

**Outline.** The rest of the paper is structured as follows: Section 2.2 provides an overview of the main institutional changes and macroeconomic trends affecting Brazilian labor markets from 1988 to 2012. Section 2.3 summarizes the administrative datasets used in our empirical analysis and discusses sample selection and variable definitions. Section 2.4 provides descriptive statistics on trends in earnings inequality in Brazil during this time. Section 2.5 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 2.6 presents our main empirical results as well as checks on the validity of our empirical framework. Finally, Section 2.7 summarizes our key findings and concludes.

## 2.2 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 bar-

---

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

riers. 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 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. (2016) 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

---

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

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

## 2.3 Data

Our analysis uses two confidential administrative datasets from Brazil: the *Relação Anual de Informações (RAIS)* contains earnings and demographic characteristics of workers as

---

<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.

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.

### 2.3.1 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 substantial variation in pay across establishments within firms, our firm-level anal-

---

<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).

ysis 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).

---

<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.

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

**Descriptive statistics.** Table 2.1 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 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.

---

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

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.

### 2.3.2 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.<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 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,

---

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



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 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 2.2 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–

---

<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.

- Bagger, Jesper, Fabien Postel-Vinay and Jean-marc Robin. 2014. "Tenure , Experience , Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics." *American Economic Review* pp. 1–54.
- 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*. National Bureau of Economic Research, Inc pp. 15–76.
- 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, Alex Bryson, James C Davis and Richard Freeman. 2014. "It's Where You Work: Increases in Earnings Dispersion Across Establishments and Individuals in The US." *NBER working paper* (w20447).
- Beck, Cihak, Demirguc-Kunt, Freyen and Levine. N.d. "Financial Development and Structure Dataset. Updated April 2013. World Bank."
- Bernard, Andrew B. and J. Bradford Jensen. 1995. "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987." *Brookings Papers: Microeconomics* .
- Bliss, Christopher and Nicholas Stern. 1978. "Productivity, wages and nutrition." *Journal of Development Economics* (4):331–362.