The Contribution of Workers, Workplaces, and Sorting to Wage Inequality in Mexico

Jorge Pérez Pérez* José G. Nuño-Ledesma†

December 16th, 2022 ‡

Abstract

We use a matched worker-workplace dataset comprising the near universe of formal private-sector workers in Mexico to estimate the contribution of average workplace-specific wage premia, worker-level characteristics, and assortative matching on Mexico's wage inequality between 2004 and 2018. To this end, we regress log earnings on sets of worker and workplace fixed effects and perform a decomposition of total wage variance. We find that assortative matching explains between 16% and 19% of total wage variance, while worker- and workplace-specific factors contribute between 35% to 42% and 33% to 38%, correspondingly. The importance of workplace factors in determining wage inequality correlates negatively with regional economic development: it is lowest in the North, Mexico's most-developed region, and largest in the South, the country's least-prosperous region.

Keywords: assortative matching, employment, firm wage premia, wage inequality

JEL Codes: E24, J20, J31, O15, R23

‡We received valuable insights from Alejandrina Salcedo, Cecilia García, Rodolfo Oviedo, Alex Rivadeneira, and seminar audiences at Banco de México. Karla Neri Hernández, Jorge Emilio Meléndez, Diego Mayorga, Marco González, Carolina Crispín, and Guillermo Mondragón provided great research assistance. The views and conclusions presented in this document are the exclusive responsibility of the authors and do not necessarily reflect those of Banco de México. We acknowledge financial support from Colombia Cientifica – Alianza EFI #60185 contract #FP44842- 220-2018, funded by The World Bank through the Scientific Ecosystems, managed by the Colombian Ministry of Science, Technology, and Innovation (MINCIENCIAS). The data was accessed through the Econlab at Banco de México. The EconLab collected and processed the data as part of its effort to promote evidence-based research and foster ties between Banco de México's research staff and the academic community. Inquiries regarding the terms under which the data can be accessed should be directed to econlab@banxico.org.mx.

*Banco de México. Corresponding author. Email: jorgepp@banxico.org.mx.

†University of Guelph. Email: jnuno@uoguelph.ca.

1 Introduction

In most countries, there are substantial wage inequalities between cities and sub-national regions. These gaps have risen over the last decades, attracting the attention of economists, policymakers, and the public at large (Katz et al., 1999; Acemoglu and Autor, 2011). Consequently, some countries have implemented measures to attenuate the adverse effects of growing salary disparities between communities (Kline and Moretti, 2014). At the same time, the literature that investigates plausible reasons behind the expansion of regional wage divergences is growing. Current work in this area tends to study developed economies with modern institutions. Labor markets in developing countries –with historically-persistent dissimilarities among sub-national economies– have received relatively less attention. In this paper, we leverage estimations from two-way fixed effects models to explore the role of worker-level characteristics, workplace-specific wage premia, and worker-workplace assortative matching as determinants of wage inequalities in Mexico and its geographical regions.[1]

As noted by Bassier (2022), there needs to be more research examining the forces driving inequalities in labor outcomes in developing markets in general, and in contexts with structurally high unemployment in particular. Work looking at developing countries frequently attributes a larger role in explaining wage dispersion to firms, compared to developed countries (Alvarez et al., 2018; Gerard et al., 2021; OECD, 2021; Frías et al. 2022; Bassier 2022). An important aspect of our work is investigating whether the association between lower economic prosperity and higher relevance of worker-level premia in wage inequality is present within countries. The Mexican economy offers a good case study to explore this question. Marked regional disconnections are present in Mexico; most notably between the relatively thriving North; the moderately successful Center and Center-North regions, and the substantially less affluent South.[2] Structural discrepan-

[1]Throughout the document, we use the terms "firm" and "workplace;" "worker" and "person," and "sorting" and "assortative matching" interchangeably.

[2]We use the regional classification defined by the Mexican Central Bank. The regions cluster states according to geographical proximity and economic similarity in indicators such as employment, the prevalence of the agricultural, manufacturing, and tourism sectors, and level of retail sales, among others (Banco de México, 2011). The regions contain the following Mexican states: the *North* includes Baja California, Chihuahua, Coahuila, Nuevo León, Sonora, and Tamaulipas; the *Center-North* gathers Baja California Sur, Aguascalientes, Colima, Durango, Jalisco, Michoacán, Nayarit, San Luis Potosí, Sinaloa and Zacatecas; the *Center* is comprised by Mexico City, Estado de México, Guanajuato, Hidalgo, Morelos, Puebla, Querétaro and Tlaxcala; the *South* contains Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz, and Yucatán.

cies include pronounced differences in industry specialization, different degrees of informality in local markets, and recent disjoint tendencies in the prevalence of formal employment (Alcaraz et al. 2015; Chávez-Martín del Campo and García Loredo 2015; Rangel González and Llamosas-Rosas 2021; Juárez-Torres et al. 2022). In addition, labor markets in the Northern, Central-Northern, and Central regions exhibit a high degree of integration between them and move in concert with national employment trends. In contrast, markets in the South do not share the same underlying economic cycles, and shocks stemming from this area tend not to propagate to the rest of the country (Delajara 2011; Delajara 2013). These sub-national peculiarities and the availability of detailed and reliable social security records from workers across the country, make the Mexican labor market a compelling scenario to investigate the sources of regional wage inequalities in the context of a developing economy.

To measure the contribution of workers, workplaces, and their sorting to regional wage inequalities in Mexico, we fit models *à la* Abowd, Kramarz, and Margolis (henceforth AKM; Abowd et al. 1999) for the period between 2004 and 2018. The estimated AKM models offer a good approximation of the determinants of wages, explaining over 90% of the variation in wages in all regions. Our analysis suggests that average workplace-specific wage premia explain around a third of the total salary variance, playing a remarkable role in determining wage inequality in Mexico. Although comparable to other developing economies, the ability of workplaces to set wages in Mexico is substantially stronger compared to other countries members of the OECD (OECD, 2021). More interestingly, mean workplace-specific wage premia do explain a more significant proportion of wage variations in the South compared to the rest of the country and a smaller proportion in the North. These findings provide further evidence in support of an inverse relationship between economic prosperity and the proportional role of workplaces in shaping wage inequality. Lastly, we also encounter evidence indicating that, over time, assortative matching explains an increasing proportion of the variance in salaries in our analysis period. Technologies that facilitate matching between workers and workplaces may play a role in the increased assortativeness of the Mexican labor market.

The rest of the paper proceeds as follows. In the next section, we survey the relevant literature. Section 3 describes the dataset we use. In part 4, we offer some facts about wage inequality for formal workers in Mexico using our dataset. We follow in section 5 by outlining the methodology behind our worker and workplace fixed effects models. Section 6 shows our results about the contribution of workers, workplaces and assortative matching on wage inequality in Mexico and discusses regional differences. Last, section 7 concludes.

2 Relevant Literature

Much of the existing literature explains the sustained rise in local wage disparities through productivity gaps between high- and low-skilled workers (e.g., Katz and Murphy, 1992; Juhn et al., 1993; Goldin and Katz, 2010). However, economists have long recognized that there is a workplace component to wage inequality because some pay higher wages than others to equally skilled employees (Krueger and Summers, 1988; Van Reenen, 1996; Card et al., 2013). The plant component of the variation in compensations can be due to assortative matching, a phenomenon that may emerge in markets with worker and workplace heterogeneity, wherein the most skill-intensive (and productive) workplaces hire highly skilled workers. This pairing process can aggravate geographical disparities because, for example, the regions with a prevalence of already unproductive plants may see their pool of highly productive candidates drained. When worker and workplace quality are complements in production, productivity and remunerations may increase with assortative matching. There is evidence that assortativeness is an important force in determining the wage distribution in several countries (Card et al., 2013; Card et al., 2018; Torres et al., 2018; Dauth et al., 2022).

AKM-style models can help disentangle worker- and workplace-specific contributions to wage variance. We apply the AKM techniques to our data at both national and regional levels, obtaining measures of the extent to which each factor of interest contributes to earning inequities. By providing a formal mechanism to explain regional wage differentials, our work diverges from recent efforts that describe income inequality in Mexico, but remain primarily agnostic about the economic forces underlying the observed trends. For example, Puggioni et al. (2022) use the same dataset as ours to offer a panoramic overview of the recent dynamics of income inequality in Mexico. Although insightful, their work is largely silent regarding the determinants propelling the trends they describe. We address this issue by quantifying the contributions to expanding wage differentials attributable to worker characteristics, average workplace wage premia, and labor sorting.

A related strand of the literature examines the role of worker composition and segregation within workplaces on earnings inequality (Lopes de Melo, 2018; Song et al., 2018). An important insight from these studies is that workers' earnings may vary non-monotonically with respect to the workplace type. Segregation within workplaces would result in non-linearities in the log-wage equation. The main implication for our research is that the effects retrieved from our log-linear earnings model may not admit a structural interpretation, a point already implied by Abowd et al. (1999).

We contribute to the literature on wage disparities and assortative matching in three ways. First, we complement efforts to document wage disparities within countries (Combes et al., 2008; Rice et al., 2006; Boeri et al., 2021; Gerard et al., 2021; Dauth et al., 2022). Second, we expand our understanding of the sources of wage disparities in developing countries. Third, we supplement previous work examining wage variance trends in Latin American countries culturally and economically similar to Mexico (Alvarez et al., 2018; Gerard et al., 2021). These investigations tend to report country-wide patterns resulting from wage-setting policies and non-market non-skill-based sorting, such as discrimination. To our knowledge, we provide the first study detailing the interplay between wage disparities and sorting and worker- and workplace-specific effects in Mexico.

We contribute to a growing literature using administrative data to study labor markets in developing countries. AKM models require detailed information on job and wage histories. This demanding data requirement is one of the reasons why the literature estimates AKM models primarily for countries with rich and reliable administrative data, which tend to be highly developed (e.g., Abowd et al., 1999; Gruetter and Lalive, 2009; Card et al., 2013; Dauth et al., 2022). Our data come from records from over 80% of the formal private-sector workers in Mexico, covering the period from 2004 to 2018. The information we rely on allows us to follow individual workers throughout their entire history in the (formal) labor market.

Frías et al. (2022) apply the same framework we use to a similar dataset but to different ends. They investigate the relationship between increased international trade and wage premia in Mexico. In contrast, we are interested in scrutinizing internal sources of variability in remunerations (as opposed to external factors such as out-of-country demand) and documenting their effect on overall salary inequality.

3 Data

We use social security records from the *Instituto Mexicano del Seguro Social* (IMSS), a Mexican governmental organization that assists public health, pension management, and social security. By law, all salaried workers employed in the private sector must register with IMSS. According to estimates using the National Survey of Occupation and Employment (ENOE), 83% of the formal workforce in 2022 is affiliated with IMSS. Self-employed persons can register with IMSS; if they do so, they can access some parts of the social security system. By default, self-employed workers are registered as earning the equivalent to one legal minimum salary. Records from self-employed workers represent around 0.1% of the complete IMSS database.

IMSS does not collect information regarding workers in the informal economy. Informal employment is high in Mexico, representing around 55% of total employment in 2018 (Instituto Nacional de Estadística y Geografía (INEGI), 2018). Therefore, the dataset we use excludes a substantial number of workers.[3]

The IMSS social security information is published monthly. We use records for the period between November 2004 and December 2018. The number of workers in the database was 12.8 million in November 2004 and 20.1 million by December 2018. Our wage variable of interest is the daily taxable income.[4] We also use information on the period of employment, gender, and year of birth. Wages over 25 minimum wages or 25 UMAs ("units of measure and update") are top-coded.[5] We do not have information on education or hours of work.

IMSS uses the *registro patronal* (employer registry number) as a workplace identifier. The *registro* corresponds to an employer but not a physical location. For example, workers operating in a single plant can work for more than one employer as identified by their *registro patronal*.[6] Strictly speaking, we do not report plant effects as estimated by previous research leveraging the

[3] Some formal workers in the public sector are not in the IMSS database because a separate institution manages their social security. If a worker reports more than one employment in the same workplace, we keep the job with the highest reported wage. Only 2.5% of workers reported having jobs in different workplaces in December 2018.

[4] This variable includes various forms of compensation other than salary (e.g., paid vacations and bonuses) while excluding additional non-taxable compensations.

[5] For 2018, this limit was 2,015 MXN daily, about 102 USD.

[6] The identifiers of the *registro patronal* we use are anonymized. The authors cannot precisely identify individual workers or firms within the dataset. The Mexican Central Bank's EconLab (our data supplier) constructs the masked identifiers before providing the dataset. The use of the anonymized identifiers in place of the original *registros* is inconsequential to our econometric analysis and results.

AKM methodology. In our study, the "workplace" contributor to wage variability is the "*registro patronal* component" of wage variance.

Descriptive statistics. Table 1 presents some IMSS wage data characteristics for selected years. We split our data into three time intervals: 2004-2008, 2009-2013, and 2014-2018. For any given year, our sample includes 73 to 113 million wage observations for men 25-54 years old and 39 to 69 million female wage observations for the same age range. Column (2) of the Table shows that, compared to 2005, the average real daily wage for prime-age men fell by 0.7% from 2009 to 2014, then rose by 1.5% by 2018. These changes were accompanied by an expanding spread of earned wages between 2005 to 2018, as shown in column (3). Women's average real wages increased steadily, from about 326 pesos in 2005 to about 345 pesos a day in 2018. The standard deviation of women's wages also grows over time. These trends allude to a possible intensification of wage inequality in Mexico during the period we analyze. Throughout the rest of the document, we aim to document the roles that average workplace-level remuneration premia, worker-specific traits, and the sorting of workers and workplaces according to their productivity play in determining these trends.

Table 1: Descriptive Statistics: Prime-age Workers, National Level

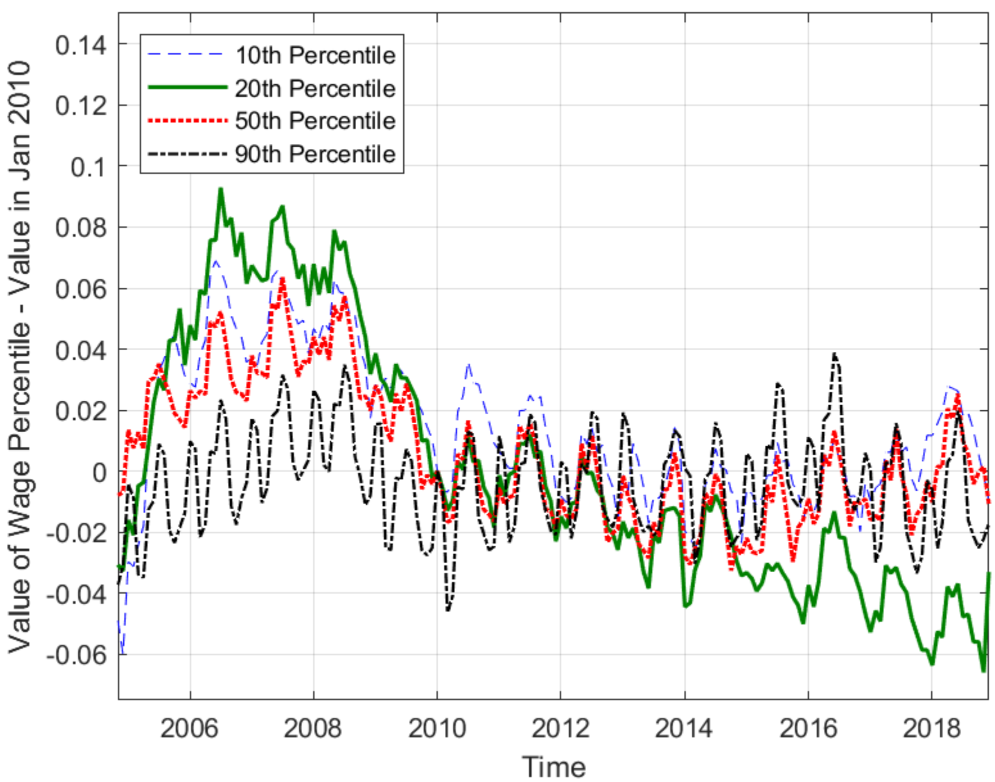
		Real wage		
	(1)	(2)	(3)	(4)
	Observations	Mean	Std. dev	Percent censored
Panel A. Full-time men				
2005	73,855,547	394.575	406.256	2.336
2009	80,069,659	394.594	403.065	2.359
2014	98,566,773	391.698	407.856	2.300
2018	113,516,335	397.765	410.689	2.626
Panel B. Full-time women				
2005	39,579,722	326.635	330.534	0.933
2009	46,347,336	332.782	336.956	1.044
2014	57,801,647	339.536	351.473	1.158
2018	69,681,317	345.607	355.114	1.455

Source: Authors' calculations using IMSS data. Observations correspond to the sum of all the monthly observations in a year. Real wages using prices of July 2018. Percent censored is the percentage of observations with wages exactly equal to the upper wage limit of 25 minimum wages or UMAs.

4 Stylized Facts on Wage Inequality in Mexico

Before outlining the methodology we employ to decompose total wage variance in Mexico's labor market, we describe overall inequality trends. Figure 1 exhibits deviations of percentiles of real daily log-wages from values of the same percentiles in 2010 for males between the ages of 25 and 54 (prime age). From 2004 to 2010, wages fell in real terms in all the percentiles shown except the 90th percentile. From 2010 to 2018, there were further real wage losses at the bottom of the distribution in the 20th percentile, with wage compression in the left tail. The 10th percentile does not decrease as much because of the presence of the minimum wage.

Figure 1: Trends in Percentiles of Log Wages for Prime-age Men



Source: Authors' calculations using IMSS data. The lines depict the values of the 10th, 20th, 50th and 90th percentile of the wages of men 25-54 years old, relative to the values of these percentiles in January of 2010.

Figure 2 presents measures of the spread of real daily wages for prime-aged men. We plot the standard deviation of log wages; and the normalized gaps in log wages between selected percentiles. If log wages followed a normal distribution, every graphed line would overlap. Log wages do not have a normal distribution at the lower and upper tails. However, the low-wage gap in the middle of the distribution does follow a distribution close to normal. Additionally, both the standard deviations and the normalized gaps follow parallel trends.

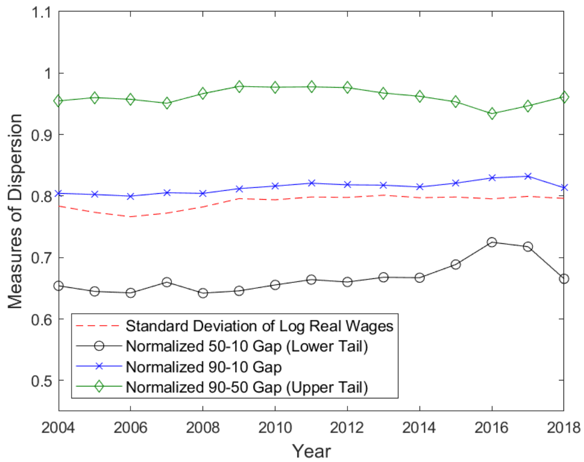
Figure 2: Upper-tail, Lower-tail and Overall Wage Inequality Trends for Prime-age Men, National Level

Source: Authors' calculations using IMSS data. Normalized percentile gaps are differences in percentiles divided by the corresponding differences in percentiles of standard normal variate. For example, the 90th-10th gap is divided by $\Phi^{-1}(0.9) - \Phi^{-1}(0.1)$, where $\Phi(\cdot)$ stands for the standard normal distribution function.

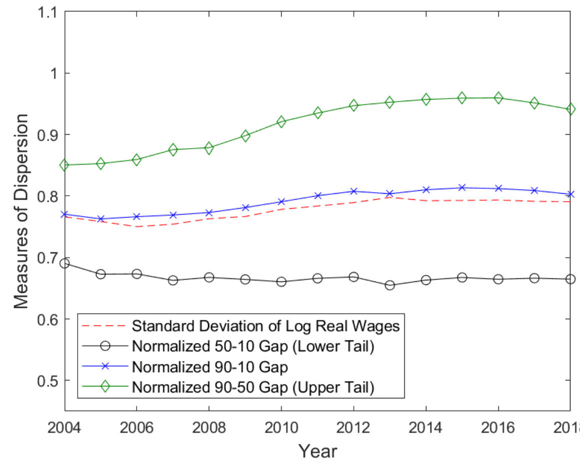
Figure 3 shows similar patterns in all of the country's sub-regions. Overall, wage inequality is higher in the Center and South. The standard deviation of wages is steady for all regions in the sample periods, except for the South. There, it decreases from 2014 to 2018. Lower-tail inequality decreased in the South in the same period.

Figure 3: Upper-tail, Lower-tail and Overall Wage Inequality Trends for Prime-age Men, Regional Level

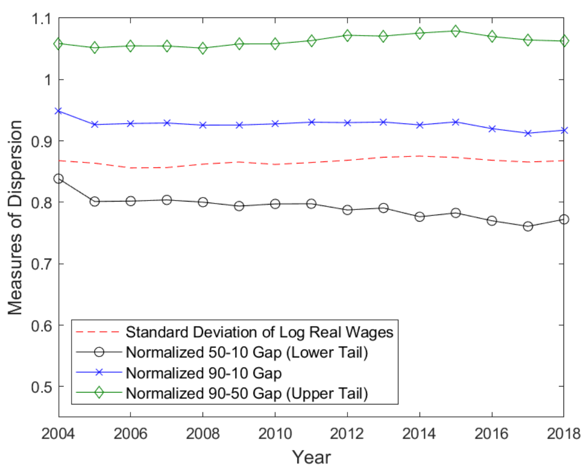
(a) North



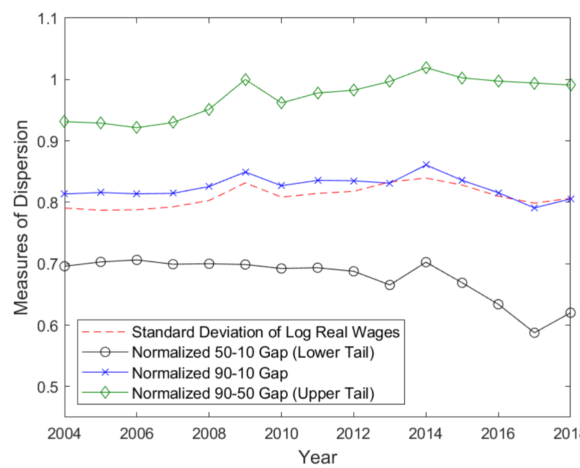
(b) Center-North



(c) Center



(d) South



Source: Authors' calculations using IMSS data. Normalized percentile gaps are differences in percentiles divided by the corresponding differences in percentiles of standard normal variate. For example, the 90th-10th gap is divided by $\Phi^{-1}(0.9) - \Phi^{-1}(0.1)$, where $\Phi(\cdot)$ stands for the standard normal distribution function.

5 Methodology

To isolate the assortativeness, worker- and workplace-specific components of the evolution of wage variability in the Mexican labor market, we follow Card et al. (2013). We begin by adopting the widely embraced econometric approach proposed by Abowd et al. (1999), where log wages are modeled as follows:

$$\ln(\text{wage}_{it}) = \alpha_i + \psi_{\mathbf{J}(i,t)} + X'_{it}\beta + r_{it}. \tag{1}$$

Here, wage_{it} is the real wage of worker i at time t. The worker fixed effects α_i are constant within any given time interval and capture worker-specific skills, abilities, and other characteristics that receive equivalent compensation across firms. Similarly, the workplace effects $\psi_{\mathbf{J}(i,t)}$ capture a similar wage premium or discount that accrues to all workers employed in the same workplace J (Card et al., 2013). The vector $X'_{i,t}$ contains observable characteristics, which in our specification include a time trend, age, age squared, and age cube.[7] We estimate equation (1) by OLS. The identification assumption is that the error term r_{it} is not correlated with the covariates or the worker and workplace dummies. We address this assumption's implications when we talk about job exchangeability below.

We define positive (negative) assortative matching as the positive (negative) correlation between worker and workplace fixed effects as measured by the covariance $Cov(\alpha_i,\ \psi_{\mathbf{J}(i,t)})$; where, by definition, the magnitudes of the worker and workplace effects increase according to their productivity. Assuming complementarity in production between workplaces and workers, the covariance between these two effects will be positive if high-quality workplaces tend to hire highly productive workers, and their remuneration is larger than that of low-productivity workers employed in the same place.

Connected set. The AKM methodology takes advantage of worker mobility across workplaces. We can only identify worker and workplace fixed effects within a "connected set" of workplaces united by a shared pool of workers with the common characteristic of having changed jobs at least once (Abowd et al., 1999). One workplace belongs to the connected set if at least one of their workers has worked or will work in a different plant at a given time interval. Note

[7] We do not include time effects since they would be highly collinear with the linear age effect (Dauth et al., 2022).

that every workplace pair does not need to be connected directly for a connected set to exist. We restrict our analysis to the largest connected set in each time interval.

To ease the comparison of our estimates to previous studies, the analysis in this section discusses estimations for men aged 25 to 54 (prime-age). For each one of the four time periods, columns (1) to (4) of Table 2 show the number of worker-year observations for prime-age males that had more than one job, the number of individuals, and the average and standard deviation of log wages. In each interval, our database has between 158 and 297 million worker-year observations corresponding to 5 to 9 million individuals. The standard deviation of wages rose from 0.77 in the 2004-2008 interval to 0.79 in 2014-2018. Average real wages have increased throughout the sample. Columns (5) to (8) of table 2 show the corresponding descriptive statistics for the largest connected set of prime-age male workers. The largest connected set contains at least 94% of all worker-year observations and 97% of all individuals in a given interval. Average wages in the connected set are slightly higher than in the overall sample, while standard deviations are marginally smaller. The large size of the connected set relative to the entire sample; the comparable mean salaries, standard deviations, and the similar trends of the average wage and salary dispersion imply that we lose little by directing our attention to said connected group.

Exchangeability. Card et al. (2013) show that if the residual term in (1) is uncorrelated with the right-hand-side variables, then, on average, a worker that moves from workplace A to workplace B should experience a wage change of the opposite sign to that experienced from a worker moving in the opposite direction. Following Card et al. (2013), Figure 4 shows an event study to examine whether this holds in our dataset. The plot presents the average wages of workers who changed jobs for each time interval in our analysis period. Workers may move from "low-wage" to "high-wage" workplaces or *vice versa*. We classify workplaces based on the quartile of the average co-worker wage in their initial job and the corresponding quartile for their final job. We then compute average wages in the years before and after the job exchange for each cell.[8]

The Figure shows that different mobility groups classified by average co-worker wage have, on average, different wage levels before and after a move. For job-changers moving down the quartile classification, before a move, average wages in the quartile of origin vary monotonically

[8] We exclude observations from establishments with only one worker. We keep only "direct" moves, that is, moves without an unemployment spell in the transition between jobs.

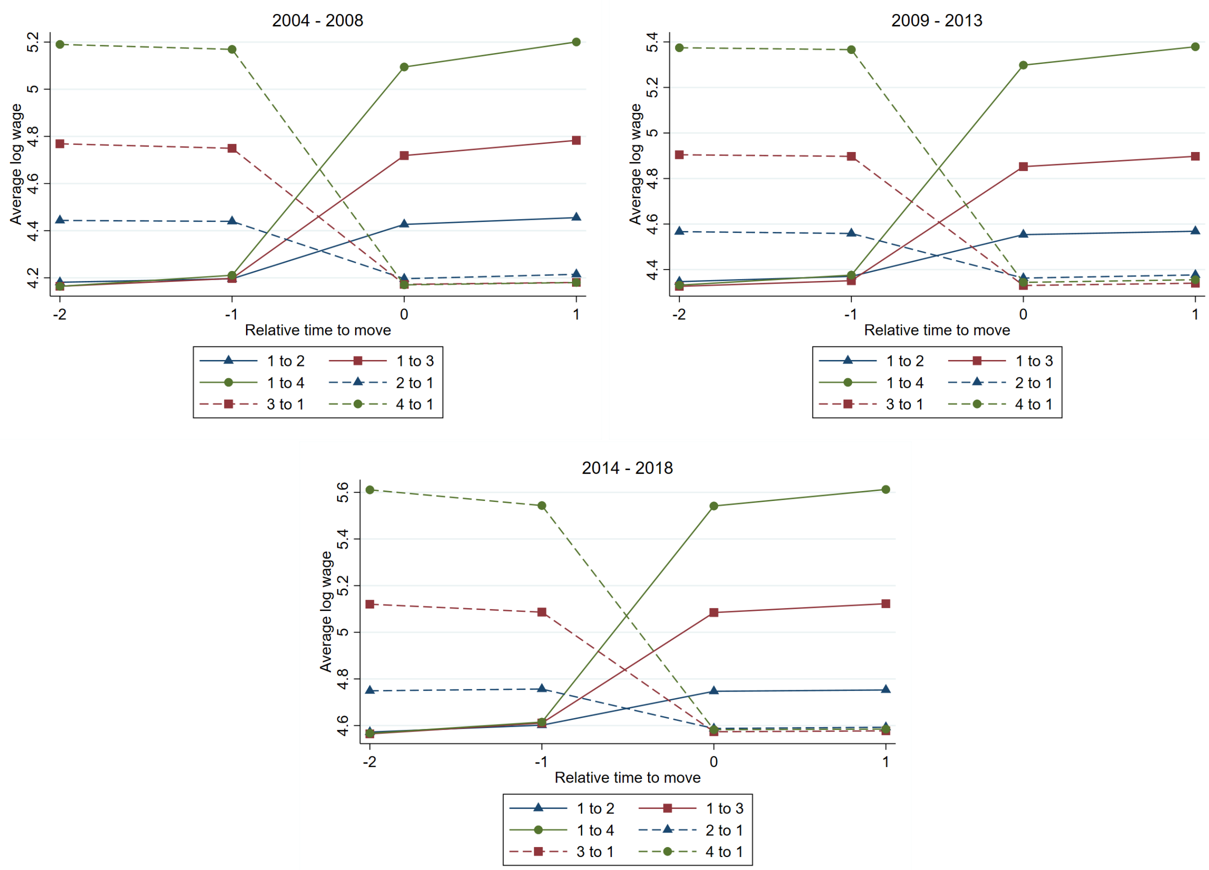
Table 2: Descriptive Statistics - Overall Sample and Workers in the Largest Connected Set

	All sample				Individuals in largest connected set			
			Log wage				Log wage	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Interval	All obs.	Individuals	Mean	Std. dev.	All obs.	Individuals	Mean	Std. dev.
Nov 2004-2008	158,543,931	5,721,179	5.525	0.772	150,458,370	5,576,345	5.556	0.772
Ratio: largest connected/all					94.91	97.51	100.61	100.01
2009-2013	226,528,652	7,072,043	5.487	0.791	216,360,702	6,920,461	5.515	0.792
Ratio: largest connected/all					95.51	97.91	100.51	100.11
2014-2018	297,395,413	9,069,558	5.488	0.793	288,394,833	8,941,908	5.507	0.793
Ratio: largest connected/all					97.01	98.61	100.31	100.01
Change from first to last interval			-0.0371	0.0211			-0.0491	0.0211

Source: Authors' calculations using IMSS microdata. Statistics for men 25 to 54 years old who had more than one job, i.e. were employed in more than one workplace. Log wage is the log of daily taxable income registered in IMSS, expressed in real terms using prices from July 2018. "Ratio: largest connected/all" is the ratio of the corresponding statistic in the largest connected set to its counterpart in the full sample.

with respect to the destination quartile. For example, average wages for workers moving from quartile 4 (the highest average co-worker salary) to quartile 1 (the lowest mean co-worker wage) are higher before the job switch than for those who go from quartile 3 to 1, and so on. Additionally, the magnitude of the absolute change in average wages when moving from one quartile to another is equivalent to the variation associated with the opposite change. Such symmetry is consistent with an additive model for wages with worker and workplace fixed effects such as the one we estimate.

Figure 4: Exchangeability: Average Log Wage Around Movement by Quartile of Average co-workers' Wages in the Origin and Destination Workplace



Source: Authors' calculations using IMSS data. The graph shows average worker wages for workers who move between an origin workplace to destination workplace, from two months before the move to one month after the move. The lines group workers according to the quartiles of average co-worker wages in the origin and destination workplaces. The panels correspond to different time intervals. We exclude observations from establishments with only one worker. We keep only "direct" moves without an unemployment spell in the transition between jobs.

Variance decomposition. Following Card et al. (2013), under the assumption that the error term and the covariates in equation (1) are uncorrelated, the variance of log wages in a given period can be decomposed as:

$$\begin{aligned}\operatorname{Var}(\ln \text{wage}_{it}) = & \underbrace{\operatorname{Var}(\alpha_i)}_{\text{workers}} + \underbrace{\operatorname{Var}(\psi_{\mathbf{J}(i,t)})}_{\text{workplaces}} + \operatorname{Var}(x'_{it}\beta) + \operatorname{Var}(r_{it}) \\ & + 2\underbrace{\operatorname{Cov}(\alpha_i, \psi_{\mathbf{J}(i,t)})}_{\text{sorting}} + 2\operatorname{Cov}(\psi_{\mathbf{J}(i,t)}, x'_{it}\beta) + 2\operatorname{Cov}(\alpha_i, x'_{it}\beta).\end{aligned} \tag{2}$$

The first term corresponds to the variance of log wages explained by time-invariant worker characteristics, while the second term corresponds to the contribution of workplace differences to wage inequality. The sorting term measures the contribution of assortative matching to wage variance.

We estimate the model in equation (1) by OLS with a pre-conditioned iterative gradient method. To compute the decomposition in equation (2), we replace the parameters with their OLS estimates and compute the sample analogs of each variance and covariance term.

6 Contributors to Wage Inequality in Mexico 2004-2018

In this section, we show estimates of the AKM model in (1) for the entire Mexican labor market. We first show a summary of the estimated model and argue that it explains a large share of the variance of wages of formal workers. Then, we highlight the increasing role of assortative matching in explaining wage inequality in Mexico. Last, we compare our estimates to those from other countries.

Table 3 shows a summary of the estimated model for each time interval: 2004-2008, 2009-2013, and 2014-2018. Our models include from 5.5 to 8.9 million worker effects and 5.2 to 6.9 million workplace effects each period. We report the standard deviations of the estimated workplace and worker effects and their correlation. We also report the model's root mean squared error (RMSE) and its adjusted R^2. The estimated model has high explanatory power, with high values of the adjusted R^2 in each interval.

The results in Table 3 show several patterns of interest. First, consider how the dispersion of worker and workplace effects follow opposing trends: the standard deviation of worker effects

Table 3: AKM Model Estimation Results

	Interval1 2004-2008	Interval2 2009-2013	Interval3 2014-2018
Worker and workplace parameters			
Number of worker effects	5,576,345	6,920,461	8,941,908
Number of workplace effects	523,701	554,593	695,749
Summary of parameter estimates			
St. dev. of worker effects	0.504	0.486	0.472
St. dev. of workplace effects	0.444	0.479	0.487
Correlation worker/workplace effects	0.212	0.231	0.259
Correlation worker effects/Xb	-0.123	-0.074	-0.104
Correlation workplace effects/Xb	-0.051	-0.045	-0.048
Goodness of fit			
St. dev. of log wages	0.772	0.792	0.793
RMSE	0.238	0.237	0.233
R Squared	0.909	0.913	0.916
Adj. R Squared	0.905	0.910	0.913

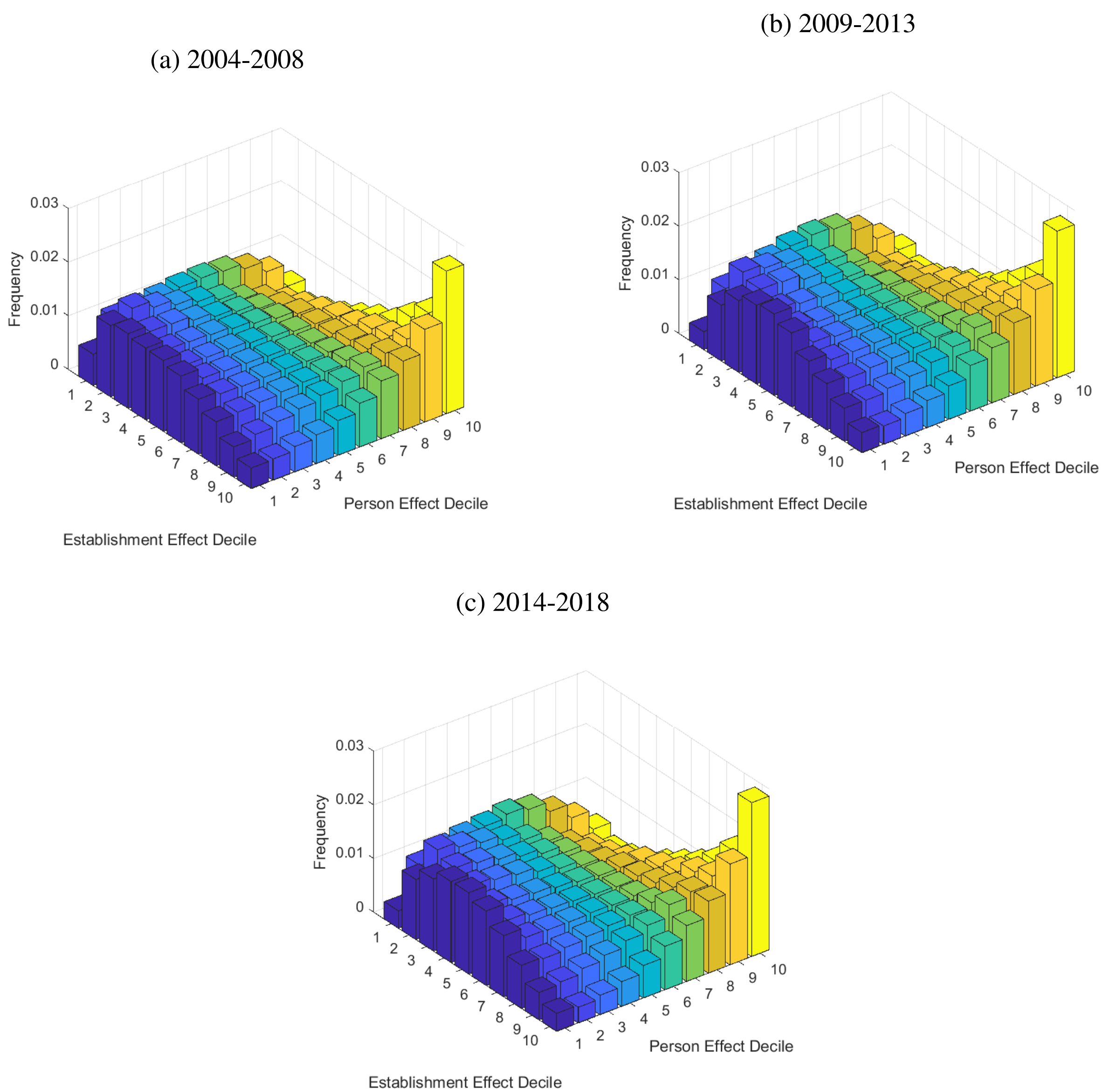
Source: Authors' calculations using IMSS data. Results from estimation of equation (1) via OLS. Observations correspond to largest connected set per time interval. "Xb" stands for covariates and includes the following controls: age, age squared, age cube, and a monthly time trend.

decreases over time while the dispersion of workplace effects increases. These patterns suggest that workplace-specific effects played an increasingly important part in propelling wage inequality trends in Mexico.

Additionally, the correlation between worker and workplace effects grows over time, which hints at an increasing influence of positive assortative matching over earnings inequality. Figure 5 offers visual evidence of this trend. We plot the joint distributions of the estimated worker and workplace effects (grouped by deciles) for 2004-2008, 2009-2013, and 2014-2018, classifying the fixed effects by deciles. Comparing figure 5's sub-figures makes clear the secular tendency for higher-wage workers to sort to workplaces with more significant wage premia.

We suspect that the democratization of the internet and the more common use of online job platforms may be drivers of the increased sorting. Starting in 2013, Mexico experienced a dramatic expansion in high-speed internet access. Between 2013 and 2020, the coverage of broadband telecommunications expanded by 227.2%, growing from 23 to 77 lines per 100 persons: the starkest increase in coverage among OECD members (Instituto Federal de Telecomunicaciones, 2021).

Figure 5: Changes in Assortative Matching: Joint Densities of Workplace and Worker Effects. National Level



Source: Authors' calculations using IMSS data. Panels depict the joint distribution of estimated worker and workplace effects from equation (1) by deciles of the marginal worker and fixed effect distributions.

Similarly, the use of job-matching platforms has expanded significantly. The proportion of job-seekers that report preferring to look for a position online grew from 71% in 2014 to 95% in 2018 (Asociación de Internet MX, 2014; Asociación de Internet MX, 2018). Along with the increased use of online job-search platforms by workers, during the same period, there has been a parallel expansion in the number of websites offering job-searching services (Asociación de Internet MX, 2018).

6.1 Decomposing Wage Variance

We now present estimates of contributions made by these two components to total wage variance. To quantify the individual contributions of worker effects, workplace effects, and their sorting, we conduct a variance decomposition analysis based on equation 2 in each period considered.

As we noted when commenting on the results from Table 3, the dispersion of worker and workplace effects trend in opposite directions. At the same time, the correlation between these factors increases over time. Table 4 shows how these opposing trends contributed to the increase in wage inequality in Mexico from 2004 to 2018. Worker effects went from accounting for a 42% share of the wage variance of prime-age workers' wage variance in 2004-2008 to less than 36% of their variance in 2014-2018. This decrease happened as the variance of wages increased by about 5%. In contrast, workplace effects account for a 4.7 percentage points (p.p.) higher share of variance in the last period compared to the initial period. Simultaneously, the variance share from the covariance of worker and workplace effects increased by 3 p.p.

The last rows of Table 4 show a counterfactual calculation following Card et al. (2013). For these counterfactuals, we keep the correlation of worker and workplace effects and the variance of workplace effects at their 2004-2008 levels and calculate the implied variance of wages for 2009-2013 and 2014-2018. These are scenarios where matching technologies do not improve over time, and the wage-setting power of workplaces keep remains constant. Without the increase in the importance of workplace effects and assortative matching in determining wages, the variance of wages would be 10% smaller in 2014-2018.

Card et al. (2013) argue that in the absence of an increase in the importance of workplaces and assortative matching, Germany's wage variance would have been about 25% lower in 2002-2009. We find that the increase in the importance of these factors in Mexico has been smaller.

Nevertheless, the importance of workplaces in wage inequality in Mexico is substantially larger. Workplace effects are more consequential to the evolution of worker-workplace sorting in Mexico, unlike most national labor markets where the AKM methodology has been applied. The high share of variance attributed to workplace effects is consistent with previous work utiliz-

Table 4: Wage Variance Decomposition, National Level

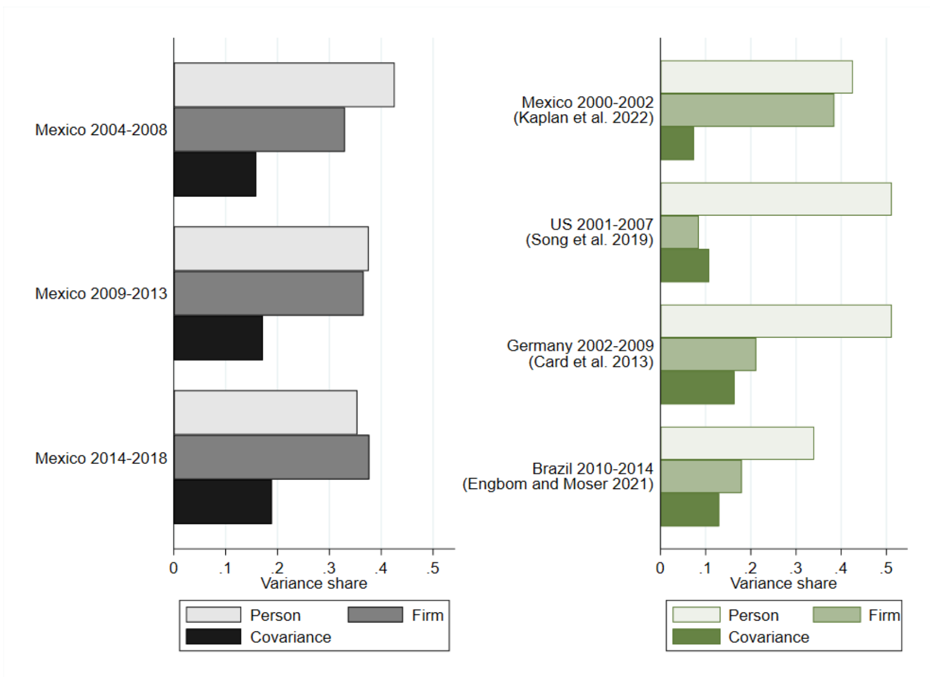
	Interval 1 2004-2008	Interval 2 2009-2013	Interval 3 2014-2018	Change from int. 1 to 3
Variance and covariance				
Total variance of log wages	0.596	0.627	0.628	0.032
Variance of worker effects	0.254	0.236	0.222	-0.032
Variance of workplace effects	0.197	0.230	0.237	0.040
Variance of covariates (Xb)	0.019	0.013	0.016	-0.004
Variance of residual	0.055	0.054	0.053	-0.002
2 Cov(worker effects, workplace effects)	0.095	0.108	0.119	0.024
2 Cov(worker effects, covariates)	-0.017	-0.008	-0.012	0.005
2 Cov(workplace effects, covariates)	-0.006	-0.005	-0.006	0.000
Variance shares				
Variance of worker effects	0.426	0.376	0.354	-0.073
Variance of workplace effects	0.330	0.366	0.377	0.047
Variance of covariates (Xb)	0.032	0.020	0.025	-0.006
Variance of residual	0.091	0.087	0.084	-0.008
2 Cov(worker effects, workplace effects)	0.159	0.172	0.189	0.030
2 Cov(worker effects, covariates)	-0.029	-0.013	-0.019	0.009
2 Cov(workplace effects, covariates)	-0.011	-0.008	-0.009	0.001
Counterfactuals for variance of log wages				
1. No rise in correl. of worker/firm effects	0.596	0.618	0.608	
2. No rise in var. of workplace effects	0.596	0.587	0.578	
3. Both 1 and 2	0.596	0.585	0.568	

Source: Authors' calculations using IMSS data. Variance decomposition results from equation (2) using the estimated worker and workplace fixed effects from equation (1). The "Variance and covariance" rows show the values of the variance of log wages in the estimation sample of prime-age men and its components. The "Variance shares" rows show the share of the overall variance in log wages in the sample attributed to each one of its components. The first three columns correspond to time intervals, and the last columns is the change from 2004-2008 to 2014-2018. The "Counterfactuals for variance of log wages" rows show the variance of wages assuming that the correlation of worker/workplace effects and the variance of workplace effects had remained constant at 2004-2008 values.

ing worker-workplace longitudinal data from Mexico before 2002 (Frías et al., 2022), and with research pointing out an increase in inequality across as opposed to within workplaces (Song et al., 2018). Figure 6 illustrates this difference. The left panel displays our estimates for the contributing shares of worker and workplace effects to total wage variance in Mexico for the considered intervals. The right panel presents equivalent estimations from previous work studying Mexico (Frías et al., 2022), the United States (Song et al., 2018), Germany (Card et al., 2013), and Brazil (Engbom and Moser, 2021). In the Mexican labor economy, both worker and workplace effects contribute equally to trends in wage inequality. Intriguingly, the command that workplaces have to determine wage differentials has increased at the same time that labor unions have lost strength and representation capacities in Mexico. In particular, according to Mexico's Ministry of Labor, the proportion of salaried workers that belong to a union diminished from 17% to 12% between 2005 and 2018 (Secretaría del Trabajo y Previsión Social, 2022).

On the other hand, the contribution of sorting (as measured by the covariance between the two effects) is roughly comparable to the shares estimated for other countries.[9]

Figure 6: Comparing Estimated Worker and Workplace Contributions to Wage Variance



Source: Authors' calculations from IMSS data, and reported values from Frías et al. (2022), Song et al. (2018), Card et al. (2013). and Engbom and Moser (2021). The left panel shows variance shares attributed to worker effects, workplace effects and their covariance in each time period from Table 4. The right panel shows equivalent variance shares for different countries from different studies.

[9]The share of variance attributed to workplace effects in Mexico is also more extensive than that of other OECD countries: see OECD (2021).

6.2 Regional Differences

We now examine how wage differences across workers, workplaces, and assortative matching –as estimated from our AKM model– contribute to wage inequality in Mexican regions. We apply the decomposition of equation (2) to the variance of wages in our estimated model sample, dividing the sample into regions.[10]

Figure 7 shows how the worker and workplace effects and their correlation contributed to wage spread in each of the four sub-national regions. In all four regions, assortative matching explains between 13% and 21% of the variance in wages. The strong ability of workplaces to influence wage disparities is also present in all four Mexican geographical regions.

The contribution of workplace-specific effects to overall wage variance correlates negatively with the level of regional development. Workplace fixed effects are relatively more important in determining wage variance in the South, followed by the Center-North, Center, and last by the northern region. The contribution of worker effects follows precisely the opposite pattern. These motifs resemble local levels of general economic development: historically, Northern and Southern Mexico have been the country's most and least economically mature regions, respectively.

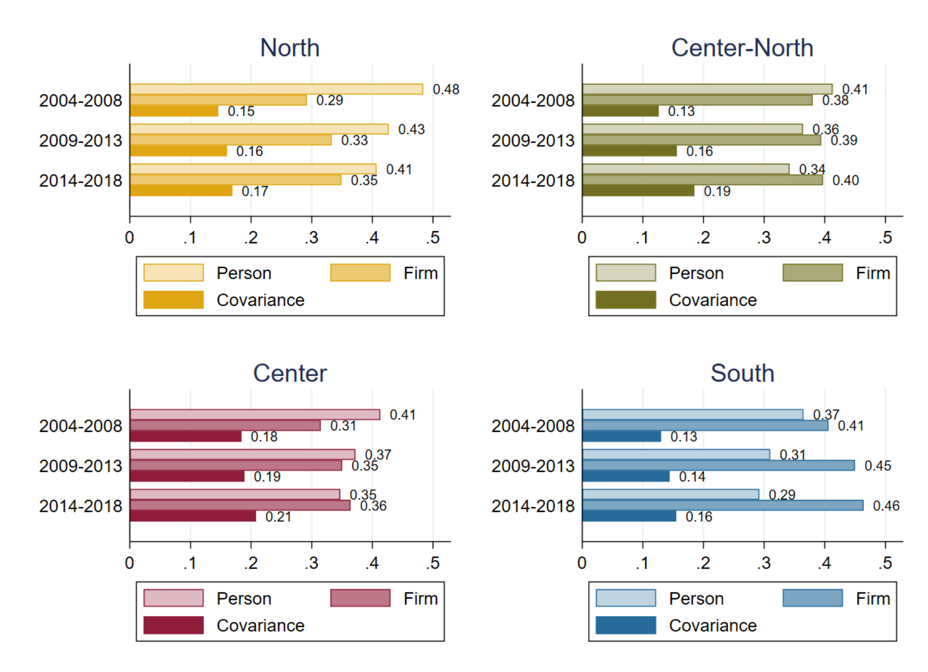
We now highlight differences in assortative matching across regions. In Figure 8, we compare the joint distribution of worker and workplace fixed effects across regions in 2014-2018. While in the Center, about 4% of workers are in the top (northern region) decile of worker-specific wage premia and work in top-decile establishments, in the South, this number is less than 2%. It does not differ much from the fractions of workplaces across worker fixed effect deciles in the bottom establishments. The North and Center-North also display stronger assortative matching patterns than the South, but they are still not as visible as those in the Center.

A possible mechanism behind these differences in the importance of workplaces in wage inequality and assortative matching is the disparity in educational attainment across regions. The South region has substantial lags in educational attainment relative to the rest of the country.[11] Lack of education and differences in education would lead to a lower variance of worker fixed

[10]Strictly speaking, since we do not re-estimate the model per region, equation (2) may not hold exactly by region because the OLS residual may correlate with covariates in each regional sub-sample. Nevertheless, the share of variance attributed to this correlation is negligible.

[11]As an example, in 2016, 33.6 and 39.7 % of the population ages 25 to 34 in Oaxaca and Chiapas (Southern region states), respectively, had not completed primary education. This contrasts with only 7.8% in Mexico City (Center) and 10.7 % in Nuevo León (North) (INEE, 2018)

Figure 7: Estimated Worker and Workplace Contributions to Wage Variability by Region



Source: Authors' calculations using IMSS data. The figure depicts the variance shares attributed to worker fixed effects, workplace fixed effects, and their covariance in the overall variance of wages in each region, using the estimates of the AKM model from equation (1) and the decomposition in equation (2).

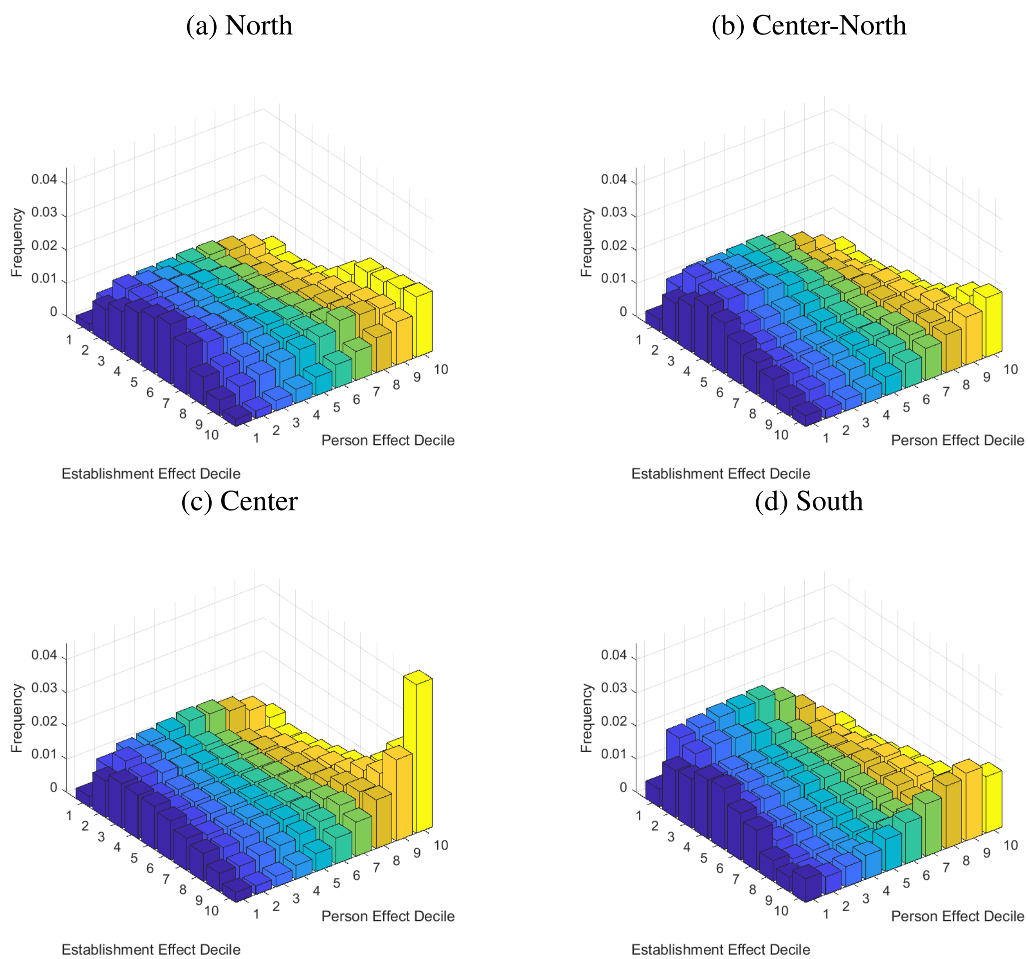
effects (which include worker educational wage premia) and a larger wage-setting capacity by workplaces.

6.3 Additional Evidence and Robustness

We now summarize additional estimation exercises to probe the robustness of our results. We use different specifications of the AKM model in equation (1), estimate the model for women and all workers, and estimate variance shares using a correction for limited mobility bias. Across all these exercises, we still see a large share of variance attributed to workplaces in Mexico and an increasing importance of assortative matching in explaining wage inequality.

Alternative model specifications. In Appendix Figure A.2, we calculate the shares of variance attributed to workers, workplaces, and assortative matching with different model specifications: excluding time trends, excluding top-coded observations, including time trends interacted with sector indicators, including controls for workplace size, and a quartic polynomial in age (Lemieux,

Figure 8: Regional Differences in Assortative Matching



Source: Authors' calculations using IMSS data. Panels depict the joint distribution of estimated worker and workplace effects from equation (1) by deciles of the marginal worker and fixed effect distributions in each region.

2006). Across all specifications, we still find that workplaces account for a large share of variance and that assortative matching is becoming increasingly important over time.

Variance decomposition for additional periods. We repeat the estimation on the prime-age men sample for every 4-year window starting in December 2004 - December 2008 and ending in December 2014 - December 2018. We plot the variance shares attributed to worker effect, workplace effects and their covariance in each period in Appendix figure A.3 A.3. The trends confirm the patterns found in table 4: the relevance of firm effects and assortative matching in explaining the variance of wages is increasing over time, while worker effects are losing importance.

Comparing men, women, and all workers. The inequality trends we document are slightly different for women. In Appendix figure A.1, we show that wages have increased from 2010 to 2018 in the 10th, 50th, and 90th percentiles of women's wage distributions.

Our AKM models are also adequate in explaining wages for women and the entire sample. In Appendix Table A.1, we show estimates of the AKM model for men, women, and all workers ages 25 to 54. The additive effects model explains a high share of the variance of log wages for women and the overall sample. For all samples, we see an increasing variance of workplace effects over time and a decreasing variance of worker effects. The correlation of worker and establishment effects is slightly larger for men in all periods.

Our findings regarding the importance of workplaces also hold for women's wages. Appendix Tables A.2 and A.3 show the variance decomposition results in equation (2) for the women and all workers samples. For women, workplace effects and the correlation of workplace and worker effects explain an increasing share of variance over time, similar to our results for men in Table 4. Workplace effects explain a lower share of the variance of overall wages for women and do not overtake workplace effects as the largest component of wage variance in 2014-2018. Nevertheless, the variance of wages for women would also be about 8% lower in 2014-2018 if the workplace and matching components had not increased their importance. The picture is similar in the sample with all workers ages 25-54.

Limited mobility bias. Andrews et al. (2012), Kline et al. (2020), and Bonhomme et al. (2022) show that there may be substantial bias in estimates of variance shares in AKM models like the one we estimate. These biases arise in settings with low worker mobility across workplaces, such that the estimate of the variance components in equation (2) has a large small-sample bias. We address this concern by re-estimating the variance decomposition in Table 4 with a corrected leave-one-out variance estimator following Kline et al. (2020). Appendix Table A.4 shows the results. Our corrected estimates of the variance components of log wages are virtually equal to those of Table 4.[12]

Variance decomposition across sectors. Table A.5 in the appendix, shows a decomposition

[12]Our relatively unchanged estimates contrast with those of Frías et al. (2022) and Engbom and Moser (2021), who find that their estimates have meaningful changes once they implement their limited-mobility-bias corrections. There are two reasons why our estimates do not change as much: First, our connected set is a large share of the entire sample and thus, we expect mobility in the connected set to remain comparable to that of the overall sample. Second, our time intervals are wider than those in Frías et al. (2022), allowing for more worker mobility in each time interval. Note that if limited mobility bias is relatively constant over time, our results in Table 4 regarding the evolution of each variance component will hold even without the bias correction.

of the wage variance across *sectors*.[13] The main patterns remain essentially unchanged. The dispersion of mean log wages expands over time at the same time as the estimated contribution of worker-specific characteristics declines. The role of assortative matching increases across all three time intervals considered.

7 Conclusion

We quantify the proportion of observed wage inequality in Mexico attributed to worker-specific characteristics, average workplace-level salary premia, and assortative matching. We use a large, matched worker-workplace database with the near universe of private-sector workers in Mexico with wage information spanning the period between 2004 and 2018. To decompose total wage variance, we leverage estimations from AKM-style models of log wages with two-way fixed effects. We observe that assortative matching plays an increasingly important part in shaping wage inequality in Mexico. In agreement with previous work looking at other developing countries, workplace-specific salary premia contribute significantly to wage inequality in the country. Interestingly, the proportion by which workplaces explain wage discrepancies is the largest (smallest) in the southern (northern) region. The workplace-specific contribution to inequality moves along regional levels of economic prosperity, being the largest (smallest) in the South (North) and historically the least (most) affluent Mexican geographical region.

Interesting avenues of research remain open for researchers wishing to expand on our work. Notably, starting in 2019, there has been a flurry of economic reforms that could directly impact the ability of workplaces to set wages. Examples include the reform of the former North-American Free Trade Agreement; the Mexican labor reform, which modified collective agreement regulations and altered formal labor dispute procedures; and, starting in 2021, the government advanced proposals to regulate the outsourcing of labor.

[13] To perform our calculations, we rely on the sector classification in the IMSS data, which can be mapped to a 3-digit NAICS classification.

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–333.

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Alcaraz, C., Chiquiar, D., and Salcedo, A. (2015). Informality and Segmentation in the Mexican Labor Market. *Banco de México Working Papers*, No. 2015-25.

Alvarez, J., Benguria, F., Engbom, N., and Moser, C. (2018). Firms and the Decline in Earnings Inequality in Brazil. *American Economic Journal: Macroeconomics*, 10(1):149–89.

Andrews, M. J., Gill, L., Schank, T., and Upward, R. (2012). High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias. *Economics Letters*, 117(3):824–827.

Asociación de Internet MX (2014). Estudio sobre los Hábitos de los Usuarios de Internet en México 2014.

Asociación de Internet MX (2018). Estudio sobre los Hábitos de los Usuarios de Internet en México 2018.

Banco de México (2011). Reporte sobre las Economías Regionales, Enero-Marzo 2011.

Bassier, I. (2022). Firms and Inequality When Unemployment is High. *Available at SSRN 4091273*.

Boeri, T., Ichino, A., Moretti, E., and Posch, J. (2021). Wage Equalization and Regional Misallocation: Evidence from Italian and German Provinces. *Journal of the European Economic Association*, 19(6):3249–3292.

Bonhomme, S., Holzheu, K., Lamadon, T., Manresa, E., Mogstad, M., and Setzler, B. (2022). How Much Should we Trust Estimates of Firm Effects and Worker Sorting? *Journal of Labor Economics*, Forthcoming.

Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics*, 36(S1):S13–S70.

Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.

Chávez-Martín del Campo, J. C. and García Loredo, K. (2015). Identificación de Clusters Regionales en la Industria Manufacturera Mexicana. *Banco de México Working Papers*, No. 2015-19.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial Wage Disparities: Sorting Matters! *Journal of Urban Economics*, 63(2):723–742.

Dauth, W., Findeisen, S., Moretti, E., and Suedekum, J. (2022). Matching in Cities. *Journal of the European Economic Association*, 20(4):1478–1521.

Delajara, M. (2011). Comovimiento Regional del Empleo Durante el Ciclo Económico en México. *El Trimestre Económico*, 78(311):613–642.

Delajara, M. (2013). Comovimiento y Concordancia Cíclica del Empleo en los Estados Mexicanos. *Economía Mexicana. Nueva época*, 22(2):297–340.

Engbom, N. and Moser, C. (2021). Earnings Inequality and the Minimum Wage: Evidence from Brazil. Working Paper 28831, National Bureau of Economic Research.

Frías, J. A., Kaplan, D. S., Verhoogen, E., and Alfaro-Serrano, D. (2022). Exports and Wage Premia: Evidence from Mexican Employer-Employee Data. *The Review of Economics and Statistics*, pages 1–45.

Gerard, F., Lagos, L., Severnini, E., and Card, D. (2021). Assortative Matching or Exclusionary Hiring? The Impact of Employment and Pay Policies on Racial Wage Differences in Brazil. *American Economic Review*, 111(10):3418–57.

Goldin, C. and Katz, L. F. (2010). *The Race Between Education and Technology*. Harvard University Press.

Gruetter, M. and Lalive, R. (2009). The Importance of Firms in Wage Determination. *Labour Economics*, 16(2):149–160.

Instituto Federal de Telecomunicaciones (2021). México entre los 3 Países con Mayor Crecimiento Anual en la Penetración de Banda Ancha Fija: OCDE.

Instituto Nacional de Estadística y Geografía (INEGI) (2018). Indicadores de Ocupación y Empleo. Cifras Oportunas durante Octubre de 2018.

Instituto Nacional para la Evaluación de la Educación (INEE) (2018). Panorama Educativo de México: Indicadores del Sistema Educativo Nacional.

Juárez-Torres, M., Puigvert-Angulo, J., and Zazueta-Borboa, F. (2022). The Role of Clusters in the Performance of the Mexican Economy. *Banco de México Working Papers*, No. 2022-06.

Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy*, 101(3):410–442.

Katz, L. F. et al. (1999). Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics*, volume 3, pages 1463–1555. Elsevier.

Katz, L. F. and Murphy, K. M. (1992). Changes in Relative Wages, 1963–1987: Supply and Demand Factors. *The Quarterly Journal of Economics*, 107(1):35–78.

Kline, P. and Moretti, E. (2014). People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs. *Annual Review of Economics*, 6(1):629–662.

Kline, P., Saggio, R., and Sølvsten, M. (2020). Leave-out Estimation of Variance Components. *Econometrica*, 88(5):1859–1898.

Krueger, A. B. and Summers, L. H. (1988). Efficiency Wages and the Inter-Industry Wage Structure. *Econometrica*, 56(2):259–293.

Lemieux, T. (2006). The "Mincer Equation" Thirty Years After Schooling, Experience, and Earnings. In *Jacob Mincer: A Pioneer of Modern Labor Economics*, pages 127–145. Springer.

Lopes de Melo, R. (2018). Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence. *Journal of Political Economy*, 126(1):313–346.

OECD (2021). *The Role of Firms in Wage Inequality*. OECD Publishing.

Puggioni, D., Calderon, M., Zurita, A. C., Bujanda, L. F., Gonzalez, J. A., and Jaume, D. (2022). Inequality, Income Dynamics, and Worker Transitions: The Case of Mexico. *Quantitative Economics*.

Rangel González, E. and Llamosas-Rosas, I. (2021). Observando la Evolución del Sector Informal desde el Espacio: Un Enfoque Municipal 2013-2020. *Banco de México Working Papers*, No. 2021-18.

Rice, P., Venables, A. J., and Patacchini, E. (2006). Spatial Determinants of Productivity: Analysis for the Regions of Great Britain. *Regional Science and Urban Economics*, 36(6):727–752.

Secretaría del Trabajo y Previsión Social (2022). Comunicado: Se Registran las Tasas de Sindicalización más Altas en los Últimos Años. Retrieved from http://www.gob.mx.

Song, J., Price, D. J., Guvenen, F., Bloom, N., and von Wachter, T. (2018). Firming Up Inequality. *The Quarterly Journal of Economics*, 134(1):1–50.

Torres, S., Portugal, P., Addison, J. T., and Guimarães, P. (2018). The Sources of Wage Variation and the Direction of Assortative Matching: Evidence from a Three-Way High-Dimensional Fixed Effects Regression Model. *Labour Economics*, 54:47–60.

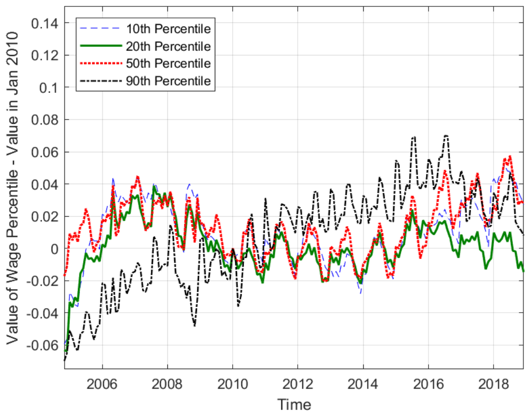
Van Reenen, J. (1996). The Creation and Capture of Rents: Wages and Innovation in a Panel of U. K. Companies. *The Quarterly Journal of Economics*, 111(1):195–226.

Online Appendix - Not for Publication

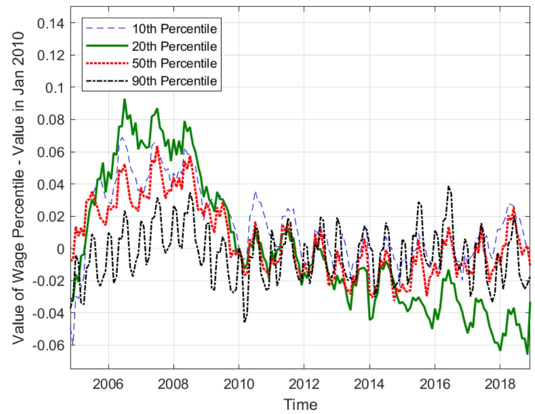
A Additional Figures and Tables

Figure A.1: Trends in Percentiles of Log Wages for Men, Women, and all Workers Ages 25-54

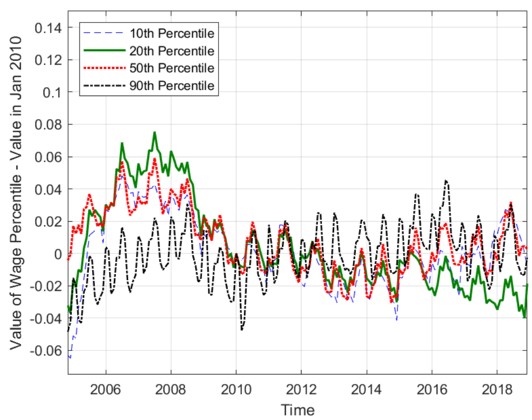
(a) Women



(b) Men



(c) All

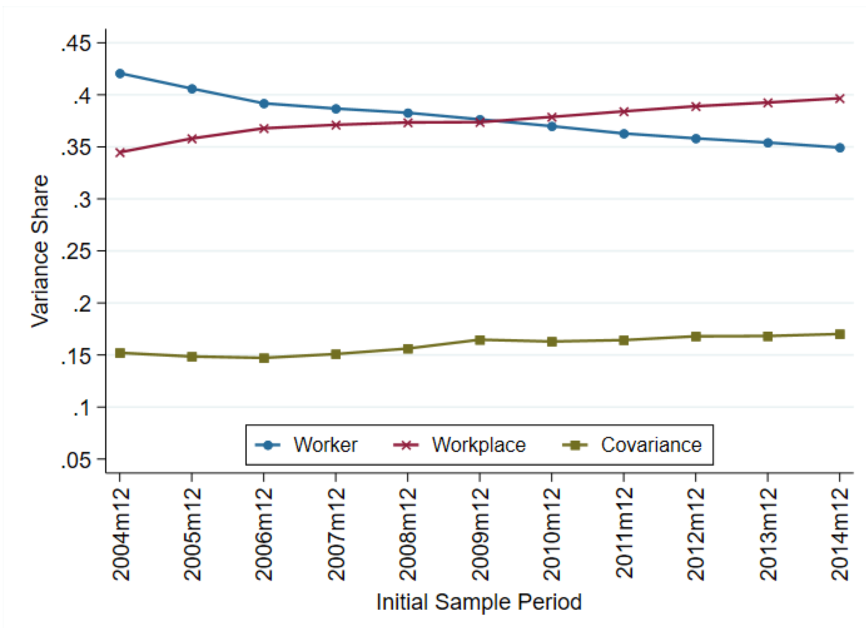


Source: Authors' calculations using IMSS data. The lines depict the values of the 10th, 20th, 50th and 90th percentile of the wages of workers 25-54 years old, relative to the values of these percentiles in January of 2010.

Figure A.2: Variance Shares Comparison Across Model Specifications

Source: Authors' calculations using IMSS data. The panels depict variance shares from variance decomposition results using equation (2). Each panel corresponds to a different model specification. Panel "Base" corresponds to the baseline estimates in Table 4, where the control set includes age, age squared, age cube and a time trend. Panel "No Time Trend" excludes the linear time trend from the control set. Panel "Exclude Topcoded" excludes top-coded observations. Panel "Time Trends by Sector" includes interactions of sector indicators ("actividad" in IMSS data) and a linear time trend. Panel "workplace Size Controls" includes a control for workplace size. Panel "Quartic in Age" includes Age to the fourth power as a control. The rows in each panel correspond to time intervals.

Figure A.3: Variance Decomposition for Additional 4-year windows



Source: Authors' calculations using IMSS data. The lines are depict variance shares from variance decomposition results using equation (2). Each time point corresponds to an estimation using a 4-year period starting in the given month.

Table A.1: AKM Model Summary: Women, Men, and All Workers Age 25-54

	Interval1 2004-2008	Interval2 2009-2013	Interval3 2014-2018
Panel A: Women			
Worker and workplace parameters			
Number of worker effects	2,600,846	3,463,498	4,824,406
Number of workplace effects	286,383	346,557	458,888
Summary of parameter estimates			
St. dev. of worker effects	0.544	0.520	0.507
St. dev. of workplace effects	0.395	0.427	0.429
Correlation worker/workplace effects	0.191	0.210	0.226
Correlation worker effects/Xb	-0.281	-0.221	-0.267
Correlation workplace effects/Xb	-0.070	-0.062	-0.057
Goodness of fit			
St. dev. of log wages	0.739	0.757	0.748
R Squared	0.920	0.921	0.920
Panel B: Men			
Worker and workplace parameters			
Number of worker effects	5,576,345	6,920,461	8,941,908
Number of workplace effects	523,701	554,593	695,749
Summary of parameter estimates			
St. dev. of worker effects	0.504	0.486	0.472
St. dev. of workplace effects	0.444	0.479	0.487
Correlation worker/workplace effects	0.212	0.231	0.259
Correlation worker effects/Xb	-0.123	-0.074	-0.104
Correlation workplace effects/Xb	-0.051	-0.045	-0.048
Goodness of fit			
St. dev. of log wages	0.772	0.792	0.793
R Squared	0.909	0.913	0.916
Panel C: All			
Worker and workplace parameters			
Number of worker effects	8,271,051	10,420,514	13,822,322
Number of workplace effects	627,949	672,769	830,982
Summary of parameter estimates			
St. dev. of worker effects	0.516	0.495	0.482
St. dev. of workplace effects	0.427	0.461	0.466
Correlation worker/workplace effects	0.215	0.232	0.254
Correlation worker effects/Xb	-0.169	-0.118	-0.158
Correlation workplace effects/Xb	-0.058	-0.052	-0.053
Goodness of fit			
St. dev. of log wages	0.764	0.781	0.778
R Squared	0.910	0.913	0.916

Source: Authors' calculations ising IMSS data. Results from estimation of equation (1) via OLS. Observations correspond to largest connected set per time interval. "Xb" stands for covariates and includes the following controls age, age squared, age cube, and a monthly time trend.

Table A.2: Wage Variance Decomposition, National Level. Women Ages 25-54

	Interval 1 2004-2008	Interval 2 2009-2013	Interval 3 2014-2018	Change from int. 1 to 3
Variance and covariance				
Total variance of log wages	0.546	0.573	0.559	0.013
Variance of person effects	0.296	0.271	0.257	-0.039
Variance of firm effects	0.156	0.183	0.184	0.029
Variance of covariates (Xb)	0.021	0.014	0.022	0.001
Variance of residual	0.044	0.045	0.045	0.001
2 Cov(person effects, firm effects)	0.082	0.093	0.098	0.016
2 Cov(person effects, covariates)	-0.045	-0.027	-0.040	0.005
2 Cov(firm effects, covariates)	-0.008	-0.006	-0.007	0.001
Variance shares				
Variance of person effects	0.542	0.473	0.460	-0.082
Variance of firm effects	0.285	0.319	0.330	0.044
Variance of covariates (Xb)	0.039	0.024	0.039	0.001
Variance of residual	0.080	0.079	0.080	-0.000
2 Cov(person effects, firm effects)	0.150	0.163	0.176	0.025
2 Cov(person effects, covariates)	-0.082	-0.047	-0.072	0.011
2 Cov(firm effects, covariates)	-0.015	-0.011	-0.013	0.002
Counterfactuals for variance of log wages*				
1. No rise in correl. of person/firm effects	0.546	0.563	0.545	
2. No rise in var. of firm effects	0.546	0.538	0.522	
3. Both 1 and 2	0.546	0.536	0.516	

Source: Authors' calculations using IMSS data. Variance decomposition results from equation (2) using the estimated worker and workplace fixed effects from equation (1). The "Variance and covariance" rows show the values of the variance of log wages in the estimation sample of prime-age men and its components. The "Variance shares" rows show the share of the overall variance in log wages in the sample attributed to each one of its components. The first three columns correspond to time intervals, and the last columns is the change from 2004-2008 to 2014-2018. The "Counterfactuals for variance of log wages" rows show the variance of wages assuming that the correlation of worker/workplaces effects and the variance of workplace effects had remained constant at 2004-2008 values.

Table A.3: Wage Variance Decomposition, National Level. All Workers Ages 25-54

	Interval 1 2004-2008	Interval 2 2009-2013	Interval 3 2014-2018	Change from int. 1 to 3
Variance and covariance				
Total variance of log wages	0.584	0.610	0.606	0.022
Variance of person effects	0.266	0.245	0.233	-0.033
Variance of firm effects	0.182	0.212	0.217	0.035
Variance of covariates (Xb)	0.019	0.013	0.017	-0.002
Variance of residual	0.053	0.053	0.051	-0.001
2 Cov(person effects, firm effects)	0.095	0.106	0.114	0.019
2 Cov(person effects, covariates)	-0.024	-0.013	-0.020	0.004
2 Cov(firm effects, covariates)	-0.007	-0.005	-0.006	0.000
Variance shares				
Variance of person effects	0.456	0.402	0.384	-0.071
Variance of firm effects	0.312	0.348	0.358	0.046
Variance of covariates (Xb)	0.033	0.021	0.028	-0.004
Variance of residual	0.090	0.087	0.084	-0.007
2 Cov(person effects, firm effects)	0.162	0.174	0.188	0.025
2 Cov(person effects, covariates)	-0.042	-0.022	-0.033	0.008
2 Cov(firm effects, covariates)	-0.012	-0.009	-0.011	0.002
Counterfactuals for variance of log wages*				
1. No rise in correl. of person/firm effects	0.584	0.602	0.588	
2. No rise in var. of firm effects	0.584	0.572	0.561	
3. Both 1 and 2	0.584	0.572	0.553	

Source: Authors' calculations using IMSS data. Variance decomposition results from equation (2) using the estimated worker and workplace fixed effects from equation (1). The "Variance and covariance" rows show the values of the variance of log wages in the estimation sample of prime-age men and its components. The "Variance shares" rows show the share of the overall variance in log wages in the sample attributed to each one of its components. The first three columns correspond to time intervals, and the last columns is the change from 2004-2008 to 2014-2018. The "Counterfactuals for variance of log wages" rows show the variance of wages assuming that the correlation of worker/workplace effects and the variance of firm effects had remained constant at 2004-2008 values.

Table A.4: Variance Decomposition with the Kline et al. (2020) Variance Estimator

	Interval 1 2004-2008	Interval 2 2009-2013	Interval 3 2014-2018
Connected Set			
Total variance of log wages	0.596	0.627	0.628
Variance of worker effects	0.254	0.236	0.222
Variance of workplace effects	0.197	0.230	0.237
2 Cov(worker effects, workplace effects)	0.095	0.108	0.119
Leave-One-Out Connected Set			
Total variance of log wages	0.596	0.627	0.628
Variance of worker effects	0.254	0.235	0.222
Variance of workplace effects	0.193	0.227	0.235
2 Cov(worker effects, workplace effects)	0.098	0.110	0.121
KSS Corrected in Leave-One-Out Connected Set			
Total variance of log wages	0.596	0.627	0.628
Variance of worker effects	0.252	0.234	0.220
Variance of workplace effects	0.193	0.226	0.234
2 Cov(worker effects, workplace effects)	0.099	0.111	0.121

Source: Authors' calculations using IMSS data. Variance decomposition results from equation (2) using the estimated worker and workplace fixed effects from equation (1). The rows in each panel show the values of the variance of log wages in the estimation sample of prime-age men and its components. The "Connected Set" Panel shows the original estimates in the connected set from Table 4. The "Leave-one-out Connected Set" panel shows estimates in the workplaces that remain in the connected set in every leave-one-out sample. The "KSS Corrected in Leave-One-Out Connected Set" shows estimates of the variance components using the correction by Kline et al. (2020). We use the "match" leave-one-out estimator, leaving out worker-workplace matches one at a time. To approximate the components, we use 50 iterations of the JILA algorithm. See Kline et al. (2020) for details.

Table A.5: Wage Variance Decomposition Across Sectors

				Change in variance	
	(1)	(2)	(3)	(4)	(5)
	Interval 1	Interval 2	Interval 3		
	2004-2008	2009-2013	2014-2018		Share
Std. dev. of mean log wages	0.359	0.374	0.377	0.0131	100.0
Std. dev. of mean worker effects	0.153	0.147	0.143	-0.0031	-23.7
Std. dev. of mean firm effects	0.255	0.270	0.275	0.0101	77.1
Correlation of mean worker effects and firm effects	0.562	0.608	0.633	0.0061	46.6

Source: Authors' calculations using IMSS data. "Std. dev. of mean log wages" is the standard deviation of average log wages across sectors. "Std. dev. of mean worker effects" is the standard deviation across sectors of the sector-averages of worker effects. "Std. dev. of mean firm effects is the standard deviation across sectors of the sector-averages of workplace effects. "Correlation of mean worker effects and firm effects" is the correlation of the sector-level average worker and firm effects. The "Change in Variance" columns show the change in the variance components and the share of variance from 2004-2008 to 2014-2018"