
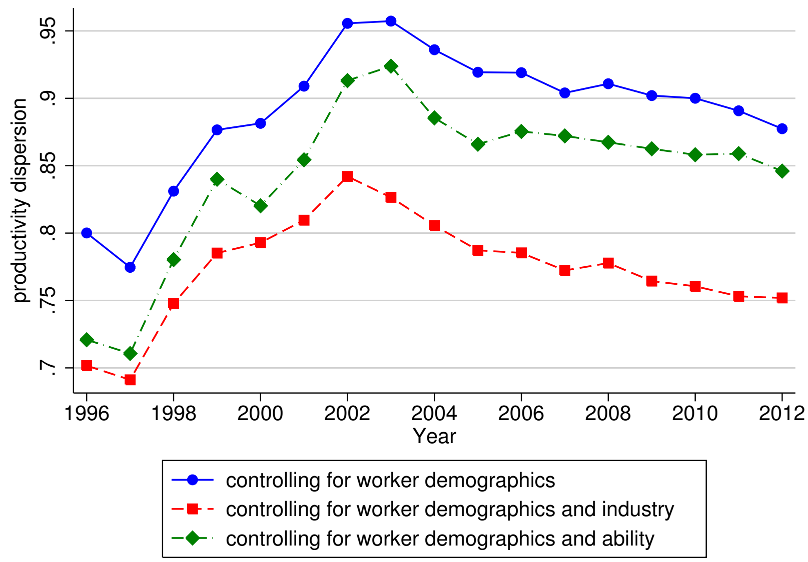


Figure 15. Evolution of dispersion of alternative productivity measures



A.3 Earnings inequality by education groups

Figure 16. Within education group inequality