NBER WORKING PAPER SERIES

HOW MUCH SHOULD WE TRUST ESTIMATES OF FIRM EFFECTS AND WORKER SORTING?

Stéphane Bonhomme
Kerstin Holzheu
Thibaut Lamadon
Elena Manresa
Magne Mogstad
Bradley Setzler

Working Paper 27368
http://www.nber.org/papers/w27368

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2020

We thank audiences at various conferences and seminars for useful comments. The opinions expressed in this paper are those of the authors alone and do not reflect the views of the Internal Revenue Service, the US Treasury Department, or the National Bureau of Economic Research. We thank SOI, Statistics Norway, the IFAU and Statistics Sweden, the Fondazione Rodolfo De Benedetti, as well as the Austrian Labor Market Service and the Austrian Federal Ministry for Social Affairs, Health, Care and Consumer Protection for granting access to all data sources. This work is a component of a larger project on income risk in the United States, conducted through the SOI Joint Statistical Research Program. Mogstad and Setzler acknowledge funding from NSF Grant SES-1851808. Bonhomme, Lamadon and Manresa acknowledge funding from NSF Grant SES-1658920.

How Much Should we Trust Estimates of Firm Effects and Worker Sorting?
Stéphane Bonhomme, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler
NBER Working Paper No. 27368
June 2020, Revised August 2020
JEL No. C23,J31,J62

ABSTRACT

Many studies use matched employer-employee data to estimate a statistical model of earnings determination where log-earnings are expressed as the sum of worker effects, firm effects, covariates, and idiosyncratic error terms. Estimates based on this model have produced two influential yet controversial conclusions. First, firm effects typically explain around 20% of the variance of log-earnings, pointing to the importance of firm-specific wage-setting for earnings inequality. Second, the correlation between firm and worker effects is often small and sometimes negative, indicating little if any sorting of high-wage workers to high-paying firms. The objective of this paper is to assess the sensitivity of these conclusions to the biases that arise because of limited mobility of workers across firms. We use employer-employee data from the US and several European countries while taking advantage of both fixed-effects and random-effects methods for bias-correction. We find that limited mobility bias is severe and that bias-correction is important. Once one corrects for limited mobility bias, firm effects dispersion matters less for earnings inequality and worker sorting becomes always positive and typically strong.

Stéphane Bonhomme
University of Chicago
1126 East 59th Street
Chicago, IL 60637
sbonhomme@uchicago.edu

Kerstin Holzheu
Sciences Po
28 Rue des Saints-Peres
75007 Paris
France
kerstin.holzheu@gmail.com

Thibaut Lamadon
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
lamadon@uchicago.edu

Elena Manresa
New York University
em1849@nyu.edu

Magne Mogstad
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
magne.mogstad@gmail.com

Bradley Setzler
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
bradley.setzler@gmail.com

1 Introduction

Abowd, Kramarz, and Margolis (1999) (AKM hereafter) proposed a statistical model that uses employer-employee data to quantify the contributions of workers and firms to earnings inequality. In the AKM model, log-earnings are expressed as a sum of worker effects, firm effects, covariates, and idiosyncratic error terms. AKM showed how to estimate worker and firm fixed-effects using linear regression. The resulting estimates can then be used to decompose the variance of log-earnings into the contributions of worker heterogeneity, firm heterogeneity, and sorting of high-wage workers to high-paying firms.

Over the past two decades, the AKM model and estimator have been frequently used to analyze earnings inequality in many developed countries.[1] This work has produced several influential yet controversial conclusions, summarized in the review article by Card et al. (2018). One key conclusion is that firm-specific wage settings are important for earnings inequality, with many studies finding that firm effects contribute approximately 20% of the overall variance of log-earnings. Another key conclusion is that the correlation between firm and worker effects is often small and sometimes negative, indicating little if any sorting of high-wage workers to high-paying firm. At the same time, evidence from Germany (Card et al., 2013) and the US (Song et al., 2019) indicate that worker sorting has been increasing over time, driving much of the rise in earnings inequality in these countries.

These empirical findings have been important, not only for quantifying the sources of earnings inequality, but also for how economists model the labor market. For example, if firm effects are a key source of inequality, then it is natural to ask why similar workers are paid differently. Indeed, the evidence of significant firm effects was instrumental in the development of labor market models with frictions (Mortensen, 2003). Furthermore, if better workers do not sort to more productive firms, then one might question the empirical importance of production complementarities for the matching of workers and firms (Shimer and Smith, 2000, Eeckhout and Kircher,

[1] See, among many others, Gruetter and Lalive (2009), Mendes et al. (2010), Card et al. (2013), Goldschmidt and Schmieder (2017), Card et al. (2016), Sorkin (2018), and Song et al. (2019). The AKM approach has also been widely used in contexts other than firms and workers, including teachers and students (e.g., Rockoff, 2004), hospitals and patients (e.g., Finkelstein et al., 2016), and banks and firms (e.g., Amiti and Weinstein, 2018).

2011).

Motivated by the importance of the findings from AKM, we ask the question: How much should we trust the AKM estimates of firm effects and worker sorting? We focus on the problem of estimation, taking as given the AKM model. In particular, we assume that mobility is conditionally exogenous given worker and firm effects, and we rule out the presence of dynamics and worker-firm complementarities. Other work has examined and relaxed these assumptions (e.g., Abowd et al., 2018, Bonhomme et al., 2019).[2] Our goal is to assess the sensitivity of AKM estimates to the incidental parameter problem that arises in the AKM model, often referred to as "limited mobility bias".

Limited mobility bias is due to the large number of firm-specific parameters that are solely identified from workers who move across firms. Abowd et al. (2004) and Andrews et al. (2008, 2012) highlighted this problem, and the simulations reported in Andrews et al. (2008) suggest the bias can be substantial. If firms are weakly connected to one another because of limited mobility of workers across firms, AKM estimates of the contribution of firm effects to wage inequality are biased upwards while AKM estimates of the contribution of the sorting of workers to firms are biased downwards. While researchers have been aware of these issues, bias corrections have not been widely implemented.[3] There could be a variety of reasons for this. As Card et al. (2018) point out, bias correction necessarily involves making potentially restrictive assumptions about the model. In addition, exact computation of fixed-effects corrections is costly, and often prohibitive in large data sets. As a result, there is yet no consensus about the magnitude of the biases, and how they might alter conclusions about labor markets and inequality.

To investigate the importance of limited mobility bias, we use a variety of data sets and methods. Empirically, we study matched employer-employee data from multiple

[2]In addition, recent work has studied worker sorting with two-sided heterogeneity using different approaches (e.g., Bagger and Lentz, 2019, Hagedorn et al., 2017, Lentz et al., 2017, and Borovickova and Shimer, 2017).

[3]Recently, Kline et al. (2020) and Lachowska et al. (2020) have applied a fixed-effects method for bias-correction to data from two regions of Italy and one US state (Washington), respectively. Bonhomme et al. (2019) and Lamadon et al. (2019) have estimated linear and nonlinear models with discrete firm heterogeneity using data from Sweden and the US. Here we develop a correlated random-effects estimator and apply both fixed-effects and random-effects methods to data from a wide range of countries.

countries, including the US and several European countries: Austria, Italy, Norway, and Sweden. These countries have different wage structures and labor market institutions. By comparing the results across countries, we shed light on whether our findings are specific to the US or common across several Western economies that could potentially differ in the importance of firm-specific wage-setting and the patterns of worker mobility across firms.

Methodologically, we take advantage of the availability of econometric techniques for bias-correction. We implement fixed-effects methods for bias-correction, originally proposed by Andrews et al. (2008) and developed further by Kline et al. (2020). In addition, we propose a correlated random-effects method. To specify the joint distribution of worker and firm effects we extend the model proposed in Woodcock (2008) in two ways: we allow for non-zero off-diagonal elements in the covariance matrix, and we allow mobility decisions to correlate with worker and firm heterogeneity by relying on a firm grouping approach as in Bonhomme et al. (2019).[4]

There are advantages and disadvantages to both the random- and fixed-effects methods for bias-correction. The random-effects method restricts the means and covariances of worker and firm effects, which is not needed in the fixed-effects method. However, reducing the number of parameters can be helpful for two reasons. First, random-effects estimates may be more precise than fixed-effects ones. We investigate the variability of bias-corrected estimators in a subsampling exercise. Second, our correlated random-effects estimator is significantly computationally less costly than (approximate) fixed-effects bias-corrected estimators. The random-effects method only relies on between-worker restrictions for estimation, does not require imposing independence assumptions over time for a given worker, and does not change the estimand relative to AKM. By comparing the results across the methods, we learn whether the conclusions about limited mobility bias are sensitive or robust to the alternative approaches to bias-correction.

Our first set of results quantifies the importance of firm effects and worker sorting as sources of earnings inequality in the US labor market. To do so, we link worker and firm tax records comprising a near universe of private sector employment during 2001-

[4] As a robustness check, Bonhomme et al. (2019) allow for within-group variation when estimating the variance of firm effects. The correlated random-effects model we propose extends this approach by allowing for within-group sorting.

2015. The AKM estimates suggest that 12% of the variance of log-earnings across American workers is due to firm effects. We then apply the fixed-effects and random-effects bias-correction procedures to the US data, finding across methods that 5-6% of all earnings variation is due to firm effects. We conclude there is substantial upward-bias in the AKM estimates of the variance of firm effects, the various bias-correction methods provide similar results to one another, and the share of earnings variation explained by firm effects is half as large once one corrects for limited mobility bias.

In our second set of results, we examine the contribution of worker sorting to earnings inequality in the US labor market, as measured by the share of log-earnings variance explained by the covariance between worker and firm effects. The AKM estimates suggest that only 1% of the variance of log-earnings across workers is due to sorting of high-wage workers to high-paying firms. Applying the fixed-effects and correlated random-effects bias-correction procedures to the US data, we find across methods that 13-15% of all earnings variation is due to sorting. This translates into correlations between worker and firm effects of at least 0.30. By contrast, the AKM estimates suggest a correlation of only 0.02. This shows that limited mobility bias is important for understanding the contribution of sorting to inequality in the cross-sectional distribution of earnings. However, we find that the biases in the sorting estimates change little over time. As a result, bias-correction does not materially change the inferences drawn about changes in worker sorting and the growth in earnings inequality over time. In fact, both the AKM and the bias-corrected estimates suggest a small decline in the importance of firm effects and a modest increase in the importance of worker sorting in the US over the period 2001-2015. This finding is broadly consistent with the AKM results of Song et al. (2019) for this time period. This confirms that bias-correction is empirically important for accurately describing the cross-sectional distribution of earnings in the US, but not for understanding the growth in earnings inequality.[5]

Our third set of results shows that the main findings about firm effects and sorting are not specific to the US, but generalize to the four European countries for which we have data. The sample selection and variable definitions are harmonized, to the extent possible, across countries. Consistent with previous research, the AKM estimates

[5]Relatedly, Card et al. (2013) focus on changes over time in inequality in Germany, arguing estimated changes over time are unlikely to be affected by limited mobility bias.

suggest that about 15-25% of the variance of log-earnings is due to firm effects in these countries. Correcting for limited mobility bias reduces the contribution of firm effects to earnings inequality in these countries to around 5-15%. Thus, we conclude there is substantial upward-bias in the AKM estimates of firm effects in each country, suggesting that firm-specific wage-setting is less important than previously thought. However, the exact magnitude of the bias-corrected estimates of firm effects is more sensitive to the choice of method in the European data as compared to the US.

Considering the AKM estimates of worker sorting, we often find a negative correlation between worker and firm effects in the European countries. By comparison, when using bias-correction methods, all of the sorting estimates become positive. We find that about 10-20% of earnings inequality is due to the sorting of better workers to better firms, with implied correlations between worker and firm effects in the 0.24-0.34 range. Thus, we conclude the AKM estimates of worker sorting are severely downward biased, obscuring the fact that high-wage workers do indeed sort into high-paying firms.

In our last set of results, we perform several robustness and specification checks. Our goals are to examine and compare the behavior of the various estimators, as well as to assess the situations in which limited mobility bias is most likely to be a problem. For example, it has been conjectured that in large samples, when feasible, one might be able to side-step the issue of limited mobility bias by restricting the estimation sample to large firms. The argument is that large firms may be better connected. We find that excluding the firms that are relatively small does help attenuate limited mobility bias. However, even if one excludes all small and mid-size firms, limited mobility bias remains a serious problem, especially for estimating the sorting of workers to firms. In addition, removing all these firms would dramatically change the population of study, and, as a result, shift attention to studying earnings inequality among workers in large firms only.

Another insight from the robustness and specification analyses is that limited mobility bias is particularly severe in short panels, yet the bias-corrected estimates do not materially change depending on the length of the panel. This suggests that bias-correction methods can be particularly useful to study the evolution of firm effects and sorting over time. This finding is consistent with the study of Washington state

by Lachowska et al. (2020), who find much less bias in longer panels. Furthermore, we show that the choice of earnings or wage measure does not alter the conclusions about limited mobility bias. In Norway, we observe hours and days worked, so we are able to compare AKM and bias-corrected estimates for annual earnings, daily wages, and hourly wages. We find that the choice of outcome measure does not materially change the results.[6]

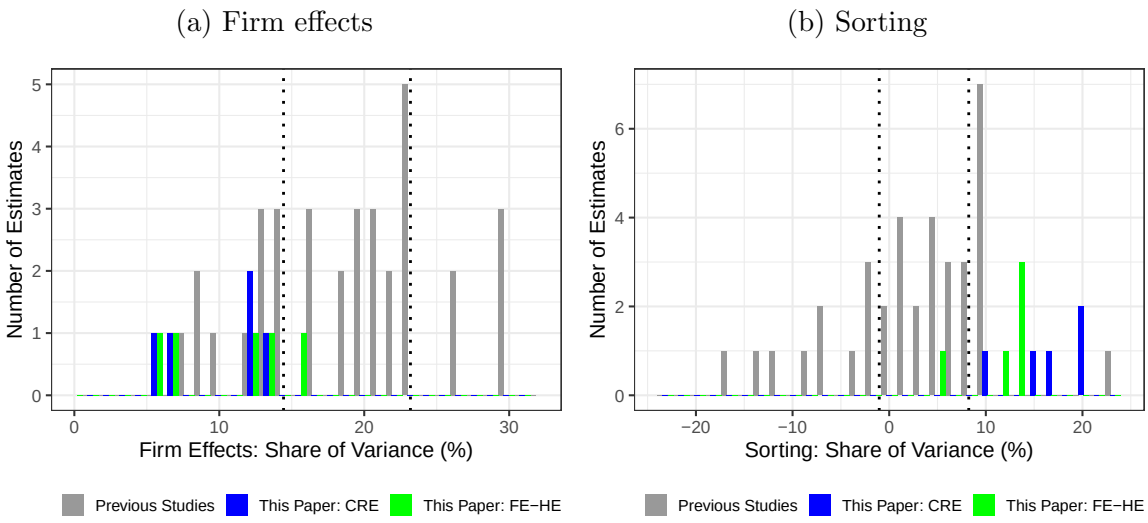
Relatedly, since the US data lacks information on hours worked or a formal measure of full-time employment, existing work has imposed a minimum earnings threshold when defining the sample and outcome measure. We find that the choice of earnings threshold does not materially change the conclusions about limited mobility bias. Lastly, we perform several robustness checks for the bias-correction methods, studying computational issues for fixed-effects methods, showing that the number of groups does not affect the correlated random-effects results, and demonstrating that posterior estimates of the variance of firm effects are close to our correlated random-effects estimates.

Taken together, these four sets of results suggest limited mobility bias is a major issue for studies using AKM to document firm effects and worker sorting. To illustrate this point and to put our results into perspective, Figure 1 compares our bias-corrected estimates to existing AKM estimates of the contribution from firm effects and worker sorting to wage or earnings inequality. We report AKM estimates from previous studies in grey. Then, for each of the five countries of study, we report estimates based on our correlated random-effects (CRE) method using the firm grouping of Bonhomme et al. (2019) in blue, and estimates based on the heteroskedastic fixed-effects method (FE-HE) of Kline et al. (2020) in green.[7] In Subfigure 1a, we focus on the contribution of firm effects. The interquartile range of estimates of the variance of firm effects in previous studies is from 14% to 23%, while the range of our bias-

[6]This finding differs from Lachowska et al. (2020), who consider quarterly UI records from Washington state (where hours information is available). In the bias-corrected estimates, they find that both the variance of firm effects and sorting are more important for earnings than for hourly wages. After bias-correction, we find sorting is more important for earnings while the variance of firm effects is more important for hourly wages.

[7]See Appendix Table D1 for a list of the 18 studies and 37 AKM estimates used in this comparison. These estimates are directly comparable to those based on CRE, as both AKM and CRE use the largest set of firms that are connected through at least one mover. We also report the results using FE-HE, which restricts attention to a subset of firms that remain connected after any given mover is removed from the sample (a leave-one-out connected component).

Figure 1: Comparison to Existing Studies



Notes: AKM estimates from previous studies in grey. Correlated random-effects (CRE) bias-corrected estimates from this paper based on the grouping of Bonhomme et al. (2019) in blue. Heteroskedastic fixed-effects (FE-HE) bias-corrected estimates from this paper using the method of Kline et al. (2020) in green. The vertical dotted lines indicate the interquartile range of estimates in previous studies.

corrected estimates is from 5% to 13% using CRE and 6% to 16% using FE-HE. In Subfigure 1b, we shift attention to the contribution of sorting. The interquartile range of estimates of the contribution of sorting in previous studies is from -1% to 8%, while the range of our bias-corrected estimates is from 10% to 20% using CRE, and 5% to 13% using FE-HE.

The paper proceeds as follows. Section 2 presents the data sources. Section 3 describes the AKM estimator and limited mobility bias. Section 4 describes the fixed-effects and random-effects methods for bias-correction. Sections 5 and 6 report our main empirical findings. Section 7 provides specification and robustness checks. Section 8 concludes.

2 Data

2.1 Data Sources

We describe the data sources we use to construct the employer-employee matched data. We start with the US and then turn to the European countries. For each country, we report specific sample selection rules due to data structure and variable availability. Next, we describe the procedure we use to harmonize the sample and variable definitions across countries.

2.1.1 United States

The US data is constructed by linking Treasury business tax filings with worker-level filings. Our sample spans 2001-2015 and our main results will focus on 2010-2015. We express all monetary variables in 2015 dollars, adjusting for inflation using the CPI. Earnings data are based on taxable remuneration for labor services reported on form W-2 for direct employees. Earnings include wages and salaries, bonuses, tips, exercised stock options, and other sources of income deemed taxable by the IRS. These forms are filed by the firm on behalf of the worker and provide the firm-worker link. We exclude workers who are employed in the public or non-profit sector by requiring that their employers file tax form 1120 (C-corporations), 1120S (S-corporations), or 1065 (partnerships). In the US data, we do not observe any information about the duration of the spell within the year. To construct a comparable sample to previous studies in other countries and in the US, we apply a full-time equivalence earnings threshold, as described below.

2.1.2 European countries

Each European country allows for the construction of a matched employer-employee data set with information on total annual earnings paid to each worker by every employer. This measure of earnings includes both direct wage payments and other sources of labor income. All data sources include information on the worker's age and gender. Countries differ in the level of detail regarding the duration of the employment spells as well as the calendar years over which data is available. In each country, we

focus on 6-year panels in the main analysis and provide results from 3-year panels for comparison, and we adjust all monetary variables for inflation.

Austria. The data, called the Arbeitsmaktdatenbank (AMDB), is co-constructed by the Austrian Labor Market Service and the Federal Ministry for Social Affairs, Health, Care and Consumer Protection using worker-level social security records. Our sample spans the years 2010-2015. A similar vintage of this data has been used by Borovickova and Shimer (2017). For each job, it includes information on start and end dates as well as total annual earnings. Given this information, we construct the daily average wage as our main outcome of interest.

Italy. The data, known as the Veneto Worker Histories, is constructed by tracking all workers in the provinces Treviso and Vicenza even if they move to other provinces in Italy. Our sample spans 1996-2001. This data is used for instance by Kline et al. (2020) and has been developed by the Economics Department at Universit Ca Foscari Venezia under the supervision of Giuseppe Tattara. For each job, it includes information on number of days worked in the year and annual earnings. Given this information, we construct the daily average wage as our main outcome of interest.

Norway. The data comes from the State Register of Employers and Employees, which covers the universe of workers and firms. Our sample spans 2009-2014. For each job, it includes information on start and end dates, annual earnings, and contracted hours. We construct the daily average wage as our main outcome of interest. Because the Norwegian data also provides hours worked per day, we also construct the average hourly wage as a secondary outcome.

Sweden. The data we use builds on the sample from Friedrich et al. (2019), and we focus on 2000-2005. The employee-employer link is built from the Register-Based Labor Market Statistics (RAMS), with access provided by the Institute for Evaluation of Labour Market and Education Policy (IFAU). The data covers the universe of workers and firms, but the sample available to us is limited to employment spells of at least two months. The sample contains information about yearly earnings,

employer identifiers and month of start and end of each spell. Given this information, we construct average monthly earnings as our main outcome of interest.

2.2 Sample harmonization and construction

To harmonize the data across countries, we apply five steps. First, as is common in the literature, whenever a worker is employed by multiple employers in the same year, we focus on the employer associated with the greatest annual earnings. Second, we restrict attention to workers employed in the private sector. Third, we restrict attention to workers who are between 25 and 60 years of age. Fourth, we adjust for differences in age and time by regressing the outcome measure on calendar year indicators and an age profile. We follow Card et al. (2018) in specifying the age profile as a third-order polynomial which is flat at age 40.

Lastly, we restrict attention to full-time equivalent (FTE) workers. Since we do not observe hours worked in US data, or a formal measure of full-time employment, we follow Lamadon et al. (2019) in defining a worker as FTE if his or her annual earnings exceed $15,000, which is approximately the annualized minimum wage and corresponds to 32.5% of the national average. To harmonize the sample selection across countries, we similarly restrict the European samples to workers with annual earnings above 32.5% of the national average. In Appendix Figure D3, we consider a range of FTE thresholds from $3,750 (about 25% of the annualized minimum wage) to $15,000 (about 100% of the annualized minimum wage). As shown in Subsection 7.5, our findings about limited mobility bias are robust to the choice of FTE threshold. In Subsection 7.6, we assess the sensitivity of the Norwegian estimates to using annual earnings (as in the US), daily wages (as in Italy, Sweden, Austria) and hourly wages as the outcome variable.

Given these harmonized samples, we prepare them for estimation by collapsing the annual observations over each 6-year panel into employment spells. Since we do not want to make assumptions about serial correlation within employment spells, we only use the mean earnings within a spell, which is sufficient to construct our estimators of interest. This approach allows for partial-year employment when constructing spells. In Subsections 7.4 (for the US) and 7.7 (for the European countries), we apply sample restrictions meant to capture only full-year employment in these spells, finding that

the conclusions are unchanged. For workers that move across employers, we further reshape the spell data into an event study format that compares the spell-level log earnings or wage measures before and after a job change. A worker that does not move across employers has only one observation. This structure effectively reduces the data to the information needed for the identification of firm effects and sorting. See Appendix A for additional details.

2.3 Descriptive Statistics

We next present descriptive information about sample sizes, distributions of moves, and earnings or wage inequality. Table 1 provides descriptive statistics for the five countries we study. It characterizes the full population (first column, under each country), the connected set (second column), and the leave-one-out set (third column). The bias-correction methods recover variance components on these two sets, as we will explain in the next two sections. These sets are constructed by computing the largest set of firms that are connected by at least one mover (connected set), and the largest set of firms that remain connected after any given mover is removed from the sample (leave-one-out set). The rows report information on the number of firms and workers, the distribution of the number of moves per firm, and certain moments of the distribution of log earnings or wages.

In Table 1 and in our main analysis, a mover is defined as a worker that is employed by at least two different firms during the sample period. In Appendix Figure D8, we consider a stricter mover definition in which a worker must be employed for at least 3 consecutive years at the first firm and at least 3 consecutive years at the second firm, only measuring earnings during intermediate years within these 3-year spells. This does not materially alter our conclusions. For further discussion, see Subsection 7.4.

Table 1 highlights several key features of the data. First, we see that at least 93% of workers belong to the connected set in each country and at least 87% belong to the leave-one-out set. By contrast, less than half of all firms belong to the connected set, and far fewer belong to the leave-one-out set. This indicates that, within each country, a large share of firms are very small, account for little of overall employment, and are not connected to other firms by movers. For further discussion of differences between the connected and leave-one-out sets, see Subsection 7.1.

Table 1: Sample Characteristics

	Austria			Italy			Norway			Sweden			US		
Set:															
Baseline Years	2010-2015			1996-2001			2009-2014			2000-2005			2010-2015		
Full Set	✓	×	×	✓	×	×	✓	×	×	✓	×	×	✓	×	×
Connected Set	×	✓	×	×	✓	×	×	✓	×	×	✓	×	×	✓	×
Leave-one-out Set	×	×	✓	×	×	✓	×	×	✓	×	×	✓	×	×	✓
Sample Counts (1,000):															
Unique Firms	446	206	140	198	92	61	233	114	78	136	63	52	7,565	2,568	1,689
(Share of Full Set)	(100%)	(46%)	(31%)	(100%)	(47%)	(31%)	(100%)	(49%)	(34%)	(100%)	(46%)	(38%)	(100%)	(34%)	(22%)
Unique Workers	3,582	3,396	3,240	1,188	1,111	1,034	1,379	1,286	1,199	1,979	1,921	1,850	59,621	55,464	52,484
(Share of Full Set)	(100%)	(95%)	(90%)	(100%)	(94%)	(87%)	(100%)	(93%)	(87%)	(100%)	(97%)	(93%)	(100%)	(93%)	(88%)
Distribution of Moves:															
Moves per Firm	2	5	8	2	4	6	2	5	7	4	10	11	2	6	8
Worker-weighted quantiles:															
10th Quantile	4	4	5	3	3	4	3	3	4	4	5	6	3	4	5
50th Quantile	52	51	56	22	22	25	26	26	29	77	77	82	56	58	67
90th Quantile	605	605	629	313	311	326	397	399	420	2,354	2,352	2,484	4,214	4,304	4,676
Log Earnings Distrib.:															
Variance	0.195	0.187	0.182	0.169	0.167	0.168	0.241	0.239	0.236	0.164	0.164	0.164	0.413	0.414	0.416
Between-firm Share	43%	46%	44%	46%	46%	45%	47%	47%	46%	31%	32%	31%	40%	40%	39%

Notes: This table displays descriptive statistics on the baseline panel data from the US and four European countries. For each country, it provides information on the full set, connected set, and leave-one-out set.

Second, while each country has a large number of moves for the median firm, a substantial share of firms have a small number of moves. For example, in the US, the majority of firms have at least 58 moves in the connected set and 67 moves in the leave-one-out set. However, ten percent of firms have only 4 moves in the connected set and 5 moves in the leave-one-out set.

Third, while earnings or wage inequality varies substantially across countries, the between-firm share of variance tends to be more similar, ranging from 30% in Sweden to 45% in Austria and Italy. The between-firm component captures differences across firms in mean log earnings or wages. Thus, it may reflect firm effects or systematic heterogeneity in the workers that firms hire. To disentangle these two components, the AKM model takes advantage of workers moving across firms, as formalized in Section 3.

Before describing the AKM model and estimator, in Appendix Figure D1 we present an event study of the earnings changes experienced by workers moving between different types of firms, in the US sample. Following Card et al. (2013) and Card et al. (2018) we define firm groups based on the average pay of coworkers. As in previous studies we find that workers who move to firms with more highly-paid coworkers experience earnings raises while those who move in the opposite direction

experience earnings decreases of similar magnitude, and that the gains and losses for movers in opposite directions between any two groups of firms seem fairly symmetric. By comparison, earnings do not change materially when workers move between firms with similarly paid coworkers. In addition, the earnings profiles of the various groups are all relatively stable in the years before and after a job move. This lends some support to the mobility assumption in the AKM model, that workers do not select their firm based on idiosyncratic earnings growth.

3 AKM Estimator and Limited Mobility Bias

In this section we first describe the AKM estimator of Abowd et al. (1999), and we then provide initial evidence on the presence of bias in the US and Sweden.

3.1 Model, Estimator and Biases

The AKM model is

$$Y_{it} = X_{it}'\beta + \alpha_i + \psi_{j(i,t)} + \varepsilon_{it}, \tag{1}$$

where Y_{it} are the log-earnings of worker i in period t, X_{it} are exogenous covariates such as age or calendar time, α_i is the unobserved worker effect, $j(i,t)$ is the firm where i works at t, $\psi_{j(i,t)}$ is the unobserved firm effect, and ε_{it} is an idiosyncratic error term. We denote as N the number of workers, J the number of firms, and T the number of time periods. Following AKM, we assume that the following mean independence condition holds:

$$\mathbb{E}(\varepsilon_{it} \,|\, X_{11}, ..., X_{NT}, j(1,1), ..., j(N,T), \alpha_1, ..., \alpha_N, \psi_1, ..., \psi_J) = 0. \tag{2}$$

Throughout this paper we assume that (2) holds in model (1). Assuming that shocks ε_{it} are mean independent of past and future firm indicators rules out endogenous mobility and state dependence, which are important in dynamic models with wage posting or sequential bargaining. In addition, (1) and (2) imply that the conditional mean of log-earnings is additive in worker and firm effects. Additivity rules out interactions between worker effects α_i and firm effects $\psi_{j(i,t)}$ that may be economically

relevant.[8]

In this model, we will focus on the contributions of firm effects and sorting in the following variance decomposition

$$\mathrm{Var}(Y_{it} - X_{it}'\beta) = \underbrace{\mathrm{Var}(\alpha_i)}_{\text{Worker effects}} + \underbrace{\mathrm{Var}(\psi_{j(i,t)})}_{\text{Firm effects}} + \underbrace{2\mathrm{Cov}(\alpha_i, \psi_{j(i,t)})}_{\text{Sorting}} + \underbrace{\mathrm{Var}(\varepsilon_{i,t})}_{\text{Residual}}. \quad (3)$$

We now describe the AKM estimator of the "Firm effects" and "Sorting" components in this decomposition.

The AKM fixed-effects estimator treats $\alpha = (\alpha_1, ..., \alpha_N)'$ and $\psi = (\psi_1, ..., \psi_J)'$ as parameter vectors. It will be convenient to write (1) and (2) in vector form, as

$$Y = X\beta + A\gamma + \varepsilon, \quad \mathbb{E}(\varepsilon \,|\, X, A, \gamma) = 0, \quad (4)$$

where Y and ε are $NT \times 1$, X is a matrix with NT rows, and A is a matrix with NT rows and $N + J$ columns.[9] The vector $\gamma = (\alpha', \psi')'$ includes worker and firm effects, and the matrix $A = [A_W \; A_F]$ depends on worker and firm indicators.

The slope parameter β can be estimated using OLS after partialling out worker and firm indicators.[10] For simplicity, in the presentation we will treat β as known, and redefine $Y_{it} - X_{it}'\beta$ as the outcome variable. That is, we will work with the model

$$Y = A\gamma + \varepsilon, \quad \mathbb{E}(\varepsilon \,|\, A, \gamma) = 0. \quad (5)$$

We start by assuming that $A'A$ is non-singular. This requires working within a connected component of the firm-worker graph (Abowd et al., 2002), and imposing one normalization on γ, e.g., one of the firm effects being equal to zero. With some abuse of notation we still denote as A the resulting selection of rows and columns of the A matrix, and we redefine N, J, T appropriately. Then, the AKM estimator of

[8] See, e.g., Abowd et al. (2018) and Bonhomme et al. (2019). An interesting question would be to study the properties of the bias-correction methods we use when the AKM model is misspecified.

[9] Note the conditioning on α and ψ in (4) is not necessary here, since we are treating them as deterministic parameters. In random-effects methods below we will be treating α and ψ as random.

[10] Formally, denote as $A^\dagger$ the Moore-Penrose inverse of A, and as $M_A = I - AA^\dagger$ the residual "hat" projection matrix. The AKM estimator of β is $\widehat{\beta} = (X'M_AX)^{-1}(X'M_AY)$. When $A'A$ is non-singular, $M_A = I - A(A'A)^{-1}A'$, however M_A remains well-defined under singularity.

worker and firm effects is the least-squares estimator

$$\widehat{\gamma} = (A'A)^{-1}A'Y.$$

As in other studies using AKM, we are interested in the variance components in (3), such as the variance of firm effects $\text{Var}(\psi_{j(i,t)})$ and the covariance between worker and firm effects $\text{Cov}(\alpha_i, \psi_{j(i,t)})$. Variance components can be written as quadratic forms in γ; that is, $V_Q = \gamma' Q \gamma$ for some matrix Q. Note that Q typically depends on A, although we leave the dependence implicit in the notation. The AKM or "fixed-effects" (FE) estimator of V_Q is then

$$\widehat{V}_Q^{\text{FE}} = \widehat{\gamma}' Q \widehat{\gamma}.$$

To see that $\widehat{V}_Q^{\text{FE}}$ is biased, note that

$$\mathbb{E}[\widehat{V}_Q^{\text{FE}}] = V_Q + \underbrace{\mathbb{E}[\varepsilon' A (A'A)^{-1} Q (A'A)^{-1} A' \varepsilon]}_{=\text{Bias}_Q}, \tag{6}$$

where the expectations are conditional on A and γ. The expected AKM estimator $\mathbb{E}[\widehat{V}_Q^{\text{FE}}]$ differs from the true variance component V_Q in general, due to the presence of the bias term Bias_Q. Note that the bias is due to V_Q being quadratic in γ. In contrast, the AKM estimates $\widehat{\gamma}$ of the level of worker and firm effects are unbiased under (4).

As explained by Andrews et al. (2008), the bias intuitively arises from an insufficient number of job movers in the firm. As a result of "limited mobility bias", the AKM variance of firm effects tends to be overstated. In turn, the covariance between worker and firm effects tends to be negatively biased, since worker effects and firm effects enter (1) additively. Jochmans and Weidner (2019) show that the magnitude of the bias is inversely related to the degree of connectivity of the firm-worker graph. A limiting case is when the graph is disconnected, i.e., when $A'A$ is singular. Within a connected component, the bias can still be large when connectivity is weak. An implication of their analysis is that the structure of the bias is complex, since it depends on the (large) matrix A of worker and firm indicators. Hence, the magnitude of the bias is ultimately an empirical question.

3.2 Empirical Illustration of Limited Mobility Bias

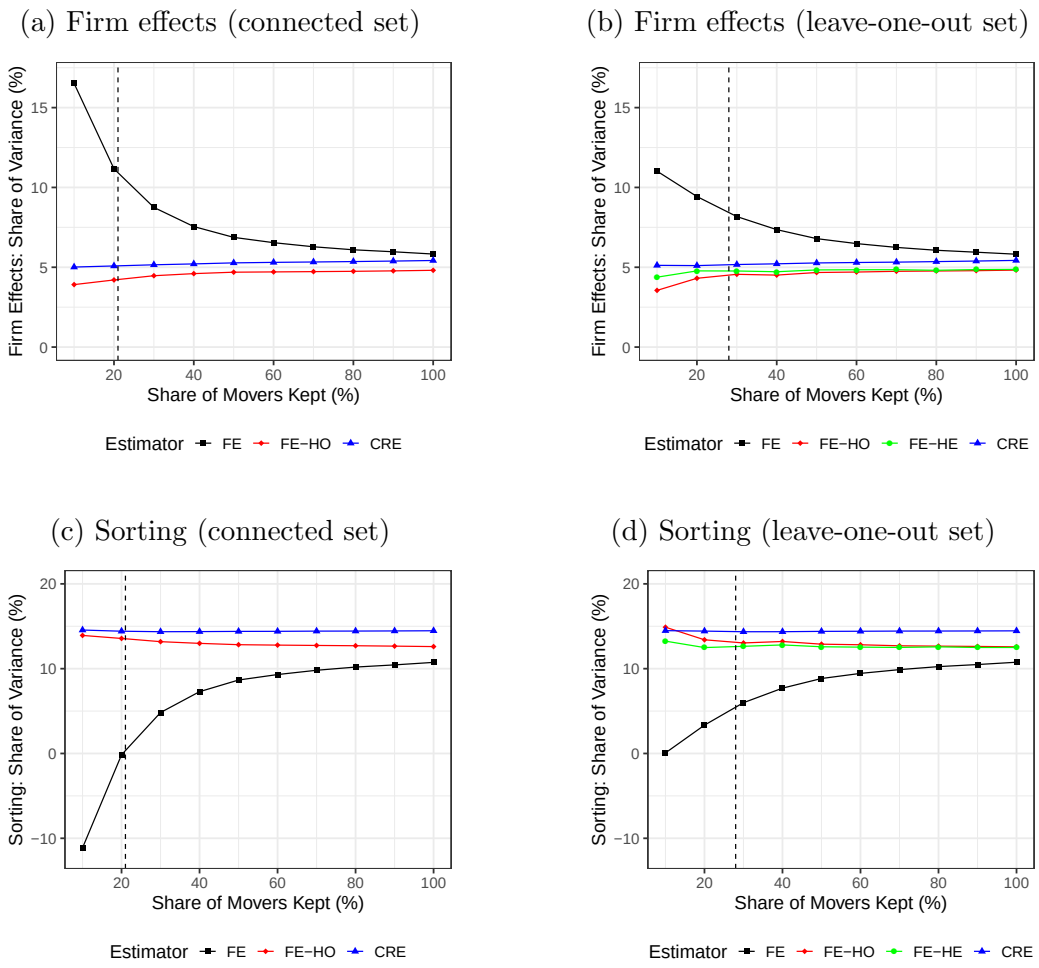
To get a sense of the scope for limited mobility bias, an informal approach is to apply the estimator of Abowd et al. (1999) (FE) to alternative samples of workers and firms that are comparable except for the number of movers per firm. Figure 2 and Figure 3 present the results from such an analysis for the US and for Sweden, respectively, using a subsampling strategy inspired by Andrews et al. (2008, 2012).

In each figure, we consider a subsample of firms with a relatively large number of movers; that is, at least 15 movers per firm over a six year period. Then, we apply the FE estimator to this sample. Next, we remove movers randomly within each firm, with a pre-specified sampling probability, before re-estimating the FE estimator. This allows us to investigate how the AKM estimates change as we reduce the number of movers yet keep the set of firms approximately the same. It is important to observe that we start losing a significant number of firms as we approach samples with only 10% of the movers (see Appendix Figure D2). Thus, we only interpret this exercise as a motivation for the use of the bias-correction methods that we will describe in the next section. For completeness, we perform this exercise for both the connected set and the leave-one-out connected set of firms.[11] For each set, we also report bias-corrected estimates both for the correlated random-effects and the fixed-effects methods. We will discuss these bias-corrected estimates in Section 4.

Subfigure 2a provides estimates of the contribution of firm effects to earnings inequality, i.e., $\text{Var}(\psi_{j(i,t)})/\text{Var}(Y_{it})$, for the connected set. Focusing on the FE estimates in the black line, we find that the variance of firm effects declines monotonically as the number of movers per firm increases. Consistent with limited mobility bias, the fewer the number of movers per firm, the larger the variance of firm effects. For approximately the same set of firms, the estimated variance of firm effects is about twice as large (11 percent) if we only keep 20% of the movers within each firm (on average, 6 movers per firm) as compared to what we obtain if we keep all the movers per firm (at a minimum 15 and, on average, 28 movers per firm). By way of comparison, there are around 6 movers per firm in the full estimation sample, which roughly corresponds to the number of movers per firm when randomly removing 79%

[11] The heteroskedastic fixed-effects method for bias-correction of Kline et al. (2020) recovers estimates of the variance components on the leave-one-out connected set.

Figure 2: Evidence on Limited Mobility Bias in the United States



Notes: In this figure, we consider the subset of firms in the US with at least 15 movers. We randomly remove movers within each firm and re-estimate the variance of firm effects and covariance between firm and worker effects using the various estimators. For each estimator, we repeat this procedure several times then average the estimates across repetitions. The procedure allows us to keep the connected or leave-one-out set of firms approximately the same and examine the bias that results from having fewer movers available in estimation. The vertical dashed line approximates the point at which movers per firm in this sample matches movers per firm in the full sample.

of movers in the sample with originally 15 or more movers per firm, as indicated by a dashed vertical line. Subfigure 2b repeats this analysis for the leave-one-out set. The results are similar for the leave-one-out set, though FE is subject to less limited mobility bias, reflecting that the leave-one-out set has more movers per firm.

Subfigure 2c provides estimates of the contribution of worker sorting to earnings inequality, i.e., $2\,\mathrm{Cov}(\alpha_i, \psi_{j(i,t)})/\,\mathrm{Var}(Y_{it})$, for the connected set. Focusing again on the FE estimates in the black line, we find that the covariance between worker and firm effects increases monotonically as the number of movers per firm increases. For approximately the same set of firms, the FE estimate of the contribution of worker sorting to earnings inequality is about 10% when we keep all movers per firm. However, if we only keep 20% of the movers within each firm, the FE estimates turn negative. Subfigure 2d repeats this analysis for the leave-one-out set. The results are again broadly similar for the leave-one-out set.

In Figure 3, we again consider a sample of firms with at least 15 movers, but this time for our data from Sweden. As in the analyses for the US data, we remove movers randomly within firms before re-estimation. Focusing on the FE estimates in the black line, the pattern is largely similar to the US data, with the contribution to wage inequality of firm effects (respectively, worker sorting) monotonically increasing (resp., decreasing) for the FE estimator as the number of movers per firm decreases. Also in this case, the results are broadly similar for the leave-one-out set.

4 Bias-correction: methods and illustration

In this section we outline the methods we use to deal with the presence of limited mobility bias in AKM estimates, and we then illustrate the methods empirically.

4.1 Methods

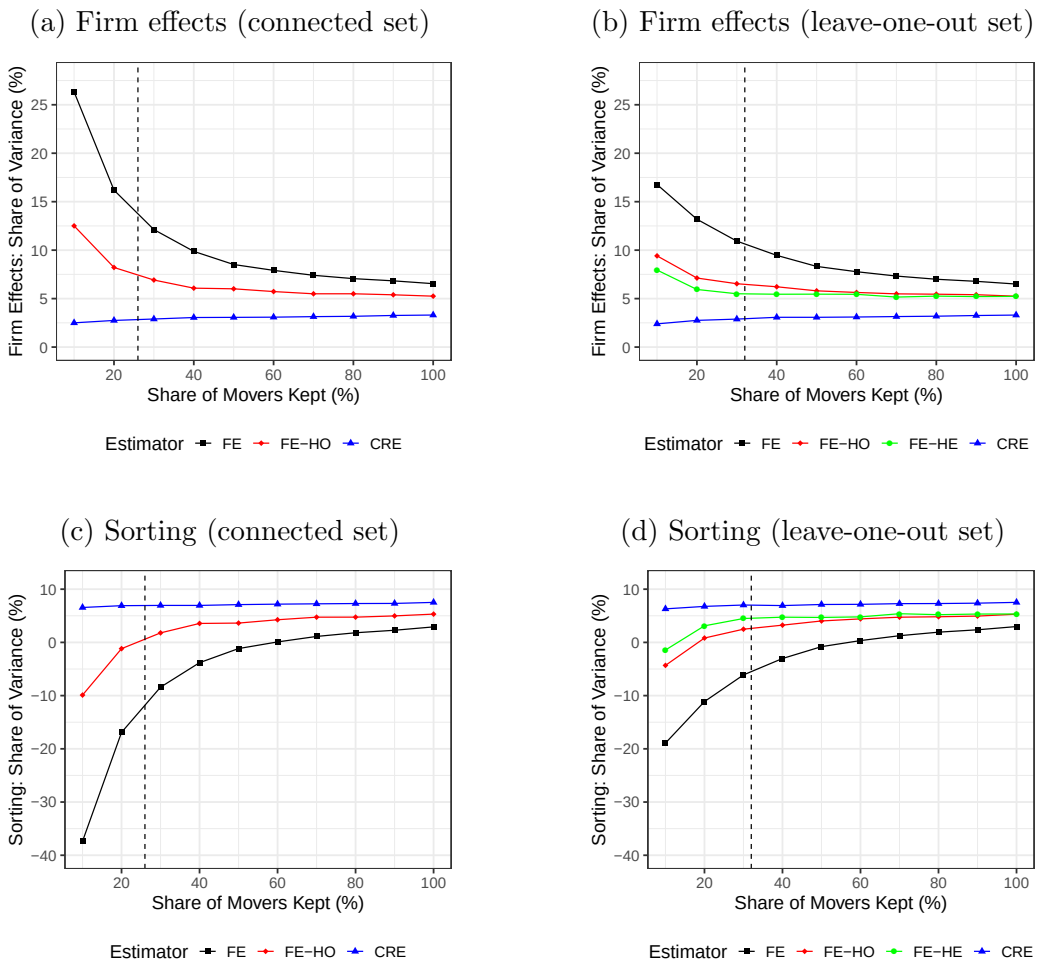
We now describe the fixed-effects and random-effects methods we use for bias-correction.

Fixed-effects. Andrews et al. (2008) note that the bias in (6) can be written as

$$\mathrm{Bias}_Q = \mathrm{Trace}\left(A(A'A)^{-1}Q(A'A)^{-1}A'\Omega(A)\right),$$

where $\Omega(A) = \mathrm{Var}(\varepsilon \,|\, A)$ is the covariance matrix of errors. Andrews et al. (2008) propose an estimator of the bias in the homoskedastic case, under the assumption

Figure 3: Evidence on Limited Mobility Bias in Sweden



Notes: In this figure, we consider the subset of firms in Sweden with at least 15 movers. We randomly remove movers within each firm and re-estimate the variance of firm effects and covariance between firm and worker effects using the various estimators. For each estimator, we repeat this procedure twenty times then average the estimates across repetitions. The procedure allows us to keep the connected or leave-one-out set of firms approximately the same and examine the bias that results from having fewer movers available in estimation. The vertical dashed line approximates the point at which movers per firm in this sample matches movers per firm in the full sample.

that $\Omega(A) = \sigma^2 I$, for I the identity matrix. Specifically, they construct

$$\widehat{\text{Bias}}_Q^{\text{FE}-\text{HO}} = \widehat{\sigma}^2 \, \text{Trace}\left((A'A)^{-1}Q\right),$$

using an unbiased estimator of the variance.[12] Under homoskedastic, independent observations, $\widehat{\text{Bias}}_Q^{\text{FE}-\text{HO}}$ is unbiased for Bias_Q, so a bias-corrected estimator of V_Q is

$$\widehat{V}_Q^{\text{FE}-\text{HO}} = \widehat{V}_Q^{\text{FE}} - \widehat{\sigma}^2 \,\text{Trace}\left((A'A)^{-1}Q\right).$$

In a recent contribution, Kline et al. (2020) propose a heteroskedastic generalization. Under the assumption that $\Omega(A)$ is diagonal, they estimate its diagonal elements using the jackknife, as

$$\widehat{\sigma}_{it}^2 = Y_{it}(Y_{it} - \widehat{\alpha}_i^{-(i,t)} - \widehat{\psi}_{j(i,t)}^{-(i,t)}),$$

where $\widehat{\alpha}_i^{-(i,t)}$ and $\widehat{\psi}_{j(i,t)}^{-(i,t)}$ are AKM estimates on a subsample where observation (i,t) has been taken out. In particular, computing the estimator requires focusing on a set that remains connected when any (i,t) observation has been taken out: a "leave-one-out" set. Hence, for this method the estimand changes relative to AKM. Letting $\widehat{\Omega}(A)$ be the diagonal matrix with diagonal elements $\widehat{\sigma}_{it}^2$, the following estimator is unbiased under heteroskedastic, independent observations:

$$\widehat{V}_Q^{\text{FE}-\text{HE}} = \widehat{V}_Q^{\text{FE}} - \text{Trace}\left(A(A'A)^{-1}Q(A'A)^{-1}A'\widehat{\Omega}(A)\right).$$

Kline et al. (2020) provide conditions under which $\widehat{V}_Q^{\text{FE}-\text{HE}}$ is consistent, and they derive its limiting distribution.

When implementing these methods we will collapse observations at the spell level. This ensures the above estimators are unbiased in the presence of serial correlation within spell, under homoskedasticity and heteroskedasticity respectively.[13] However, in practice, exact computation of $\widehat{V}_Q^{\text{FE}-\text{HO}}$ and $\widehat{V}_Q^{\text{FE}-\text{HE}}$ requires computing the trace of a large matrix inverse, which is prohibitive in most samples we use. For this reason we will rely on approximation methods (Gaure, 2014, Kline et al., 2020); see Subsection 7.8 and Appendix B.

[12] The estimator $\widehat{\sigma}^2 = (NT-N-J)^{-1}Y'(I-A(A'A)^{-1}A')Y$ is unbiased for σ^2 when observations are independent and homoskedastic.

[13] As pointed out in Kline et al. (2020), when T=2, FE-HE estimators of firm effects and sorting components are also robust to the presence of serial correlation between spells. In the empirical analysis, we focus on 6 years of data to compare to the literature, and we collapse earnings observations at the spell level; in particular, a stayer is only observed once. See Appendix B for details.

Random-effects. Random-effects methods are popular in many panel data applications, yet they are rarely used in matched employer-employee settings. Here we introduce a correlated random-effects (CRE) estimator for variance components. Compared to fixed-effects estimators, the CRE estimator requires modeling the means and covariances of worker and firm effects. However, CRE depends on a smaller number of parameters. This parsimony is helpful for computational tractability, and to obtain more precise estimates.

Our starting point is the random-effects specification in Woodcock (2008). Woodcock postulates that the conditional distribution of worker and firm effects $\gamma = (\alpha', \psi')'$ given worker and firm indicators A has mean $\mu(A)$ and variance $\Sigma(A)$.[14] In his specification, neither μ nor Σ depend on A, and Σ is diagonal. Woodcock uses this model as a prior for the worker and firm effects, and computes posterior estimates. We relax this specification in two ways. First, we allow $\Sigma(A)$ to be non-diagonal. Second, we allow $\mu(A)$ and $\Sigma(A)$ to depend on A. Both modifications improve over the baseline specification since the CRE model allows for job mobility to depend on worker and firm effects, and more generally it explicitly allows for sorting patterns.

Specifically, we impose three orthogonality conditions on $\Sigma(A)$ and the covariance matrix $\Omega(A)$ of ε_{it}:

$$\text{Cov}(\alpha_i, \psi_j) = 0 \text{ for } (i,j) \in \mathcal{S}_1, \tag{7}$$

$$\text{Cov}(\psi_j, \psi_{j'}) = 0 \text{ for } (j,j') \in \mathcal{S}_2, \tag{8}$$

$$\text{Cov}(\varepsilon_{it}, \varepsilon_{i't'}) = 0 \text{ for } t\,,\, t',\, i \neq i', \tag{9}$$

where all covariances are conditional on A but we omit the dependence in the notation. Here $\mathcal{S}_1$ contains worker-firm pairs (i,j) such that i never works in j at any point in the sample, and $\mathcal{S}_2$ contains firm pairs (j,j') where $j \neq j'$.

(7) and (8) are conditions about the covariance structure of worker and firm effects. Such conditions are not needed in fixed-effects approaches. Allowing the mean vector $\mu(A)$ and the variance matrix $\Sigma(A)$ to depend on worker and firm indicators A will be helpful to relax these conditions by restricting the sets $\mathcal{S}_1$ and $\mathcal{S}_2$. Indeed, assuming that (8) holds for all firm pairs may be empirically strong, if for example firms j and

[14] The model in Woodcock (2008) also accounts for covariates, which we abstract from in the presentation.

j' that are close to each other in economic distance have correlated effects ψ_j and $\psi_{j'}$ because they share the same suppliers. In the model we implement we will group firms, and we will only assume that ψ_j and $\psi_{j'}$ are uncorrelated conditional on j and j' belonging to different firm groups.[15] Likewise, we will only assume that α_i and ψ_j are uncorrelated in (7) when i never visits the group of firm j. In turn, (9) is an assumption on the covariance structure of ε_{it}. Note that this condition does not restrict the covariance matrix $\Omega(A)$ beyond cross-worker covariances.

Based on (7)-(8)-(9), if one is willing to assume in addition that α_i, ψ_j, and ε_{it} are independent of A, one can build a simple CRE specification that depends on only three parameters: the variance of firm effects and the covariance between worker and firm effects, which are our parameters of interest, and the covariance between the worker effects of two workers who are employed in the same firm at some point in time. Hence this model is very parsimonious. Moreover, the parameters can be recovered from cross-worker covariance restrictions.

As an example, consider two workers i and i' who work in the same firm in period t. Both i and i' move between t and t', and i' (respectively, i) moves to a firm where i (resp., i') never works. In this case the variance of firm effects can be recovered from

$$\begin{aligned}
\mathrm{Cov}(Y_{it'} - Y_{it}, Y_{i't'} - Y_{i't}) &= \mathrm{Cov}(\psi_{j(i,t')} - \psi_{j(i,t)} + \varepsilon_{it'} - \varepsilon_{it}, \\
&\qquad \psi_{j(i',t')} - \psi_{j(i',t)} + \varepsilon_{i't'} - \varepsilon_{i't}) \\
&= \mathrm{Cov}(\psi_{j(i,t')} - \psi_{j(i,t)}, \psi_{j(i',t')} - \psi_{j(i',t)}) \\
&= \mathrm{Cov}(\psi_{j(i,t)}, \psi_{j(i',t)}) \\
&= \mathrm{Var}(\psi_{j(i,t)}), \qquad (10)
\end{aligned}$$

and the covariance between worker and firm effects can be recovered from

$$\begin{aligned}
\mathrm{Cov}(Y_{it'} - Y_{it}, Y_{i't'}) &= \mathrm{Cov}(\psi_{j(i,t')} - \psi_{j(i,t)} + \varepsilon_{it'} - \varepsilon_{it}, \alpha_{i'} + \psi_{j(i',t')} + \varepsilon_{i't'}) \\
&= \mathrm{Cov}(\psi_{j(i,t')} - \psi_{j(i,t)}, \alpha_{i'} + \psi_{j(i',t')}) \\
&= \mathrm{Cov}(\psi_{j(i,t')} - \psi_{j(i,t)}, \alpha_{i'}) \\
&= -\mathrm{Cov}(\psi_{j(i',t)}, \alpha_{i'}). \qquad (11)
\end{aligned}$$

[15] A related approach would be to only consider firms j and j' in $\mathcal{S}_2$ that do not directly share a worker (i.e., a mover), although they might share workers indirectly through other firms j''.

To derive both (10) and (11) we have used the model in the first line, (2) and (9) in the second line, and (8) in the third line. In the last line, we have used that $j(i,t) = j(i',t)$ to derive (10), and we have used (7) to derive (11). In addition, this simple CRE model implies a number of overidentifying restrictions. Covariance restrictions such as (10) and (11) are the basis of our strategy to estimate the CRE model.

It is important to observe that assuming that α and ψ are independent of A would be restrictive. For example, this assumption would require mobility across firms not to depend on worker or firm effects. To build a flexible specification, we allow $\mu(A)$ and $\Sigma(A)$ to depend on A by using the grouping strategy of Bonhomme et al. (2019). Specifically, we cluster firms into K groups on the basis of their empirical earnings distributions. We use the k-means clustering algorithm for the grouping, and use $K = 10$ in our baseline specification. Given this grouping, we allow the means and variances of worker and firm effects to depend on the groups, but not on the worker and firm identities within these groups. Similarly, we allow the covariances in $\Sigma(A)$ to depend on the groups (or pairs of groups), while imposing some homogeneity assumptions so as to limit the number of parameters; see Appendix B for a detailed description. The CRE model still has many fewer parameters than the AKM fixed-effects model.

We estimate the CRE parameters by minimum distance based on mean restrictions and cross-worker covariance restrictions that exploit (7), (8) and (9). These restrictions are linear in parameters so implementation is straightforward. We provide details in Appendix B. We will report CRE estimates of variance components,

$$\widehat{V}_Q^{\text{CRE}} = \widehat{\mu}(A)'Q\widehat{\mu}(A) + \text{Trace}(\widehat{\Sigma}(A)Q). \tag{12}$$

When the firm groups are defined in terms of observable categories such as industry or commuting zone, consistency of CRE under (7), (8) and (9) follows from standard conditions for minimum distance. In addition, efficiency could be achieved using optimal weights. In our implementation, we tailor the groups to the data and construct them based on earnings using the k-means algorithm. Bonhomme et al. (2017) provide conditions for consistency of k-means clustering and estimators based on them under continuous heterogeneity, albeit in a single-agent panel data setting.

Consistency requires K to tend to infinity with the sample size. We report results based on $K = 10$ groups, and we document robustness with respect to this choice for a range of K.

In addition, in some specifications we will report posterior estimates in the spirit of empirical Bayes shrinkage. To construct posterior estimators, we will interpret our CRE model, in particular (7), (8) and (9), as a prior on the worker and firm effects. Under additional Gaussianity assumptions we will then compute posterior estimates of the variance of firm effects. This will provide a useful check, since under correct specification CRE and posterior estimates should be similar.[16]

4.2 Empirical Illustration of Bias-correction

In Figure 2, we illustrate empirically the homoskedastic fixed-effects bias-correction method of Andrews et al. (2008) (FE-HO), and our correlated random-effects method based on the firm grouping of Bonhomme et al. (2019) (CRE), and compare them to FE. As described in Section 3, this figure considers the subsample of firms with at least 15 movers. Next, we remove movers randomly within firms before applying the FE, FE-HO, and CRE estimators to each random subsample, which keeps the connected set of firms approximately the same until the share of movers kept approaches 10% (see Appendix Figure D2). We repeat this exercise for the leave-one-out set of firms, which allows us to also compare results to the heteroskedastic fixed-effects bias-correction method of Kline et al. (2020) (FE-HE).

Considering again Subfigure 2a, but now focusing on the blue (CRE) and red (FE-HO) lines, we see that, in contrast with FE, the two bias-corrected estimates are nearly identical across samples, suggesting that they are robust to the number of movers per firm. The FE estimator and the bias-corrected estimators approximately coincide when including all movers per firm. In Subfigure 2b we repeat this analysis for the leave-one-out set, which allows us to include the FE-HE bias-correction (in green). In this case also, all three bias-correction methods behave similarly, and in sharp contrast with FE, these estimators seem approximately insensitive to limited mobility bias.

Turning to Subfigure 2c, focusing on the blue and red lines, we see that CRE

[16] See Appendix B for the formula of the posterior estimator of V_Q.

and FE-HO bias-corrected estimates of the contribution of worker sorting to earnings inequality are also quite similar to each other, and the estimates do not vary much with the sample. In particular, bias-corrected estimates are always positive while FE estimates in samples with few movers are negative. In Subfigure 2d we repeat this analysis for the leave-one-out set, finding that the three bias-corrected estimators, now including FE-HE, behave similarly.

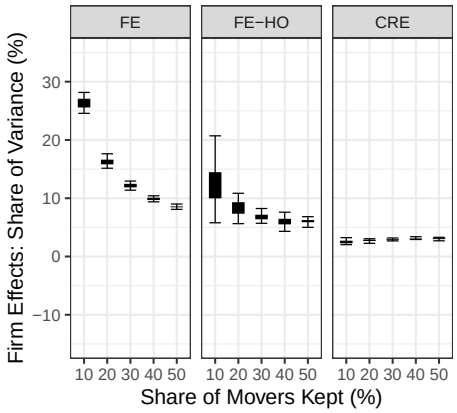
In Figure 3, we show in blue, red and green our three bias-corrected estimators, using our data for Sweden. As in the analyses for the US data, we find in Sweden that FE and the various bias-corrected estimators become more similar to one another when considering the sample of firms with at least 15 movers. When we remove movers randomly within firms, the bias-corrected CRE estimator (in blue) is approximately unchanged. In comparison to the US data, the FE-HO (in red) and FE-HE (in green) vary more across samples in the Swedish data when firms have few movers. However, like in the US, all bias-corrected estimates are very different from the FE estimates, and they suggest that FE severely over-estimates the variance of firm effects and under-estimates the covariance between worker and firm effects.

In the next sections, we report results based on both the fixed- and random-effects methods for bias-correction. The rationale for using a variety of methods is that they rely on different modeling strategies. In particular, while FE-HO and FE-HE involve a very large number of worker and firm fixed-effects, CRE depends on a smaller number of parameters and therefore can be more precise. To illustrate this, Figure 4 presents the range (whiskers) and the interquantile range (solid bar) of the estimates from the random draws of Swedish data. Whereas Figure 3 presents the mean across random draws of the data, Figure 4 presents the variability across these random draws.

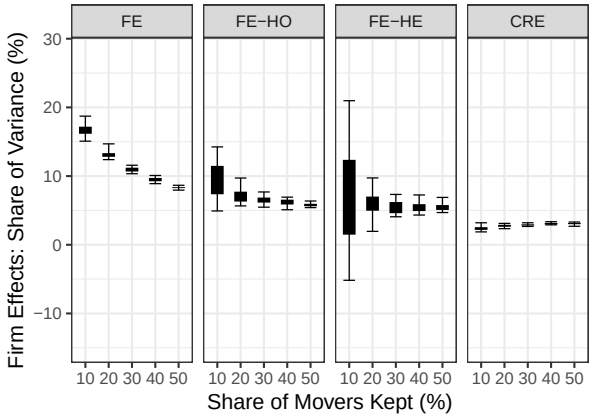
In Subfigure 4a, we consider variability in estimators of the contribution of firm effects to wage inequality for the connected set of firms. When keeping only 20% of movers per firm, we find that the range of estimates for the FE-HO estimator is 6% to 11%, while the range of estimates for the CRE estimator is only 2% to 3%. In Subfigure 4b we consider the leave-one-out set of firms, which allows us to include the FE-HE estimator. When keeping only 20% of movers per firm, we find that the range of estimates for the FE-HO estimator is 6% to 10% and the range of estimates for the FE-HE estimator is 2% to 10%, while the range of estimates for the CRE

Figure 4: Evidence on Variability of the Estimators in Sweden

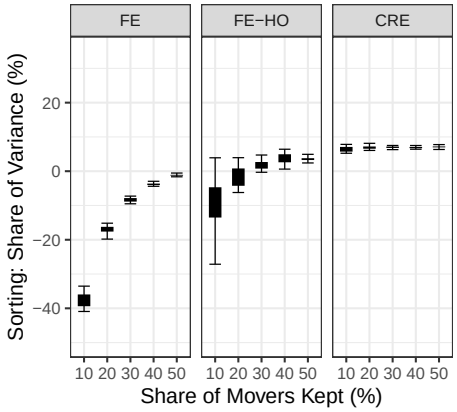
(a) Firm effects (connected set)



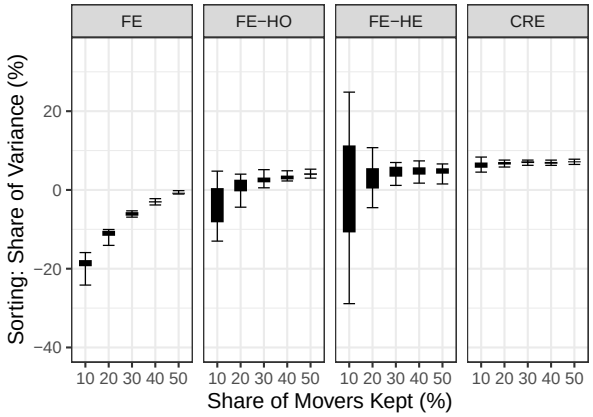
(b) Firm effects (leave-one-out set)



(c) Sorting (connected set)



(d) Sorting (leave-one-out set)



Notes: In this figure, we consider the subset of firms in Sweden with at least 15 movers. We randomly remove movers within each firm and re-estimate the variance of firm effects and covariance between firm and worker effects using the various estimators. For each estimator, we repeat this procedure twenty times, and report the overall range (whiskers) and interquartile range (solid bar) of estimates across these repetitions. The procedure allows us to keep the connected or leave-one-out set of firms approximately the same and examine the variability in the estimators when there are fewer movers available in estimation.

estimator is only 2% to 3%. In Subfigure 4c we focus on the contribution of worker sorting to wage inequality for the connected set of firms. When keeping only 20% of movers per firm, we find that the range of estimates for the FE-HO estimator is -6% to 4%, while the range of estimates for the CRE estimator is only 6% to 8%. In

Subfigure 4d we consider the leave-on-out set of firms. When keeping only 20% of movers per firm, we find that the range of estimates for the FE-HO estimator is -4% to 4% and the range of estimates for the FE-HE estimator is -5% to 11%, while the range of estimates for the CRE estimator is 6% to 8%. These findings suggest that the CRE estimator is less variable than FE-HO and FE-HE across samples.

5 Results for the United States

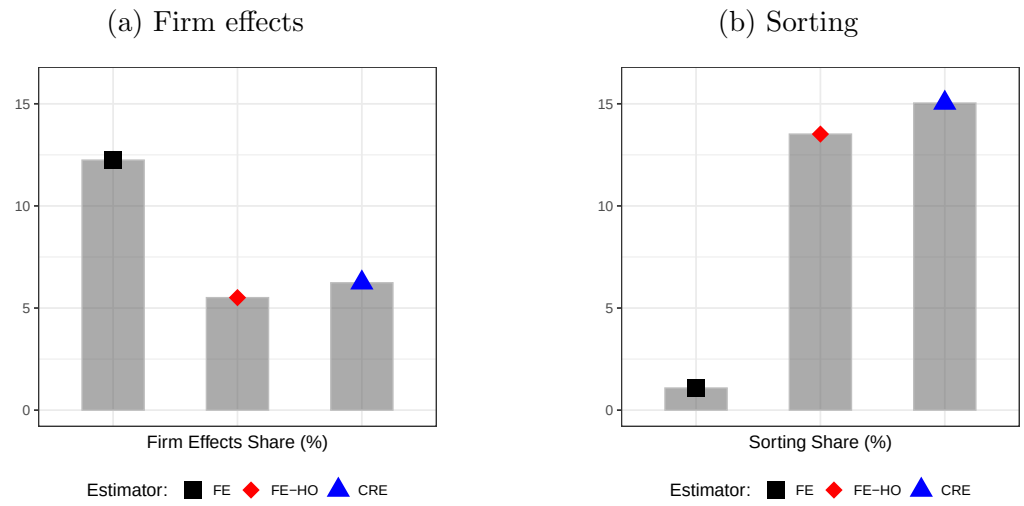
5.1 Firm Effects and Sorting

We now turn attention to quantifying the importance of firm effects and worker sorting as sources of earnings inequality in the US labor market. Figure 5 provides the main results. Subfigure 5a provides estimates of the variance of firm effects. The fixed effects (FE) estimator suggests that 12% of all earnings variation is due to firm effects. This number falls to 5% when using the homoskedastic fixed-effects estimator (FE-HO), and 6% when using our correlated random-effects estimator (CRE). We conclude there is substantial upward-bias in the FE estimator of firm effects, the two correction methods provide similar results to one another, and the share of earnings variance due to firm effects is substantially lower once one corrects for limited mobility bias.

Subfigure 5b provides the main results on the contribution to inequality of the sorting of workers to firms in the US. The FE estimator suggests that only 1% of all earnings variation is due to sorting. By comparison, this number rises to 13% when using FE-HO and 15% when using CRE. When translating these numbers into correlations, it is important to observe that estimating the correlation between worker and firm effects requires estimating the variance of worker effects, and stronger assumptions would be needed to recover the variance of worker effects (for example, one could assume a particular dependence structure within and between job spells). However, as long as the covariance is positive, it is easy to compute the following lower bound on the correlation,

$$\text{Corr}(\alpha_i, \psi_{j(i,t)}) \geq \frac{\text{Cov}(\alpha_i, \psi_{j(i,t)})}{\sqrt{\text{Var}(\psi_{j(i,t)})}\sqrt{\text{Var}(Y_{it}) - \text{Var}(\psi_{j(i,t)}) - 2\,\text{Cov}(\alpha_i, \psi_{j(i,t)})}}. \quad (13)$$

Figure 5: Firm Effects and Sorting in the United States



Notes: In this figure, we provide FE, FE-HO, and CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the United States. We consider the connected set of firms.

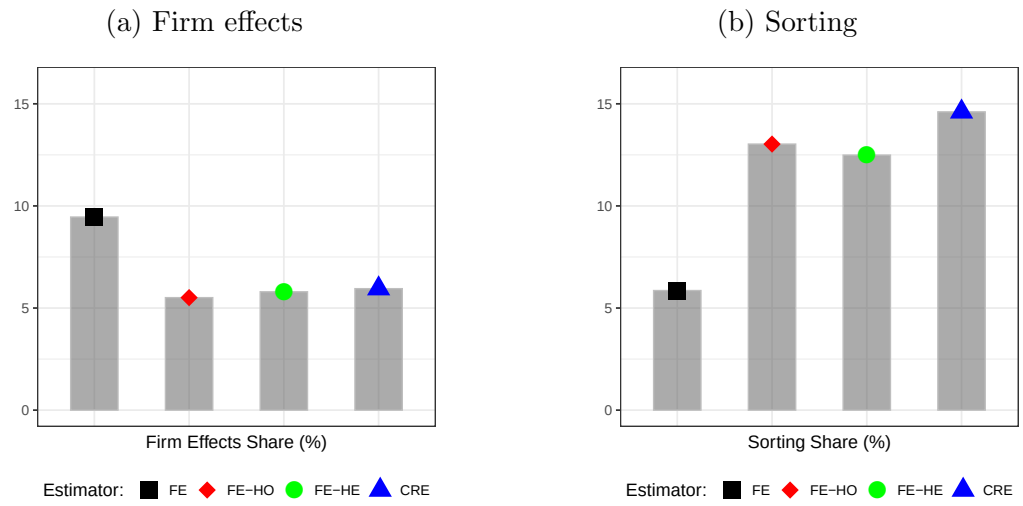
Using this lower bound, the above results translate into correlations between worker and firm effects of 0.32 when using FE-HO and 0.34 when using CRE. By contrast, FE suggests only a correlation of 0.02.[17]

In Figure 6, we present the same exercise for the leave-one-out set. Restricting to the leave-one-out set is necessary to implement the heteroskedastic fixed-effects estimator (FE-HE). The results on the leave-one-out set are broadly similar to those we obtain on the connected set. One exception is the FE estimate of sorting, which increases from 1% to 6%. This is not surprising, as the leave-one-out set excludes many of the least connected firms, thus reducing bias in the FE estimates. Yet, FE biases still appear substantial, and all three bias-correction methods behave similarly.

Next, we apply the estimators to quantify the sources of inequality within the 20 smallest states in the US. In these small states, we can compute the FE-HO and FE-HE estimators exactly, without requiring a numerical approximation. We will use this as a check of the approximation in Subsection 7.8. In Figure 7, we present the leave-one-out sample results for the smallest 20 US states. We plot CRE on the x-axis

[17] In Appendix C, we relate our sample and findings to those in Sorkin (2018) and Song et al. (2019) on US data.

Figure 6: Firm Effects and Sorting in the United States, Leave-one-out Set



Notes: In this figure, we provide FE, FE-HO, FE-HE, and CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the US. We consider the leave-one-out set of firms.

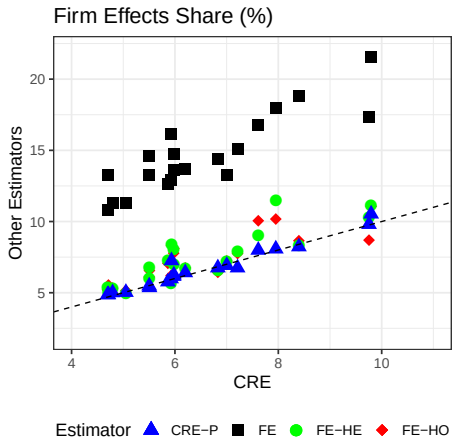
and various alternate estimators on the y-axis, so that the 45-degree line represents equality between CRE and the alternate estimators. The FE estimator is denoted by grey squares, FE-HE by green circles, and FE-HO by red diamonds. The blue triangles denote posterior CRE estimates that we will discuss in Subsection 7.9. In the small states, we find that FE suggests a large role for firm effects and small role for sorting, while all of the bias-corrected estimates provide similar (though not identical) results, as indicated by the bunching of these estimates close to the 45-degree line.

5.2 Changes over Time

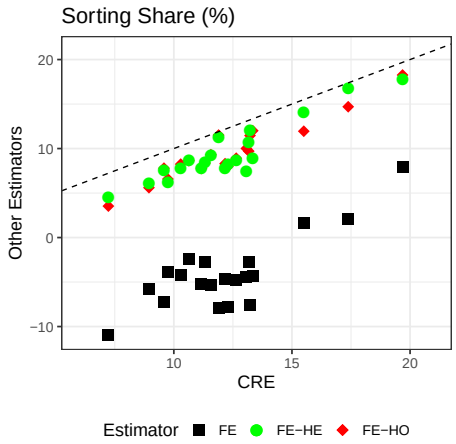
We now investigate changes over time in the contribution of firm effects and sorting to earnings inequality. We compare our baseline estimates from the final years in our sample window, 2010-2015, to the estimates we obtain for 2001-2006. The results are presented in Figure 8. The main insight from this figure is that bias-correction is important for obtaining reliable estimates of the contribution of firm effects and sorting to earnings inequality in a given time period but not for capturing how their contribution to inequality changes across time periods. The reason is that limited

Figure 7: Leave-one-out Set: Small US States

(a) Firm effects



(b) Sorting



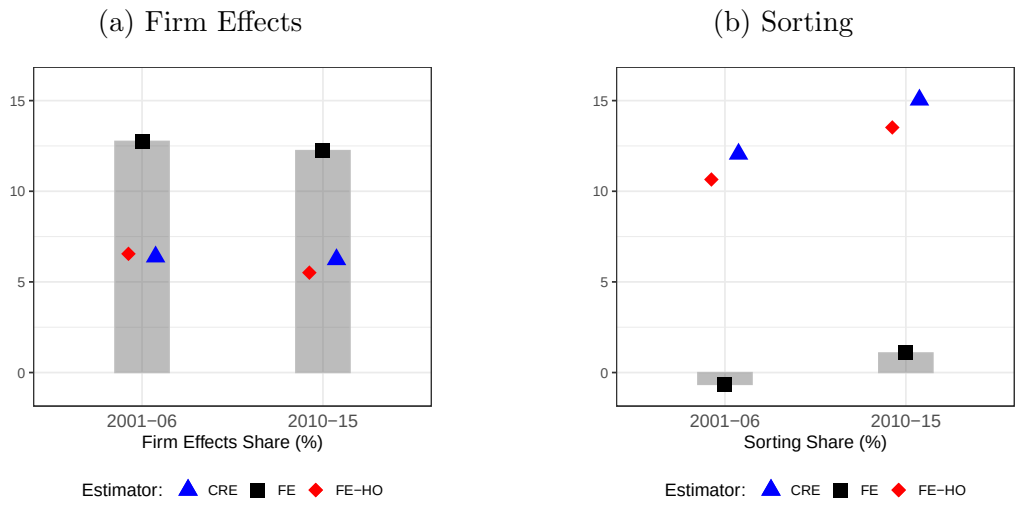
Notes: In this figure, we provide FE, FE-HO, FE-HE, and CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the 20 smallest US states. We consider the leave-one-out set of firms within each state. CRE estimates are displayed on the x-axis, and the dashed 45-degree line represents equality between CRE and the alternate estimators. The posterior CRE estimator (CRE-P) for firm effects is also displayed (subfigure a).

mobility bias, while sizable, does not change materially over time in our US sample.

To see this, consider Subfigure 8a. In this figure, the FE estimator suggests that firm effects explain around 12% of earnings variation in both time periods, while FE-HO and CRE estimates decline slightly over time. Importantly, the difference between FE and the bias-corrected estimates is very similar in both time periods. Subfigure 8b provides estimates of the contribution of sorting to earnings inequality over time. There is some evidence that sorting has become stronger, with FE, FE-HO, and CRE estimates all increasing by around 3 percentage points. However, the difference between FE and the bias-corrected estimates is nearly unchanged. In Appendix Figure D4, we find similar patterns for the leave-one-out set. We conclude that, even though bias in the FE estimates is large within each time period, the bias is not changing over time.

We now shift attention to describing the inequality within and between firms and its evolution. To do so, we follow Song et al. (2019) in expressing the variance of

Figure 8: Firm Effects and Sorting in the United States over Time



Notes: In this figure, we provide FE, FE-HO, and CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the US. We consider the connected set of firms. We compare the 6-year panel during 2001-2006 to the 6-year panel during 2010-2015.

log-earnings between firms as:

$$\underbrace{\text{Var}(\mathbb{E}[Y_{it}|j(i,t)])}_{\text{Between-firm}} = \underbrace{\text{Var}(\psi_{j(i,t)})}_{\text{Firm effects}} + \underbrace{2\text{Cov}(\alpha_i, \psi_{j(i,t)})}_{\text{Sorting}} + \underbrace{\text{Var}(\mathbb{E}[\alpha_i|j(i,t)])}_{\text{Segregation}}. \quad (14)$$

These three subcomponents capture distinct sources of inequality between firms: dispersion of firm pay premiums (Firm effects); sorting of high earning workers into high paying firms (Sorting); and a term which reflects differences in the quality of the workforce across firms (Segregation). Both worker sorting and segregation reflect non-random allocation of workers to firms. Sorting matters for aggregate inequality. However, a change in segregation by itself does not affect earnings inequality if the overall variance of worker effects $\text{Var}(\alpha_i)$ remains constant, since an increase in segregation will be offset by a reduction in the within-firm variance of worker effects. Segregation does, however, matter for the relative importance of inequality within versus between firms.

Table 2 estimates Equation (14) in our US sample. It decomposes the between-firm variation in log-earnings into components due to firm effects, sorting, and segregation.

Table 2: Between-Firm Variation in the US

Panel A.	**Total Decomposition**					
Years:	**2001-2006**			**2010-2015**		
Between Firm Share	34%			40%		
Within Firm Share	66%			60%		
Panel B.	**Share of Between Firm Variation**					
Years:	**2001-2006**			**2010-2015**		
	FE	**FE-HO**	**CRE**	**FE**	**FE-HO**	**CRE**
Firm Effects	37%	19%	19%	31%	14%	16%
Sorting	-2%	31%	35%	3%	34%	38%
Segregation	65%	50%	46%	66%	52%	46%

Notes: In this table, we provide FE, FE-HO, and CRE estimates of the contribution to between-firm earnings inequality of firm effects, sorting, and segregation in the US. We consider the connected set of firms. We compare the 6-year panel during 2001-2006 to the 6-year panel during 2010-2015.

It does this for both the 2001-06 and 2010-15 samples. In Panel A, it shows that the share of earnings variation that is between firms has increased over time, from 34% in 2001-06 to 40% in 2010-15. In Panel B, it estimates the share of between-firm variance due to firm effects, sorting, or segregation when using the FE, FE-HO, or CRE estimator. The FE estimator suggests nearly twice as much between-firm variation is due to firm effects than the bias-corrected estimators. The FE estimator suggests approximately zero between-firm variation is due to sorting, while the bias-corrected estimators find that at least one-third of between-firm variation is due to sorting. Finally, the FE estimator suggests approximately two-thirds of between-firm variation is due to segregation, while the bias-corrected estimators find that around one-half of between-firm variation is due to segregation.

Comparing the results in Panel B over time, we see that there has been a modest decline in the role of firm effects and a modest increase in the role of sorting. Our findings of the roles of firm effects and sorting in explaining the evolution of the between-firm component of inequality are qualitatively consistent with Song et al. (2019), albeit their analysis uses FE and may thus suffer from limited mobility bias.

Table 2 reveals, however, that this bias does not change materially over the time intervals we consider. This confirms that bias-correction is empirically important for accurately describing the cross-sectional distribution of earnings in the US, but not for understanding the growth in earnings inequality.[18]

In Appendix Table D2, we shift attention to the leave-one-out set. This allows us to compare all estimators on the same sample, including the FE-HE estimator. The results are broadly similar.

6 Results for European Countries

We now compare estimates of firm effects and sorting in the US to those we obtain in the four European countries, Austria, Italy, Norway, and Sweden. As described in Section 2, the sample selection and variable definitions are harmonized, to the extent possible, across countries. As in the analysis for the US data, we compare firm effect and sorting estimates across bias-correction methods, and present results separately for the connected set and the leave-one-out set. Given that some previous studies of European countries have used relatively short panel data, we present results from both 3-year and 6-year panels.

Figure 9 presents the main results for the connected set in the various countries. Subfigure 9a focuses on estimates of the share of earnings inequality due to firm effects for the 6-year panel. The FE estimator suggests 23-24% of earnings variance is due to firm effects in Italy and Norway, whereas 15-18% is due to firm effects in Austria and Sweden. When using the FE-HO bias-correction, we find a range of reductions in the estimates from about one-fifth (Austria and Italy) to about one-half (Norway and Sweden) relative to FE. The bias-correction becomes stronger when using CRE, with estimates across countries in the 5-13% range, implying reductions in the estimates ranging from about one-half to about two-thirds relative to FE.

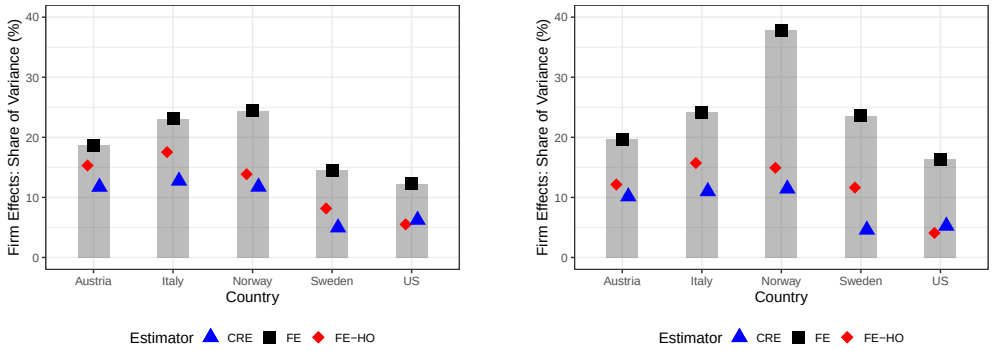
Subfigure 9b repeats these estimates for the shorter 3-year panel. The FE estima-

[18] Song et al. (2019) also argue that increases in sorting and segregation caused a large increase in between-firm inequality from 1981 to 2013. At first sight, it would seem like this is inconsistent with our findings. However, most of these increases happen before our data start. During the intervals since 2001 that we consider, Song et al. (2019) report modest increases in the contributions to between-firm inequality from sorting and segregation and a modest decrease from firm effects, consistent with our FE estimates.

Figure 9: Firm Effects and Sorting across Countries

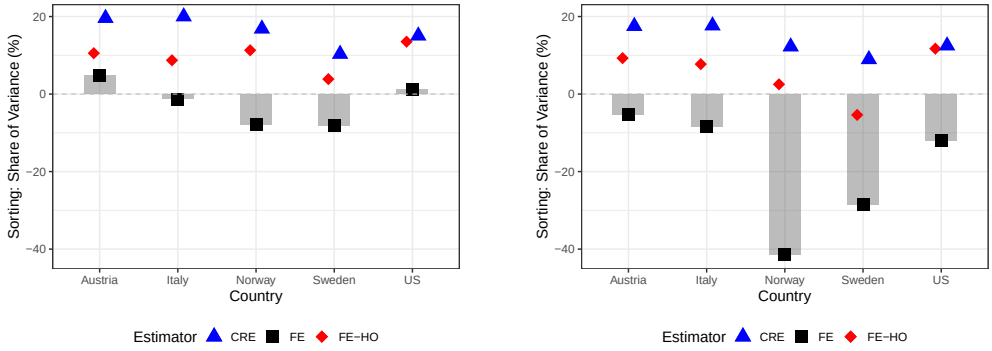
(a) Firm Effects (6-year panel)

(b) Firm Effects (3-year panel)



(c) Sorting (6-year panel)

(d) Sorting (3-year panel)



Notes: In this figure, we provide FE, FE-HO, and CRE estimates of the contribution to earnings or wage inequality of firm effects (subfigures a and b) and the sorting of workers to firms (subfigure c and d) in Austria, Italy, Norway, Sweden, and the US. We consider the connected set of firms within each country for 6-year panels (subfigures a and c) and 3-year panels (subfigures b and d).

tor suggests an even greater role for firm effects, explaining at least 20% of variance in each country with an upper estimate of about 38% (Norway). However, the CRE estimates remain in the 5-13% range, suggesting the FE estimator is much more biased in shorter panels with fewer movers per firm.

In sum, we conclude there is substantial upward-bias in the FE estimator of firm effects in each country, FE is more biased in shorter panels, CRE provides a larger bias-correction than FE-HO, and the share of earnings variance due to firm effects is

substantially smaller compared to what the FE estimator suggests.

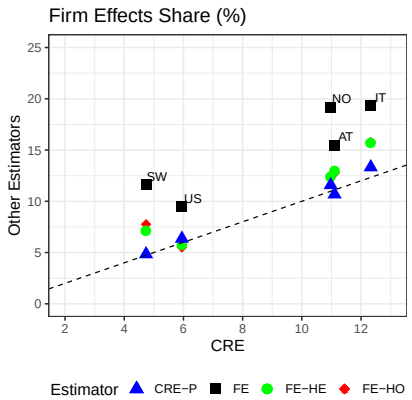
Subfigure 9c provides the main results on the contribution to inequality of the sorting of workers to firms in the various countries. When using FE, we find a negative estimate of the share of earnings variation due to sorting in all but one European country. FE estimates range from -8% (Norway and Sweden) to 5% (Austria). However, when using either the FE-HO or CRE bias-correction, all of the sorting estimates become positive. FE-HO finds estimates of the sorting contribution ranging from 4% (Sweden) to 11% (Austria and Norway), while CRE finds estimates ranging from 10% (Sweden) to 20% (Austria and Italy). The CRE estimator of sorting translates into a lower bound on the correlation between worker and firm effects (given by equation (13)) ranging from 0.24 (Sweden) to 0.34 (Austria and Italy). Subfigure 9d repeats this analysis for the shorter 3-year panels. FE suggests a negative contribution of sorting in each country, while CRE finds nearly the same estimates as in the longer 6-year panel, which implies a greater bias-correction.

Overall, we conclude the FE estimator for sorting is downward-biased and typically of the wrong sign, FE is more downward-biased in shorter panels, CRE provides a stronger bias-correction than FE-HO, and the bias-corrected share of earnings variance due to sorting tends to be substantial and is always larger than the contribution from firm effects.

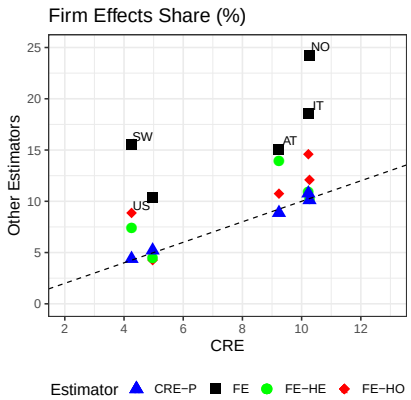
In Figure 10, we consider the leave-one-out set. This allows us to compare all estimators on the same sample, including the FE-HE estimator. We plot the CRE estimator on the x-axis and various alternate estimators on the y-axis, so that the 45-degree line represents equality between CRE and the alternate estimators. The FE estimator is denoted by grey squares, FE-HE by green circles, and FE-HO by red diamonds. The blue triangles denote posterior CRE estimates that we will discuss in Subsection 7.9. In Subfigure 10a, we provide estimates of the share of earnings variance due to firm effects for the longer 6-year panel. We see that the FE estimator is much higher than CRE in each country. By way of comparison, FE-HO and FE-HE line up well along the 45-degree line for some countries, while the estimators are somewhat larger than CRE in other countries. We repeat this analysis for the 3-year panel in Subfigure 10b, finding a similar pattern but the FE estimates are even further from the 45-degree line. In Subfigures 10c and 10d, we provide estimates of the share

Figure 10: Leave-one-out Set: Various Countries

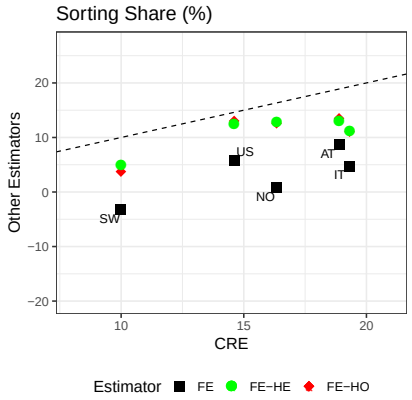
(a) Firm effects (6-year)



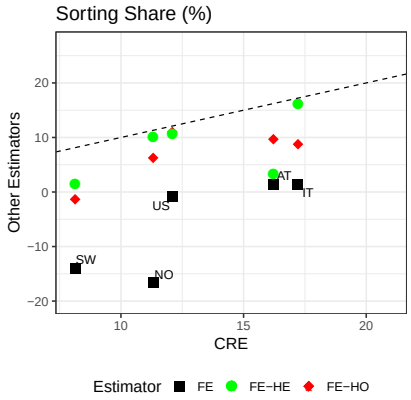
(b) Firm effects (3-year)



(c) Sorting (6-year)



(d) Sorting (3-year)



Notes: In this figure, we provide FE, FE-HO, FE-HE, and CRE estimates of the contribution to earnings or wage inequality of firm effects (subfigures a and b) and the sorting of workers to firms (subfigure c and d) in Austria, Italy, Norway, Sweden, and the US. We consider the leave-one-out set of firms within each country for 6-year panels (subfigures a and c) and 3-year panels (subfigures b and d). CRE estimates are displayed on the x-axis, and the dashed 45-degree line represents equality between CRE and the alternate estimators. The posterior CRE estimator (CRE-P) for firm effects is also displayed (subfigures a and b).

of earnings variance due to the sorting of workers to firms using the 6-year and 3-year panels, respectively. We see that the FE estimates are far below the CRE estimates, while FE-HO and FE-HE produce estimates that are very close to the CRE estimates across the various countries.[19]

Overall, the results on European countries show less agreement about the exact magnitudes of the bias-corrected estimates. Yet, the results paint a similar picture, suggesting that limited mobility bias is materially affecting the estimates of firm effects and sorting in each of these countries.

7 Robustness and Specification Checks

7.1 Connected versus leave-one-out set

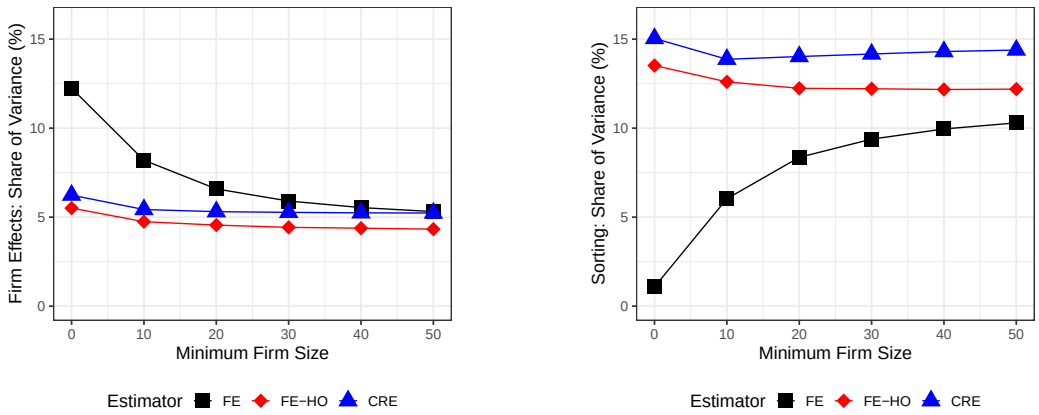
To apply the FE-HE estimator (and compare it to the other bias-correction methods), it is necessary to focus on the leave-one-out connected set of firms. A natural concern is that the leave-one-out set differ from the connected set in the composition of workers, moves, and firms. In Table 1, we saw that most workers from the connected set are also included in the leave-one-out set. However, around half of all firms in the connected set are excluded from the leave-one-out set. Appendix Figure D5 investigates whether or not this change in sample composition affects estimation results for the connected and leave-one-out sets. This figure compares the FE and CRE estimates of the contributions of firm effects and sorting to earnings inequality across countries and across the 20 small US states. As expected, the FE estimator becomes less biased on the leave-one-out set, with a reduction in the estimates of the contribution of firm effects and an increase in the estimates of the contribution of sorting. This is because the leave-one-out set is more strongly connected and is thus less susceptible to limited mobility bias. By contrast, CRE estimates are nearly identical on the connected and leave-one-out sets.

[19] In Appendix C, we relate our sample and findings to those in Kline et al. (2020) on Italian data.

Figure 11: Firm Size Restrictions in the US (Connected Set)

(a) Firm Effects

(b) Sorting



Notes: In this figure, we provide FE, FE-HO, and CRE estimates of the contribution to earnings and wage inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the US. We restrict the sample to firms with at least the number of workers indicated on the x-axis. We consider the connected set of firms for each restricted sample.

7.2 Restricting firm size in the US

One possible strategy to reduce limited mobility bias is to restrict firm size. For example, Song et al. (2019) restrict to firms with at least 20 workers, and Sorkin (2018) restricts to firms with at least 15 workers. Since large firms tend to have more movers, they also tend to be more connected. In Figure 11 we explore this possibility in our US sample by restricting the sample to firms with at least 10, 20, 30, 40, or 50 workers. This corresponds to an increase in the number of movers per firm from about 5 (baseline) to about 45 (minimum 50 workers per firm).

As expected, we find that the bias in the FE estimator diminishes as the minimum firm size rises. However, it is necessary to exclude a large share of workers and firms to limit mobility bias. For the share of variance due to firm effects, there is little remaining bias when minimum firm size is 30. For the share of variance due to sorting, there is non-neglible bias even when minimum firm size is 50. When interpreting results, it is important to observe that such restrictions change the population of study. Indeed, only 2 in 3 workers, 1 in 3 moves, and 1 in 20 firms remain in the

sample when the minimum firm size is 50 workers. Thus, it is not obvious why one is interested in earnings inequality only among this selective set of firms. In addition, while biases tend to be smaller for larger firms in our US sample there is no guarantee this will happen in other samples, since the structure of the bias depends on the network of workers and firms in complex ways (Jochmans and Weidner, 2019).

7.3 Very short panels in the US

The number of moves per firm depends strongly on the time period considered. A longer time period may help to attenuate bias by capturing more workers moving across firms. However, a longer time period also makes the assumption of time-invariant fixed-effects more difficult to justify. In Figure 9, we compared 6-year panels to 3-year panels for each country, finding that the FE estimate was more biased in the 3-year panels while the CRE results were nearly identical. To further investigate the performance of the estimators in short panels, in Appendix Figure D6, we split our baseline sample from the US during 2010-2015 into each two-year time period and apply our estimators to these 5 short panels. We find that the FE estimator becomes much more biased, with the share of variance due to firm effects rising from 12% in the 6-year panel to more than 20% in the 2-year panels, and the share of variance due to sorting falling from 1% in the 6-year panel to below -20% in the 2-year panels. Reassuringly, the bias-corrected estimates do not materially change when shortening the panel.

7.4 Mover definition: stricter definition in the US

In the European countries, our data includes start and end dates of employment spells, so we know the year in which a move occurs. However, we do not observe start and end dates in the US. To harmonize the mover definition across countries, we define a change in primary employer across years as a move, and measure earnings across all years during which the firm was the primary employer. As a check on the importance of this mover definition, we consider a stricter mover definition for the US in which a worker must be employed for at least 3 consecutive years at the first firm and at least 3 consecutive years at the second firm, only measuring earnings

during intermediate years in these multi-year spells. Appendix Figure D7 provides a diagram to help visualize the difference in these mover definitions and the timing of earnings measurement.

Imposing the strict mover definition in the US sample substantially decreases the number of movers during our sample period. Only 1 in 60 moves satisfies this particular "3-year/3-year" structure of FTE employment spells during 2010-2015. Appendix Figure D8 compares the estimates obtained under the baseline and strict definitions of movers. The FE estimate of the contribution of firm effects to earnings variation rises from 12% to 17% (the bias-corrected estimates are both around 5%), and the FE estimate of the contribution of sorting to earnings variation decreases from about 1% to about -17% (the bias-corrected estimates are both around 14%). Yet, the CRE estimates are nearly identical under the two definitions, despite the substantial change in sample composition.

7.5 Full-time equivalence: earnings thresholds in the US

In US data, we do not observe hours worked or a formal measure of full-time employment. Like Lamadon et al. (2019), we consider a worker to be full-time equivalent (FTE) if annual earnings exceed $15,000. In Appendix Figure D3, however, we consider a range of FTE thresholds from $3,750 (about 25% of the annualized minimum wage) to $15,000 (about 100% of the annualized minimum wage). As expected, total variance increases as the FTE threshold decreases. However, the share of variation explained by two-way fixed-effects is relatively stable, ranging from 12.2% to 14.6% across the range of FTE thresholds. In existing work with US data, Song et al. (2019) (using SSA data from all US states and a 25% FTE threshold) and Sorkin (2018) (using LEHD data for a subset of states and a 25% FTE threshold) find using FE that firm effects explain 9 percent (Song et al., 2019) and 14 percent (Sorkin, 2018) of the variation in log-earnings.[20]

[20] See Appendix C for a comparison with our sample and findings.

7.6 Annual earnings, daily wages, and hourly wages

In many employer-employee data sets, one does not observe hourly wages but instead observes annual earnings or average earnings over an employment spell. When applying the FE estimation, one must then take a stand on the proper measure of wages or earnings. The data from Norway is an exception, as we have accurate measures of days and hours worked in this data set.

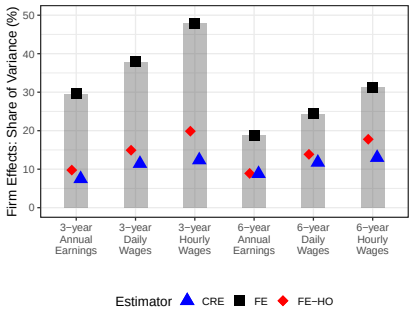
In Figure 12, we compare results on annual earnings, daily wages, and hourly wages for the same set of workers. We provide the comparison for the 6-year and 3-year panels. The FE estimate of the contribution of firm effects rises substantially when using a higher-frequency measure. In the 6-year (3-year) panel, it rises from about 19% (30%) for annual earnings to about 31% (48%) for hourly wages. The three bias-correction methods yield similar results across outcome measures. In the 6-year (3-year) panel, the CRE estimate of the contribution of firm effects rises from about 9% (8%) for annual earnings to about 13% (12%) for hourly wages. These estimates imply that FE is more biased when using higher-frequency outcome measures, and the bias-corrected estimate of the contribution of firm effects to inequality remains economically modest and somewhat greater for higher-frequency measures. A similar pattern is observed for the estimates of sorting, where FE suggests much stronger negative sorting when using hourly wages, but CRE finds substantial positive sorting with similar point estimates across outcome measures.

7.7 Full-year Employment Spells

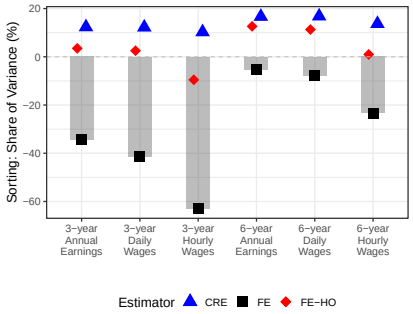
Since we do not observe start and end dates of employment spells in the US data, our main analysis allows for workers to change employers during the year. However, we observe start and end dates of employment spells for the European countries and can use these dates to restrict the sample to full-year employees. Appendix Figure D9 presents results for the sample of workers that are employed by a single firm for the entire calendar year in each European country. The bias-corrected estimates are materially unchanged when using this alternate sample definition, except the FE-HO estimate of sorting in Italy, which shows some sensitivity.

Figure 12: Norway: Annual Earnings, Daily Wages, and Hourly Wages

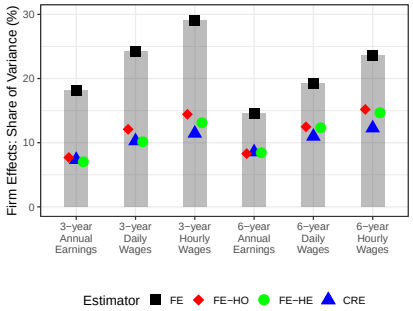
(a) Firm effects (connected set)



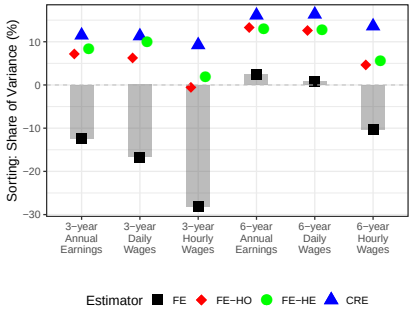
(b) Sorting (connected set)



(c) Firm effects (leave-one-out set)



(d) Sorting (leave-one-out set)



Notes: In this figure, we provide FE, FE-HO, FE-HE, and CRE estimates of the contribution to earnings or wage inequality of firm effects (subfigures a and c) and the sorting of workers to firms (subfigures b and d) in Norway. We consider the connected set of firms (subfigures a and b) and the leave-one-out set of firms (subfigures c and d) for the 6-year panel and the 3-year panel. We compare results for three outcome measures: log annual earnings, log daily wages, and log hourly wages.

7.8 Approximate versus exact computation of fixed-effects estimators

Due to the large sample size in the US, we cannot compute the FE-HO and FE-HE estimators exactly, and the estimates are computed using an approximate method following Gaure (2014) and Kline et al. (2020). A natural worry is that the approximation may perform poorly. In order to investigate this possibility, we apply the estimators to the 20 smallest US states where we can feasibly compute the exact and approximate solutions. In Appendix Figure D10 we plot exact versus approximate

FE-HO estimators in the connected set in panel (a), and exact versus approximate FE-HE estimators in the leave-one-out set in panel (b). The results show that exact and approximate solutions are close to each other, suggesting that at least in these samples the numerical approximation works well.

7.9 CRE: number of groups and posterior estimator

In our baseline CRE estimation, we cluster firms into 10 groups. One may worry that 10 groups is too restrictive. Appendix Figure D11 compares CRE estimates by number of groups in our US sample. We find that, as we increase the number of groups from 10 to 50, the estimates remain nearly identical for the earnings variation due to firm effects and sorting.

While we report random-effects estimates of variance components based on (12), using the CRE specification as a Bayesian prior we can also compute posterior estimates. Such estimates enjoy robustness properties when the CRE model is misspecified (Bonhomme and Weidner, 2019). In Figure 7 for the 20 small US states, and in Figure 10 for the various countries, we compare CRE to the other estimators. In the left graph for the variance of firm effects we also report the posterior CRE estimator (CRE-P). We find that CRE-P is almost identical to CRE. This is to be expected if the CRE model is correctly specified. Lastly, in Appendix Figure D12 we report posterior estimates for a random-effects specification that does not condition on firm groups. We estimate the firm effects variance to be less than half the CRE estimate. This suggests that accounting for the firm groups in the random-effects specification is important.

8 Conclusions

In this paper we revisit two influential yet controversial conclusions based on the model and estimator proposed by Abowd, Kramarz, and Margolis (1999): that firm effects explain around 20% of the variance of log-earnings, which points to the importance of firm-specific wage-setting for earnings inequality, and that the correlation between firm and worker effects is small and sometimes negative, which indicates little if any sorting of high-wage workers to high-paying firm. Using several econo-

metric methods and data sources for multiple countries, we find that limited mobility bias is substantial. Once bias is accounted for, firm effects dispersion matters less for earnings inequality and worker sorting becomes always positive and typically strong.

It is important to observe, however, that these conclusions rely on correctly specifying the model of earnings and the processes of worker and firm heterogeneity. There are several reasons why the AKM model may be misspecified, for example, both the assumptions that earnings are log-additive and that worker and firm heterogeneity are constant over time may be violated. To address these concerns, one possibility is to develop methods for bias-correction that are robust to misspecification. Another possibility is to enrich the model by, for example, incorporating worker-firm interactions and dynamic processes of worker and firm productivity. In this spirit, Bonhomme et al. (2019) estimate worker-firm interactions while allowing for state dependence and endogenous mobility in Sweden, while Lamadon et al. (2019) allow for worker-firm interactions and dynamic productivity processes of workers and firms in their study of the US labor market.

References

Abowd, J. M., R. H. Creecy, and F. Kramarz (2002): "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," *Unpublished Manuscript*.

Abowd, J. M., F. Kramarz, P. Lengermann, and S. Pérez-Duarte (2004): "Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States," *Unpublished Manuscript*.

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

Abowd, J. M., K. L. McKinney, and I. M. Schmutte (2018): "Modeling endogenous mobility in earnings determination," *Journal of Business & Economic Statistics*, 1–14.

Alvarez, J., F. Benguria, N. Engbom, and C. Moser (2018): "Firms and the decline in earnings inequality in brazil," *American Economic Journal: Macroeconomics*, 10, 149–89.

Amiti, M. and D. E. Weinstein (2018): "How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data," *Journal of Political Economy*, 126, 525–587.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008): "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171, 673–697.

——— (2012): "High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias," *Econ. Lett.*, 117, 824–827.

Bagger, J. and R. Lentz (2019): "An empirical model of wage dispersion with sorting," *The Review of Economic Studies*, 86, 153–190.

Bonhomme, S., T. Lamadon, and E. Manresa (2017): "Discretizing unobserved heterogeneity," *University of Chicago, Becker Friedman Institute for Economics Working Paper*.

——— (2019): "A distributional framework for matched employer employee data," *Econometrica*, 87, 699–739.

Bonhomme, S. and M. Weidner (2019): "Posterior average effects," *arXiv preprint arXiv:1906.06360*.

Borovickova, K. and R. Shimer (2017): "High Wage Workers Work for High Wage Firms," *Unpublished Manuscript*.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 36, S13–S70.

Card, D., A. R. Cardoso, and P. Kline (2016): "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, 131, 633–686.

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

Drineas, P., M. Magdon-Ismail, M. W. Mahoney, and D. P. Woodruff (2012): "Fast approximation of matrix coherence and statistical leverage," *Journal of Machine Learning Research*, 13, 3475–3506.

Eeckhout, J. and P. Kircher (2011): "Identifying sortingin theory," *The Review of Economic Studies*, 78, 872–906.

Engbom, N. and C. Moser (2018): "Earnings inequality and the minimum wage: Evidence from Brazil," *Federal Reserve Bank of Minneapolis-Opportunity and Inclusive Growth Institute Working Paper*, 7, 18–50.

Finkelstein, A., M. Gentzkow, and H. Williams (2016): "Sources of Geographic Variation in Health Care: Evidence from Patient Migration," *The Quarterly Journal of Economics*, 131, 1681–1726.

Friedrich, B., L. Laun, C. Meghir, and L. Pistaferri (2019): "Earnings dynamics and firm-level shocks," Tech. rep., National Bureau of Economic Research.

Gaure, S. (2014): "Correlation bias correction in two-way fixed-effects linear regression," *Stat*, 3, 379–390.

Goldschmidt, D. and J. F. Schmieder (2017): "The rise of domestic outsourcing and the evolution of the German wage structure," *The Quarterly Journal of Economics*, 132, 1165–1217.

Goux, D. and E. Maurin (1999): "Persistence of interindustry wage differentials: a reexamination using matched worker-firm panel data," *Journal of labor Economics*, 17, 492–533.

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

Hagedorn, M., T. H. Law, and I. Manovskii (2017): "Identifying equilibrium models of labor market sorting," *Econometrica*, 85, 29–65.

Hutchinson, M. F. (1990): "A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines," *Communications in Statistics-Simulation and Computation*, 19, 433–450.

Iranzo, S., F. Schivardi, and E. Tosetti (2008): "Skill dispersion and firm productivity: An analysis with employer-employee matched data," *Journal of Labor Economics*, 26, 247–285.

Jochmans, K. and M. Weidner (2019): "Fixed-Effect Regressions on Network Data," *Econometrica*, 87, 1543–1560.

Kline, P., R. Saggio, and M. Sølvsten (2020): "Leave-out estimation of variance components," *to appear in Econometrica*.

Lachowska, M., A. Mas, R. D. Saggio, and S. A. Woodbury (2020): "Do Firm Effects Drift? Evidence from Washington Administrative Data," Tech. rep., National Bureau of Economic Research.

Lamadon, T., M. Mogstad, and B. Setzler (2019): "Imperfect competition, compensating differentials and rent sharing in the US labor market," Tech. rep., National Bureau of Economic Research.

Lentz, R., S. Piyapromdee, and J.-M. Robin (2017): "On Worker and Firm Heterogeneity in Wages and Employment Mobility: Evidence from Danish Register Data," *Unpublished Manuscript*.

Lopes de Melo, R. (2018): "Firm wage differentials and labor market sorting: Reconciling theory and evidence," *Journal of Political Economy*, 126, 313–346.

Mendes, R., G. J. van den Berg, and M. Lindeboom (2010): "An Empirical Assessment of Assortative Matching in the Labor Market," *Labour Econ.*, 17, 919–929.

Mortensen, D. (2003): *Wage dispersion: why are similar workers paid differently?*, MIT press.

Rockoff, J. E. (2004): "The impact of individual teachers on student achievement: Evidence from panel data," *American economic review*, 94, 247–252.

Shimer, R. and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–369.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019): "Firming up inequality," *The Quarterly Journal of Economics*, 134, 1–50.

Sorkin, I. (2018): "Ranking firms using revealed preference," *The Quarterly Journal of Economics*, 133, 1331–1393.

Woodcock, S. D. (2008): "Wage differentials in the presence of unobserved worker, firm, and match heterogeneity," *Labour Economics*, 15, 771–793.

——— (2015): "Match effects," *Research in Economics*, 69, 100–121.

ONLINE APPENDIX

A Construction of event study data

In this section we describe the procedure we employ to go from an unbalanced panel of data over T years to an event study format at the spell level, with earnings before and after a move for movers, and one earning per spell for stayers.

1. **Original data:** The raw data across countries contains the variables (worker ID, firm ID, year, log earnings, spell length information). A unique row of data is defined by a (worker ID, employer ID, year) triplet. The spell length information has a different level of precision in different countries; for example, in Sweden the data has monthly spell information, the US has no spell information, and Italy has the number of days worked.

2. **Select largest earning employer:** As is common in the literature, in the event that a worker receives earnings from multiple firms within a given year, we start by selecting the (employer ID) within each (worker ID, year) associated with the highest annual earnings.

3. **Construct log-earnings measures:** We construct an earnings measure as the reported yearly earnings divided by the reported spell length. In the US, this does not change the measure in any way since the reported spell length is the same for all spells. In other countries we get a measure of monthly-earnings or daily-earnings respectively.

4. **Residualize log-earnings measures:** We residualize log earnings using OLS regression on calendar year indicators and a third-order polynomial in age. Following Card et al. (2018), the age profile is restricted to be flat at age 40.

5. **Collapse years into spells:** We assign a unique (spell ID) to each time-consecutive sequence of (worker ID, employer ID) pairs. We collapse the data by taking the mean of the residualized log-earnings within each spell ID. The resulting data has variables (worker ID, employer ID, spell ID, begin year of spell, end year of spell, log-earnings). A unique row of data is defined by a

(worker ID, spell ID) pair, or alternatively, a unique (worker ID, begin year of spell) pair.

6. **Extract stayer spells and mover spell pairs:** We collect all workers with only one spell in a dataset of stayers with (worker ID, employer ID, log-earnings, begin year of spell, end year of spell). Next, we collect all pairs of consecutive spells into a movers event-study dataset where the variables are (worker ID, employer ID 1, employer ID 2, log-earnings 1, log-earnings 2). Employer ID 1 and employer ID 2 are the employer identifiers at two consecutive spells for a given worker. These employers ID's are different by construction. Log-earnings 1 is the mean log-earnings at employer ID 1, before the job change, and log-earnings 2 is the mean log-earnings at the second employer. Employer ID 1 and employer ID 2 are defined in chronological order based on spell begin year.

7. **Weighting used in variance decompositions:** We compute the variance decompositions weighted by person-event as constructed in the previous step. This means that each move is counted once and each stayer is counted once. Given that in most of our samples individuals rarely have more than one move, this is almost identical to weighting by individuals.

B Estimation and computation

In what follows we describe the approach when working with an event-study data format. This means that each worker i is either a stayer with one log-earnings (at the only employer), or he is a mover with at most two log-earnings (one at the employer before the move, and one at a different employer after the move). An advantage of this data structure, relative to other panel data formats, is that it does not require the researcher to make assumptions about serial correlation within job spells.

Estimation of FE-HO. We follow Andrews et al. (2008). The first step in the estimation procedure is to extract the variance σ^2 of the residual. As noted in the text we use the following expression which provides an unbiased estimator under

homoskedasticity:

$$\widehat{\sigma}^2 = (NT - N - J)^{-1} Y'(I - A(A'A)^{-1}A')Y.$$

Importantly, job stayers do not contribute to the estimation of this variance since they only have a single spell observation per individual. This is because the data is in event-study form, if this was not the case one should worry about the fact that the formula assumes away serial correlation within job spells.

The next step is to compute the trace formula. When the design matrix A is not too large, we directly invert the matrix and compute:

$$\widehat{\text{Bias}}_Q^{\text{FE-HO}} = \widehat{\sigma}^2 \,\text{Trace}\left((A'A)^{-1}Q\right).$$

Estimation of FE-HO: Approximation. When the design matrix is too large to be fully inverted we rely on trace approximation methods. To be precise, we will use the Hutchinson stochastic trace estimator introduced in Hutchinson (1990), and proposed in the present context in Gaure (2014) and Kline et al. (2020), whereby the trace is approximated by

$$T_p = \frac{1}{p}\sum_{i=1}^{p} r_i'(A'A)^{-1}Qr_i,$$

where the r_i are i.i.d. Rademacher random vectors. This procedure only requires solving p linear systems, instead of trying to invert the matrix. It can be easily parallelized and in practice only a few draws seem to be sufficient to approximate the trace well.

Estimation of FE-HE. We refer to Kline et al. (2020) for a full description of their approach. Here we first outline the method while abstracting from computational feasibility concerns. The first step requires computing the leverage coefficients for each spell observation (i,t). This is done by computing:

$$\widehat{\sigma}_{it}^2 = \frac{Y_{it}\left(Y_{it} - \widehat{\alpha}_i - \widehat{\psi}_{j(i,t)}\right)}{1 - P_{it,it}},$$

where

$$P_{it,it} = A_{it} \left(A'A\right)^{-1} A'_{it}.$$

This expression however does not recover the $\widehat{\sigma}^2_{it}$ for the stayers since they only have one spell-observation. In order to be able to compute the trace correction for the covariance in a sample that includes both stayers and movers, we then make an homogeneity assumption that σ^2_{it} for stayers is equal to the average among movers at the same firm $j(i,t)$; that is,[21]

$$[\widehat{\sigma}^2_{it}]^{stayer} = \widehat{\mathbb{E}}_{i'} \widehat{\sigma}^2_{i't} \text{ for movers } i' \text{ in } j(i,t) \ .$$

Next, we construct the trace correction expression

$$\text{Trace} \left[A \left(A'A\right)^{-1} Q \left(A'A\right)^{-1} A' \widehat{\Omega}(A) \right],$$

where $\widehat{\Omega}(A) = \text{diag}[\widehat{\sigma}^2_{it}]$. We compute this formula directly whenever inverting the matrix $A'A$ is computationally feasible.

Estimation of FE-HE: Approximation. There are two computational bottlenecks when computing the FE-HE estimator. One is the computation of the trace expression, for which we rely on the same Hutchinson trace estimator described above. This approximation performs very well in our experience.

The second computational bottleneck is the computation of $P_{it,it}$, which requires effectively inverting the $A'A$ matrix. This expression does not benefit from the same aggregation property that computing the trace does. Indeed, the $P_{it,it}$ enter the expression of $\widehat{\sigma}^2_{it}$ as inverses. This is a difficult computational problem that is actively researched (Drineas et al., 2012). We decided to apply the procedure described in the computational appendix of Kline et al. (2020). Since we have $P_{it,it} = A_{it} \left(A'A\right)^{-1} A'_{it}$, if we could solve for Z in

$$(A'A)Z = A',$$

[21] As an alternative one could consider the following. First, compute the variance of firm effects in differences using movers and re-weight. Second, compute the covariance among movers using the leave-one-out procedure. Finally, compute the covariance for the stayers by using the covariance of their log-earnings with the estimated firm effects.

we would simply get $P_{it,it} = A'_{it} Z_i$. We draw a set of p random vectors r_i as in the Hutchinson approach, and to combine them into a matrix R_p with p columns, and solve instead

$$(A'A)\tilde{Z} = (R_p A)',$$

and use $\tilde{P}_{it,it} = A'_{it}\tilde{Z}_i$. We thus use the following approximation:

$$\tilde{P}_{it,it} = A'_{it}(A'A)^{-1}A'R'_p,$$

which requires solving only p linear system instead of inverting $A'A$ fully.

In practice, using a small p tends to give some estimates $\tilde{P}_{it,it}$ that are not strictly less than 1. Since $(1-P_{it,it})$ enters in the denominator of $\widehat{\sigma}^2_{it}$, this can cause unbounded $\widehat{\sigma}^2_{it}$'s. We choose to increase p until all $\tilde{P}_{it,it}$'s are < 1. This requires p to be in the order of thousands.

Estimating firm groups in CRE. Let us now describe how we estimate the firm groups that we use to build the CRE specification. Accounting for the groups allows one to correlate worker and firm effects to mobility patterns, as we will explain in the next paragraph. To estimate the firm grouping $\{k_j,\, j = 1, ..., J\}$, we follow Bonhomme et al. (2019) and cluster firms together based on earnings information. For example, using mean log-earnings one can estimate the partition by minimizing

$$\sum_{j=1}^{J} n_j(\overline{Y}_j - \mu(k_j))^2,$$

with respect to $\mu(1), ..., \mu(K)$ and $k_1, ..., k_J$, where n_j is firm size, and $\overline{Y}_j$ is the mean log-earnings in firm j. In practice we add information beyond means by including the full earnings distribution function, evaluated at a grid of 20 points (20 percentiles of the overall earnings distribution). For computation we use Lloyds' algorithm for k-means, with 30 starting values. Consistency of k-means is not straightforward to establish in this context, due to the presence of within-k firm heterogeneity. In single-agent panel data, Bonhomme et al. (2017) provide conditions for consistency and asymptotic normality of functions of the heterogeneity such as variance components as K tends to infinity together with the sample size.

CRE specification. Specifying the random-effects model consists in listing the restrictions that we impose on the vector $\mu(A)$ and the square matrices $\Sigma(A)$ and $\Omega(A)$. $\Omega(A)$ captures the error structure of the residuals across observations and has a number of rows equal to the number of observations. $\mu(A)$ and $\Sigma(A)$ describe the mean and variance of γ, and have respective length and number of rows equal to the number of workers plus the number of firms.

To be exhaustive, we need to specify how each entry in these matrices and vectors depends on A. To do so, we note that the γ vector contains three distinct types of elements: workers with only one employer, workers with multiple employers (i.e., movers), and firms. We describe the specification of $\mu(A)$ and $\Sigma(A)$ by listing the elements of $\mu(A)$ and $\Sigma(A)$ for each of these three types of entries. Throughout, we assume the data is in event study format, and hence movers will have exactly two employers. We also make use of a firm grouping structure, where k_j denotes the group of firm j and we write $k_{it} = k_{j(i,t)}$ to simplify the notation.

We assume that $\mu(A)$ does not depend on worker and firm identities beyond firm groups. We denote

$$\begin{aligned}
&\mathbb{E}[\alpha_i \,|\, A] = \mathbb{E}[\alpha_i \,|\, k_{i1}] = \mu_\alpha(k_{i1}) \text{ for stayers,} \\
&\mathbb{E}[\alpha_i \,|\, A] = \mathbb{E}[\alpha_i \,|\, k_{i1}, k_{i2}] = \mu_\alpha(k_{i1}, k_{i2}) \text{ for movers,} \\
&\mathbb{E}[\psi_j \,|\, A] = E[\psi_j \,|\, k_j] = \mu_\psi(k_j).
\end{aligned}$$

The matrix $\Sigma(A)$ consists of variances and covariances of worker effects and firm effects. We assume that $\Sigma(A)$ does not depend on worker and firm identities beyond firm groups. We denote, for any firm j,

$$\text{Var}[\psi_j \,|\, A] = \text{Var}[\psi_j \,|\, k_j] = \Sigma_{\psi\psi}(k_j).$$

For the off-diagonal terms, we assume that $\text{Cov}[\psi_j, \psi_{j'} | k_j, k_{j'}] = 0$ for $k_j \neq k_{j'}$ and leave the covariance within unrestricted. In estimation we do not estimate within-group covariances. It is important to also note that this does not restrict the covariance at the group level, since the $\mu_\psi(k)$ are unrestricted. Next, for any firm j and

any two movers i and i' we denote:

$$\begin{aligned}
\text{Cov}[\psi_j, \alpha_i \,|\, A] &= \text{Cov}[\psi_j, \alpha_i \,|\, j, j(i,1), j(i,2)] \\
&= \mathbf{1}\big[j(i,1){=}j \text{ or } j(i,2){=}j\big]\Sigma^m_{\alpha\psi}(k_j), \\
\text{Cov}[\alpha_i, \alpha_{i'} \,|\, A] &= \text{Cov}[\alpha_i, \alpha_{i'} | j(i,1), j(i,2), j(i',1), j(i',2)] \\
&= \mathbf{1}\big[j(i,1){=}j(i',1)\big]\Sigma^m_{\alpha\alpha'}(k_{j(i,1)}) + \mathbf{1}\big[j(i,2){=}j(i',2)\big]\Sigma^m_{\alpha\alpha'}(k_{j(i,2)}) \\
&\quad + \mathbf{1}\big[j(i,2){=}j(i',1)\big]\Sigma^m_{\alpha\alpha'}(k_{j(i,2)}) + \mathbf{1}\big[j(i,1){=}j(i',2)\big]\Sigma^m_{\alpha\alpha'}(k_{j(i,1)}).
\end{aligned}$$

For any firm j and any two stayers i and i' we denote

$$\begin{aligned}
\text{Cov}[\psi_j, \alpha_i \,|\, A] &= \text{Cov}[\psi_j, \alpha_i \,|\, j, j(i,1)] = \mathbf{1}\big[j(i,1){=}j\big]\Sigma^s_{\alpha\psi}(k_j), \\
\text{Cov}[\alpha_i, \alpha_{i'} \,|\, A] &= \text{Cov}[\alpha_i, \alpha_{i'} \,|\, j(i,1), j(i',1)] = \mathbf{1}\big[j(i,1){=}j(i',1)\big]\Sigma^s_{\alpha\alpha'}(k_{j(i,1)}).
\end{aligned}$$

For any given stayer i and any given mover i' we denote:

$$\begin{aligned}
\text{Cov}[\alpha_i, \alpha_{i'} \,|\, A] &= \text{Cov}[\alpha_i, \alpha_{i'} \,|\, j(i,1), j(i',1), j(i',2)] \\
&= \mathbf{1}\big[j(i,1){=}j(i',1)\big]\Sigma^{sm}_{\alpha\alpha'}(k_{j(i,1)}) + \mathbf{1}\big[j(i,1){=}j(i',2)\big]\Sigma^{sm}_{\alpha\alpha'}(k_{j(i,1)}).
\end{aligned}$$

Finally, we let the diagonal along workers unspecified since we will be focusing on the variance of firm effects and the covariance between worker and firm effects.[22]

As a reminder, the approach in Woodcock (2008) would set $\mu_\alpha(k){=}\mu_\alpha$, $\mu_\psi(k){=}\mu_\psi$, and $\Sigma_{\psi\psi}(k){=}\Sigma_{\psi\psi}$, as well as $\Sigma^m_{\alpha\psi}(k){=}\Sigma^s_{\alpha\psi}(k){=}\Sigma^m_{\alpha\alpha'}(k){=}\Sigma^s_{\alpha\alpha'}(k){=}\Sigma^{sm}_{\alpha\alpha'}(k){=}0$. Based on this specification, Woodcock focused on posterior estimates.

CRE estimation. Here we describe how we estimate the quantities that we use to reconstruct our two main parameters of interest (that is, the variance of firm effects and the covariance), as presented in equation (12). This involves the vector $\mu(A)$ and a subset of the elements in $\Sigma(A)$.

[22] A natural specification would be to allow for the variance of the worker effects of stayers to be group-specific and for the variance of the worker effects of movers to depend on the group pairs.

First we estimate all elements in $\mu(A)$ as

$$\min_{\mu_\alpha(k,k'),\mu_\alpha(k),\mu_\psi(k)} \sum_{i\,:\,stayer} \Big(Y_{i1} - \mu_\psi(k_{i1}) - \mu_\alpha(k_{i1})\Big)^2 + \sum_{i\,:\,mover} \Big(Y_{i1} - \mu_\psi(k_{i1}) - \mu_\alpha(k_{i1},k_{i2})\Big)^2 + \sum_{i\,:\,mover} \Big(Y_{i2} - \mu_\psi(k_{i2}) - \mu_\alpha(k_{i1},k_{i2})\Big)^2.$$

Next, it turns out that the elements in $\Sigma(A)$ enter equation (12) only through the following group aggregates. Specifically we define for $(t,t',p) \in \{1,2\}^3$ and compute:

$$C^m_{tt'}(p) = \widehat{\mathbb{E}}_{(i,i')\in S^m_p}\left[\Big(Y_{it} - \mu_\alpha(k_{i1},k_{i2}) - \mu_\psi(k_{it})\Big)\Big(Y_{i't'} - \mu_\alpha(k_{i'1},k_{i'2}) - \mu_\psi(k_{i't'})\Big)\right],$$

where the set S^{m}_p of pairs of workers consists of movers leaving the same firm and moving to a different firm group, or alternatively moving to the same firm and coming from two different firm groups; that is,

$$S^m_p = \{(i,i'{\neq}i) \text{ movers, s.t. } j(i,p){=}j(i',p), k_{i,-p}{\neq}k_{i',-p}, k_{i,-p}{\neq}k_{i,p}, k_{i',-p}{\neq}k_{i',p}\}.$$

Similarly, we define for $(t',p) \in \{1,2\}^2$ and compute:

$$C^s_{t'}(p) \;=\; \widehat{\mathbb{E}}_{(i,i')\in S^s_p}\left[\Big(Y_{it} - \mu_\alpha(k_{i1}) - \mu_\psi(k_{i1})\Big)\Big(Y_{i't'} - \mu_\alpha(k_{i'1},k_{i'2}) - \mu_\psi(k_{i't'})\Big)\right],$$

where

$$S^s_p = \{(i,i'{\neq}i), i \text{ stayer}, i' \text{ mover, s.t. } j(i,1){=}j(i',p), k_{i',-p}{\neq}k_{i1}\}.$$

To see the mapping between the sufficient elements of $\Sigma(A)$ in equation (12) and the previously defined group aggregates, note that:

$$\begin{aligned}
C^m_{22}(1) = C^m_{11}(2) &= \widehat{\mathbb{E}}_k\big[\Sigma^m_{\alpha\alpha'}(k)\big],\\
C^m_{12}(1) = C^m_{12}(2) &= \widehat{\mathbb{E}}_k\big[\Sigma^m_{\alpha\alpha'}(k) + \Sigma^m_{\alpha\psi}(k)\big],\\
C^m_{11}(1) = C^m_{22}(2) &= \widehat{\mathbb{E}}_k\big[\Sigma_{\psi\psi}(k) + \Sigma^m_{\alpha\alpha'}(k) + 2\Sigma^m_{\alpha\psi}(k)\big],
\end{aligned}$$

where $\widehat{\mathbb{E}}_k$ denote means, weighted by group sizes. In turn, the covariances based on

combinations of stayers and movers give:

$$C_2^s(1) = C_1^s(2) = \widehat{\mathbb{E}}_k\big[\Sigma_{\alpha\alpha'}^{sm}(k) + \Sigma_{\alpha\psi}^{m}(k)\big],$$
$$C_1^s(1) = C_2^s(2) = \widehat{\mathbb{E}}_k\big[\Sigma_{\psi\psi}(k) + \Sigma_{\alpha\alpha'}^{sm}(k) + \Sigma_{\alpha\psi}^{s}(k) + \Sigma_{\alpha\psi}^{m}(k)\big].$$

Lastly, given the estimated μ's and C's we construct the variance components appearing in equation (12).

CRE posterior. Under an additional joint normality assumption of γ and ε given A, a posterior estimator $\widehat{V}_Q^{\mathrm{P}}$ of V_Q is given by the posterior mean of $\gamma' Q \gamma$ in the Gaussian model; that is:

$$(\widehat{\Sigma}(A)^{-1}\widehat{\mu}(A) + A'\widehat{\Omega}(A)^{-1}Y)'\widehat{B}(A)^{-1}Q\widehat{B}(A)^{-1}(\widehat{\Sigma}(A)^{-1}\widehat{\mu}(A) + A'\widehat{\Omega}(A)^{-1}Y) + \mathrm{Trace}(\widehat{B}(A)^{-1}Q),$$

where $\widehat{B}(A) = \widehat{\Sigma}(A)^{-1} + A'\widehat{\Omega}(A)^{-1}A$. Relative to the main CRE estimator, we need all the elements of $\widehat{\Sigma}(A)$, and hence specify those by imposing additional zeros and modeling the entire diagonal. There are two computational challenges. First, $\widehat{\Sigma}(A)$ is a non-sparse matrix since we model covariances between worker effects and firm effects. Second, implementation requires computing the inverse of the matrix in the trace expression. This second challenge is as for the FE-HO estimator. In the paper we focus on the computation of the posterior estimator for the variance of firm effects. This only involves the part of $\widehat{\Sigma}(A)$ between firms, which is diagonal. We approximate the trace using the Hutchinson approach, as we do for FE-HO.

C Comparisons to existing work

In this section, we compare the results obtained from the methods we use to those obtained in previous studies.

C.1 Italian data

We first compare our results on the Italian data to those from the May 2020 version of Kline et al. (2020). Rather than our baseline sample selection (described in Section 2), we use their replication code to construct a sample as similar to theirs as possible. A key difference from our baseline analysis is that we now focus only on the years 1999 and 2001. Comparing descriptive statistics of our replication sample in row 3 of Table D5 to those reported in Table 1 of Kline et al. (2020), we find that the sample counts for number of observations, movers, and firms are nearly identical, and the estimates of the total variance of daily wages are very close (0.199 compared to 0.206).

In Table D5, we also apply the FE, FE-HO, and FE-HE estimators to our Kline et al. (2020) replication sample. Our implementation of the estimators differs from Kline et al. (2020) in two ways. First, we collapse yearly data to spell level data as described in Appendix A. Second, as in our main analysis, we use only one spell observation per stayer spell rather than assuming errors are uncorrelated over time within stayer spells. This choice matters for FE-HO, but not for FE-HE.

We find that these differences in implementation do not materially change the estimates when using our replication sample. Using our replication sample, we find similar results as in Kline et al. (2020). Concretely, we compare estimates from our replication sample in row 3 of Table D5 to Table 2 of Kline et al. (2020). The contribution of firm effects to wage inequality is 19% for FE, 15% for FE-HO, and 14% for FE-HE, while Kline et al. (2020) estimate 19% for FE, 14% for FE-HO, and 13% for FE-HE. We find that the contribution of sorting to wage inequality is 6% for FE, 15% for FE-HO, and 16% for FE-HE, while Kline et al. (2020) estimate 4% for FE, 11% for FE-HO, and 16% for FE-HE.

In sum, we conclude that our implementation of the estimators delivers similar results to Kline et al. (2020) on the Italian data once we use a similar sample.

C.2 US data

We now compare our results on the US tax data to those from Song et al. (2019) (Table 3, interval 2007-2013) and Sorkin (2018) (Table 1). We differ from their papers in three key dimensions. First, we consider the full sample of W-2 tax records, whereas Sorkin (2018) considers LEHD data (UI records) from 27 states and Song et al. (2019) consider SSA earnings records for men. Second, we use a minimum earnings threshold of 100% of the annualized minimum wage, whereas Sorkin (2018) and Song et al. (2019) set the minimum earnings threshold to 25% of the annualized minimum wage. Third, since we want to include small firms when studying inequality, we do not impose a minimum firm size restriction in the baseline results. By comparison Sorkin (2018) restricts the sample to firms with a minimum of 15 workers in each year (among workers who appear at least twice in the sample) and Song et al. (2019) restrict the sample to firms with at least 20 workers in each year.

To understand the impact of the restrictions made by Sorkin (2018) and Song et al. (2019), we now consider alternative minimum earnings and minimum firm size thresholds:

Minimum earnings threshold. As discussed in detail in Subsection 7.5, we examine how our results change when imposing minimum earnings thresholds ranging from 25% to 100% of the annualized minimum wage. When using the 25% threshold, we find that the variance of log earnings is 0.82 (see Table D4). This estimate is higher than the estimate of 0.67 reported in Table 1 of Sorkin (2018), and lower than the estimate of 0.92 reported in Table 3 of Song et al. (2019) for years 2007-2013. When increasing the minimum earnings threshold, the variance of log earnings must mechanically decline, and our baseline sample (100% minimum earnings threshold) has a substantially smaller variance of 0.41. However, the between-firm share of variance is nearly constant at about 40% across all minimum earnings thresholds, which is the same number reported in Table 2 of Song et al. (2019). Shifting attention to the AKM estimates, we find that the FE estimate of the share of earnings variation due to firm effects is somewhat decreasing in the minimum earnings threshold while the share due to sorting is strongly decreasing (see Figure D3).

Minimum firm size threshold. As discussed in detail in Subsection 7.2, we examine how our results change when imposing minimum firm size thresholds ranging from 0 to 50 workers. Neither the variance of log earnings nor the between-firm share of earnings variation changes materially with the minimum firm size threshold. However, the FE estimate of the share of earnings variation due to firm effects is decreasing in the firm size threshold while the share due to sorting is increasing (see Figure 11). When imposing a minimum firm size threshold of 20 workers, the FE estimate of the share of earnings variation due to sorting rises to 9% (see Table D4), which is close to the estimates by Sorkin (2018) and Song et al. (2019) of 10% and 12%, respectively.

Taken together, the results in Table D4 help explain how our estimates compare to Sorkin (2018) and Song et al. (2019). On the one hand, imposing a higher earnings threshold in the baseline sample tends to decrease our FE estimate of the contribution of firm effects to wage inequality and decrease our FE estimate of the contribution of sorting. On the other hand, imposing a lower firm size threshold in our baseline sample for the US tends to increase our FE estimate of the contribution of firm effects to wage inequality and decrease our FE estimate of the contribution of sorting. These differences partially offset each other for the contribution of firm effects, resulting in a FE estimate of the share of earnings inequality due to firm effects at 12 percent, in between the estimates of Sorkin (2018) and Song et al. (2019) at 14 and 9 percent respectively. However, both tend to decrease our FE estimate of the share due to sorting relative to the estimates of Sorkin (2018) and Song et al. (2019).

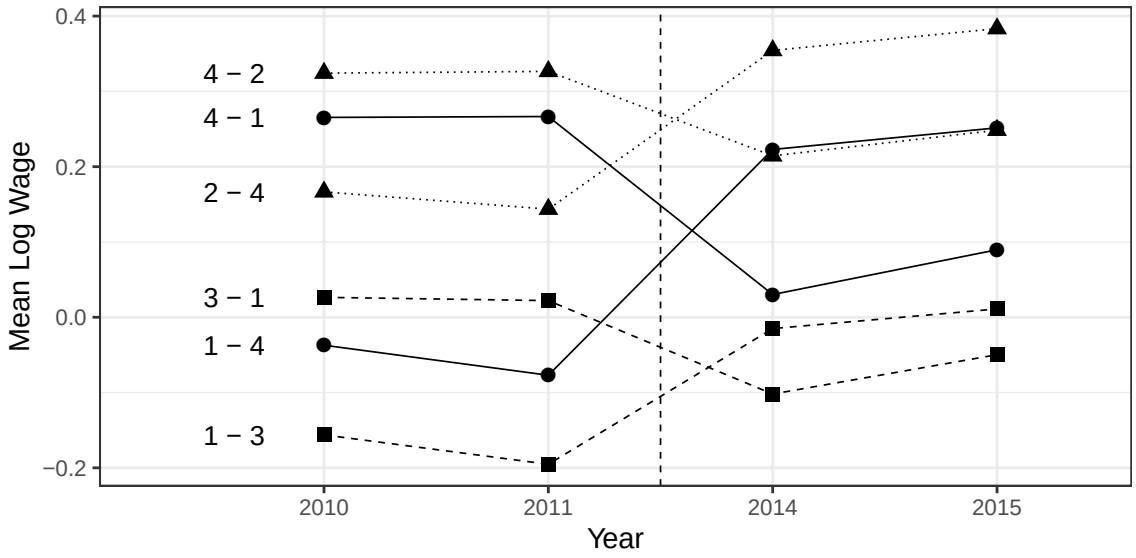
D Additional Tables and Figures

Table D1: Survey of Estimates in the Existing Literature

Paper	Country	Years	Total Var	Firm Effects	Sorting
Abowd et al. (1999)	France	1976-1987 ($\neq$ 1981, 1983)	0.269	87.0%	23.1%
Abowd et al. (2002)	France	1976-1987 ($\neq$ 1981, 1983)	0.269	30.1%	-13.6%
Abowd et al. (2002)	USA, WA	LEHD 1984-1993	0.278	19.2%	-1.0%
Abowd et al. (2004)$\star$	France	1976-1996	0.354	61.4%	-15.9%
Abowd et al. (2004)$\star$	USA	LEHD 1990-2000	0.800	16.3%	2.5%
Alvarez et al. (2018)	Brasil	1988-1992	0.750	21.3%	8.7%
Alvarez et al. (2018)	Brasil	1992-1996	0.750	22.7%	9.3%
Alvarez et al. (2018)	Brasil	1996-2000	0.690	23.2%	10.1%
Alvarez et al. (2018)	Brasil	2000-2004	0.620	21.0%	9.7%
Alvarez et al. (2018)	Brasil	2004-2008	0.530	17.0%	9.4%
Alvarez et al. (2018)	Brasil	2008-2012	0.470	14.9%	9.6%
Andrews et al. (2008)$\star$	Germany	LIAB 1993-1997, Bias Corr.	0.055	21.5%	-6.6%
Andrews et al. (2008)$\star$	Germany	LIAB 1993-1997, Not Corr.	0.057	23.5%	-9.0%
Bagger and Lentz (2019)	Denmark	1985-2003	0.097	14.4%	-1.0%
Card et al. (2013)	Germany	Universe, 1985-1991	0.137	18.2%	1.1%
Card et al. (2013)	Germany	Universe, 2002-2009	0.249	21.3%	8.2%
Card et al. (2018)$\star$	Portugal	2005-2009	0.275	22.8%	6.5%
Lopes de Melo (2018)$\star$	Brasil	1995-2005	0.601	29.9%	1.8%
Engbom and Moser (2018)	Brasil	1996-2000	0.690	23.2%	10.1%
Goldschmidt and Schmieder (2017)	Germany	IEB, 2008	0.205	26.7%	10.4%
Goldschmidt and Schmieder (2017)	Germany	IEB, 1985	0.132	21.9%	-1.9%
Goux and Maurin (1999)$\star$	France	1990-1992	0.181	12.9%	-6.1%
Goux and Maurin (1999)$\star$	France	1991-1993	0.157	30.2%	-2.5%
Goux and Maurin (1999)$\star$	France	1992-1994	0.154	65.3%	-24.0%
Goux and Maurin (1999)$\star$	France	1993-1995	0.151	19.6%	0.7%
Gruetter and Lalive (2009)	Austria	1990-1997	0.224	26.6%	-11.3%
Iranzo et al. (2008)	Italy	Manufacturing, 1981-1997	0.110	13.1%	6.4%
Kline et al. (2020)	Italy	1999-2001, AKM	0.184	19.4%	2.1%
Kline et al. (2020)	Italy	1999-2001, Homosk. Corr.	0.184	16.0%	5.3%
Kline et al. (2020)	Italy	1999-2001, Leave-out	0.184	13.0%	8.0%
Song et al. (2019)	USA	1980-1986	0.708	11.9%	2.3%
Song et al. (2019)	USA	1987-1993	0.777	9.7%	3.7%
Song et al. (2019)	USA	1994-2000	0.828	8.1%	4.6%
Song et al. (2019)	USA	2001-2007	0.884	8.5%	5.3%
Song et al. (2019)	USA	2007-2013	0.924	8.8%	5.8%
Sorkin (2018)	USA	LEHD 2000-2008	0.700	14.0%	5.0%
Woodcock (2015)	USA	2007-2013	0.410	19.5%	-0.5%

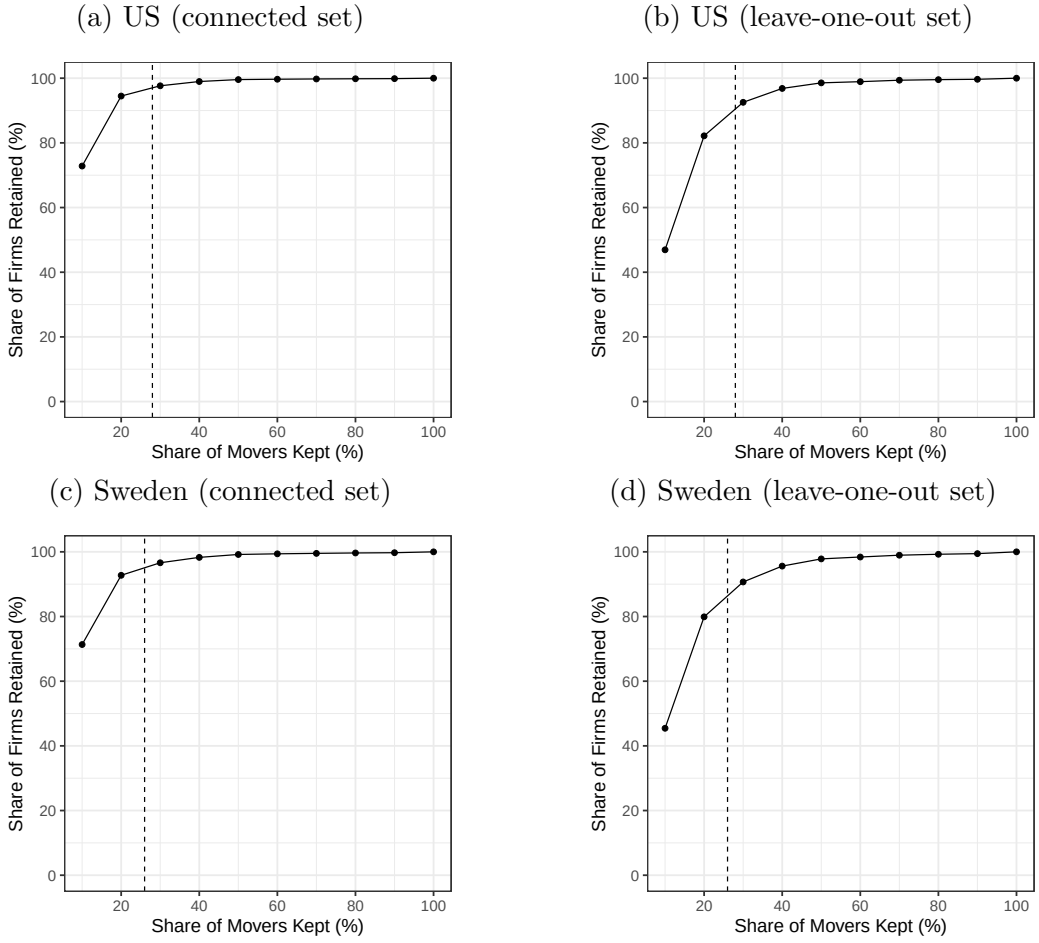
Notes: In this table, we provide a survey of estimates from a set of studies that estimate the contribution to earnings or wage inequality of firm effects and the sorting of workers to firms using the FE estimator. $*$ indicates that the total variance is not reported so we estimate it as $\text{Var}(\psi)+\text{Var}(\alpha)+2\text{Cov}(\psi,\alpha)$.

Figure D1: US Sample: Event Study around Moves



Notes: In this figure, we classify firms into four equally sized groups based on the mean earnings of stayers in the firm (with 1 and 4 being the group with the lowest and highest mean earnings, respectively). We compute mean log-earnings for the workers that move firms during 2012-2013. Note that the employer differs between event times 2012 and 2013, but we do not know exactly when the change in employer occurred. To avoid concerns over workers exiting and entering employment during these years, we do not display the transition years.

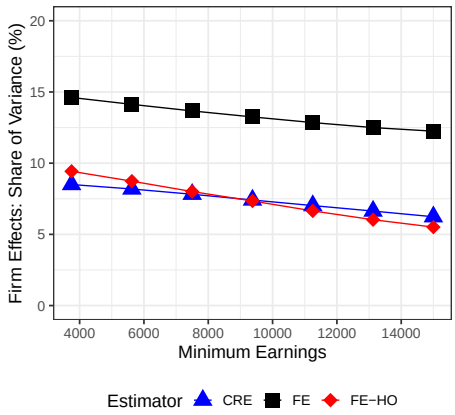
Figure D2: Retainment of Firms when Randomly Removing Movers in the US and Sweden



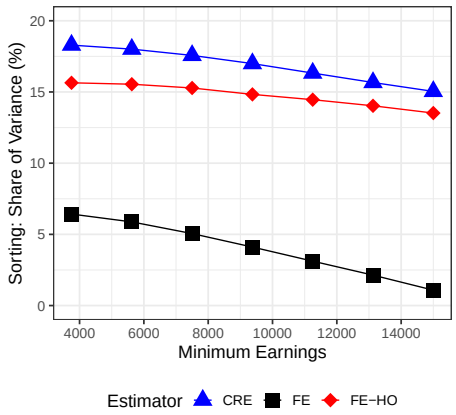
Notes: In this figure, we consider the subset of firms with at least 15 movers in the US (subfigures a and b) and Sweden (subfigures c and d). We randomly remove movers within each firm and compute the share of the initial firms retained in the connected (subfigures a and c) or leave-one-out (subfigures b and d) sets. The vertical dashed line approximates the point at which movers per firm in this sample matches movers per firm in the full sample.

Figure D3: Minimum Earnings Threshold for Defining Full-time Equivalence in the US

(a) Firm Effects: Share (%)



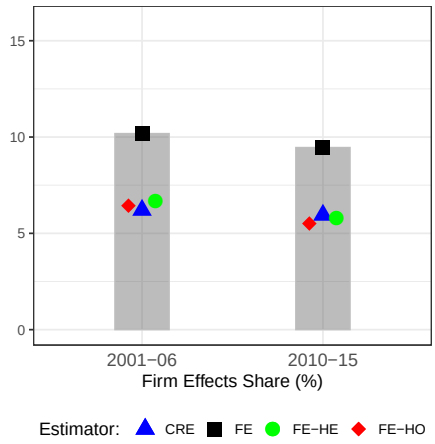
(b) Sorting: Share (%)



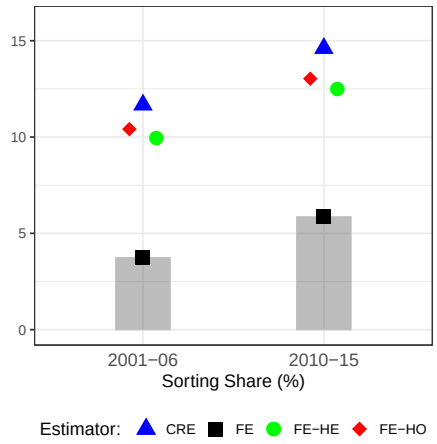
Notes: In this figure, we provide FE, FE-HO, and CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the US. We restrict the sample to workers with at least the annual earnings (at the highest-paying employer) indicated on the x-axis. We consider the connected set of firms for each restricted sample.

Figure D4: Firm Effects and Sorting in the US over Time: Leave-one-out Set

(a) Firm Effects



(b) Sorting



Notes: In this figure, we provide FE, FE-HO, FE-HE, and CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the US for the 6-year panels in 2001-2006 and 2010-2015. We consider the leave-one-out set of firms.

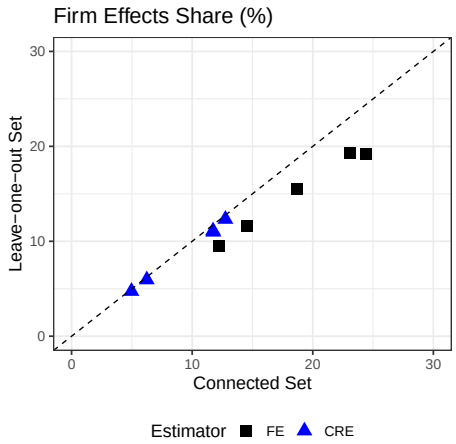
Table D2: Between-Firm Variation in the US (leave-one-out set)

Panel A.	**Total Decomposition**							
Years:	**2001-2006**				**2010-2015**			
Between Firm Share	33%				33%			
Within Firm Share	67%				67%			
Panel B.	**Share of Between Firm Variation**							
Years:	**2001-2006**				**2010-2015**			
	FE	**FE-HO**	**FE-HE**	**CRE**	**FE**	**FE-HO**	**FE-HE**	**CRE**
Firm Effects	31%	19%	20%	19%	24%	14%	15%	15%
Sorting	11%	31%	30%	35%	15%	34%	32%	38%
Segregation	58%	50%	50%	46%	61%	52%	53%	47%

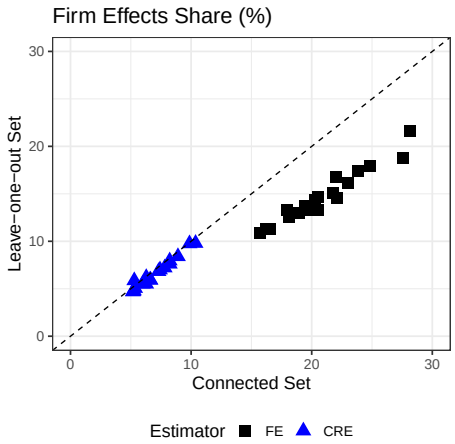
Notes: In this table, we provide FE, FE-HO, FE-HE, and CRE estimates of the contribution to between-firm earnings inequality of firm effects, sorting, and segregation in the US. We consider the leave-one-out set of firms. We compare the 6-year panel during 2001-2006 to the 6-year panel during 2010-2015.

Figure D5: Comparison of Estimators on Connected versus Leave-one-out Sets in the 5 Countries and 20 Small US States

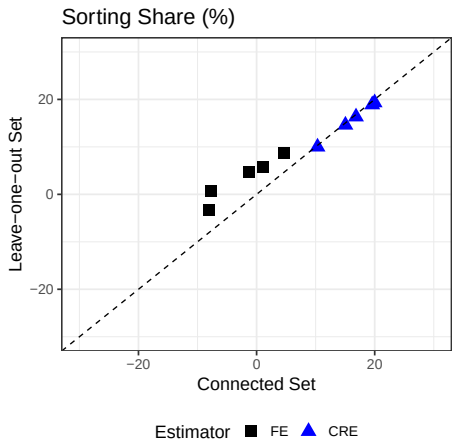
(a) Countries: Firm Effects



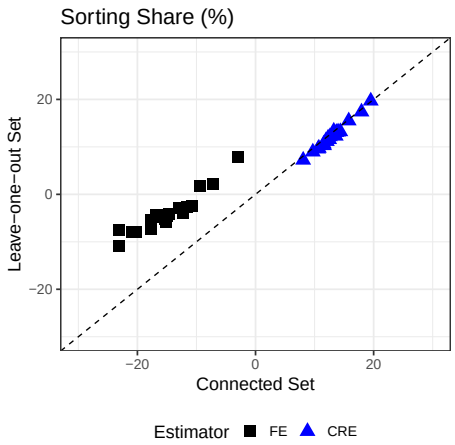
(b) States: Firm Effects



(c) Countries: Sorting



(d) States: Sorting

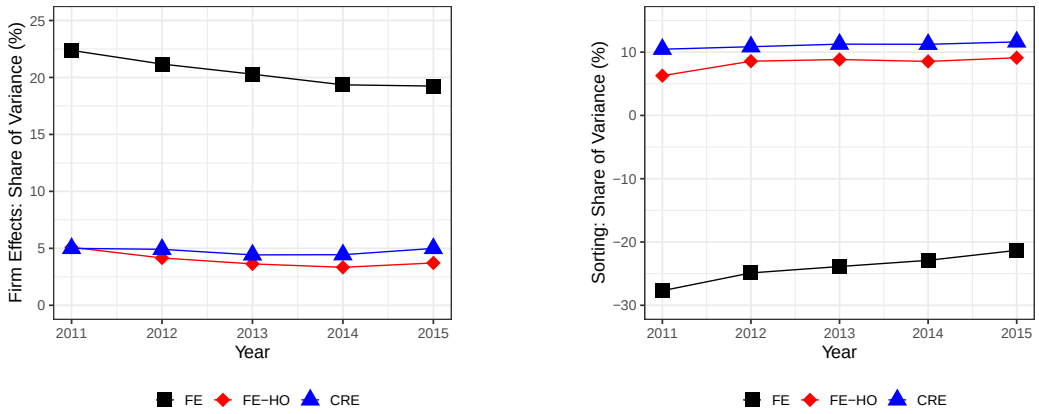


Notes: In this figure, we provide FE and CRE estimates of the contribution to earnings inequality of firm effects (subfigures a and b) and the sorting of workers to firms (subfigures c and d) in the 5 countries (subfigures a and c) and for the 20 smallest US states (subfigures b and d). We compare estimates on each connected set (x-axis) to estimates on each leave-one-out set (y-axis), so that the dashed 45-degree line represents equality between the estimates on the connected and leave-one-out sets.

Figure D6: Firm Effects and Sorting in the US: Short-Panel Estimation (Connected Set)

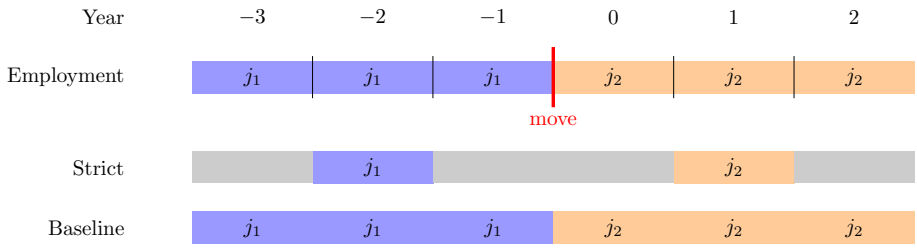
(a) Firm Effects

(b) Sorting



Notes: In this figure, we provide FE, FE-HO, and CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the US. We consider the connected set of firms, and compare estimates on each 2-year panel during 2010-2015 (the latter year of the 2-year panel is indicated on the x-axis).

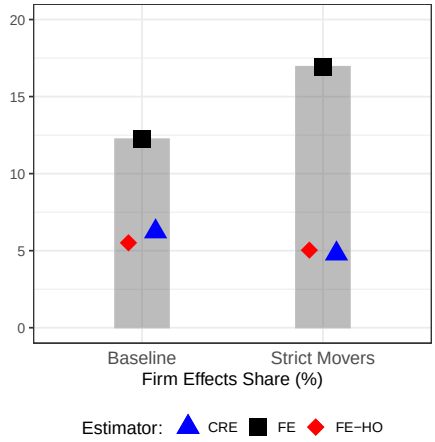
Figure D7: Visualizing Alternative Mover Definitions for the US



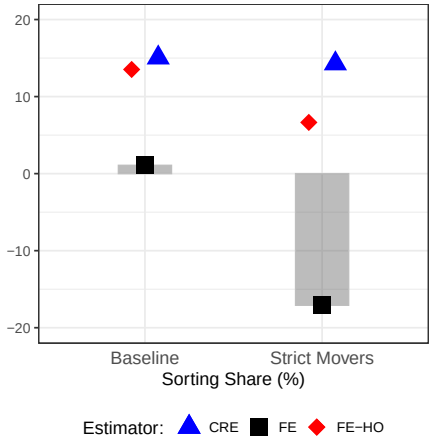
Notes: In this figure, we provide a diagram to help visualize the difference between the main definition of a mover ("Baseline") and the mover definition that uses only intermediate years within spells ("Strict").

Figure D8: Firm Effects and Sorting in the US over Mover Definitions (Connected Set)

(a) Firm Effects



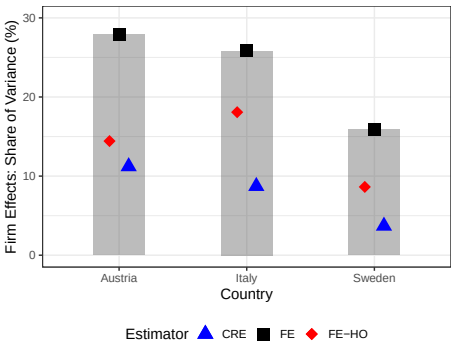
(b) Sorting



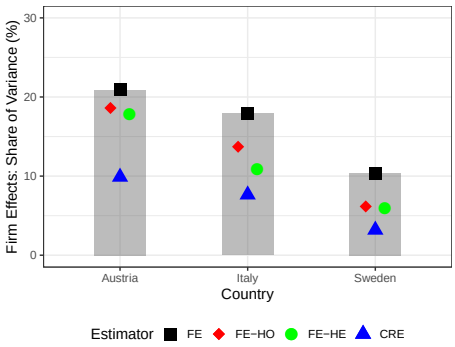
Notes: In this figure, we provide FE, FE-HO, and CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the US. We compare estimates using the baseline definition of movers and the strict definition of movers defined in the text.

Figure D9: Workers Employed the Full Year by a Single Firm

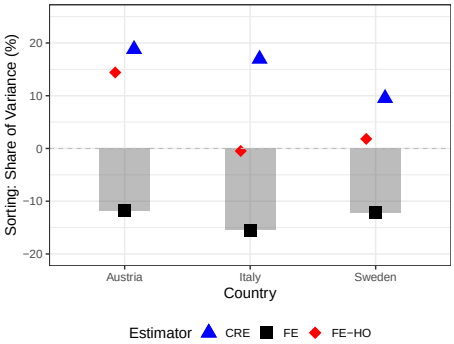
(a) Firm effects (connected set)



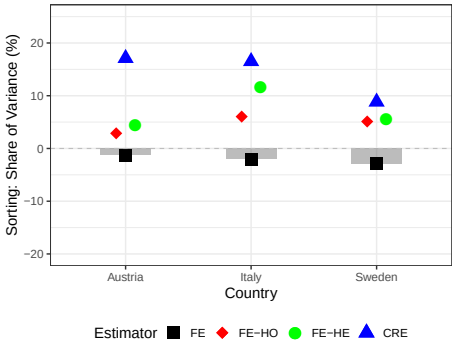
(b) Firm effects (leave-one-out set)



(c) Sorting (connected set)



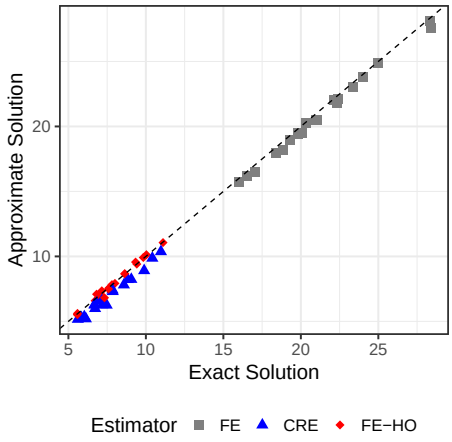
(d) Sorting (leave-one-out set)



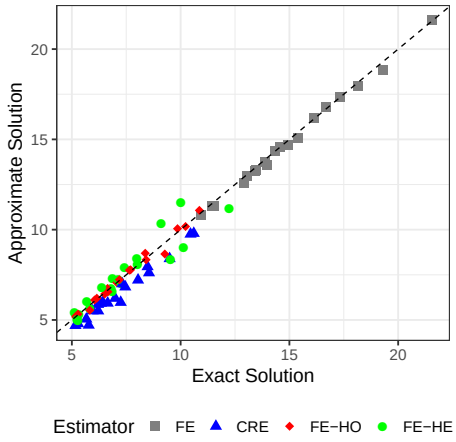
Notes: In this figure, we provide FE, FE-HO, and CRE estimates of the contribution to earnings inequality of firm effects (subfigures a and b) and the sorting of workers to firms (subfigures c and d) in Austria, Italy, and Sweden. We consider the connected (subfigures a and c) and leave-one-out (subfigures b and d) sets of firms. We consider only workers employed in the firm for the full calendar year.

Figure D10: Exact and Approximate Solutions: Firm Effects Variance (%) for the Small US States

(a) Connected Set

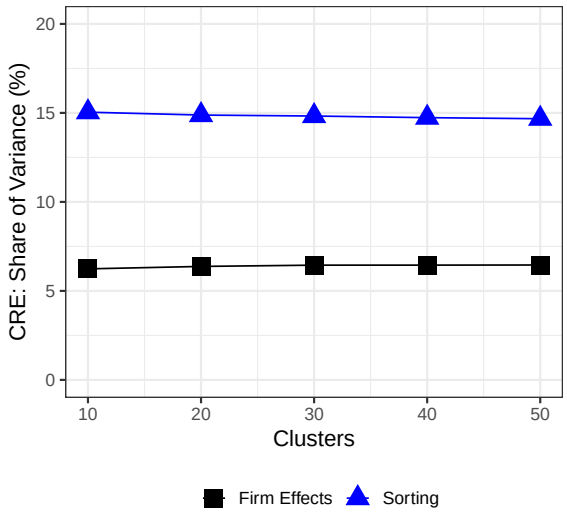


(b) Leave-one-out Set



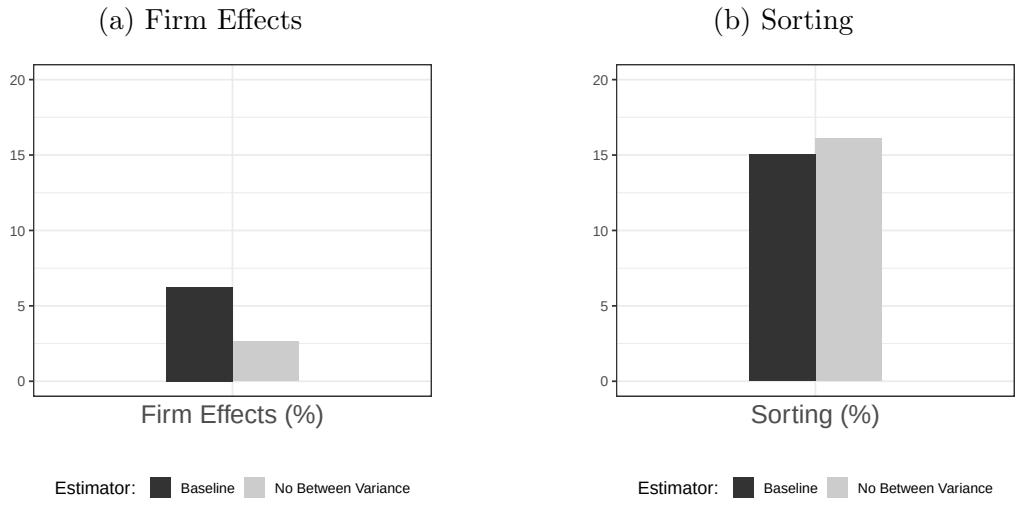
Notes: In this figure, we provide FE, FE-HO, and CRE estimates for the connected set (subfigure a) and FE, FE-HO, FE-HE, and CRE estimates for the leave-one-out set (subfigure b) of the contribution to earnings inequality of firm effects in the 20 smallest US states. We compare the exact solution (x-axis) and the approximate solution (y-axis) described in the text, so that the dashed 45-degree line represents equality between the exact and approximate solutions.

Figure D11: Number of Groups for CRE Estimates in the US (Connected Set)



Notes: In this figure, we provide CRE estimates of the contribution to earnings inequality of firm effects and the sorting of workers to firms in the US. We consider the connected set of firms, and vary the number of firm groups considered in the CRE estimation procedure (indicated on the x-axis).

Figure D12: Firm Effects and Sorting in the US over Type of CRE Estimator (Connected Set)



Notes: In this figure, we provide CRE estimates of the contribution to earnings inequality of firm effects (subfigure a) and the sorting of workers to firms (subfigure b) in the US. We compare the baseline CRE estimates to the posterior estimates for a random-effects specification that does not condition on firm groups.

Table D3: Baseline Results

			sample information					$Var(\psi)$				$2 \times Cov(\alpha, \psi)$			
	T	set	firms	workers	movers	var(y)	btw firm	FE	FE-HO	FE-HE	CRE	FE	FE-HO	FE-HE	CRE
US	6	connected	2,568	55,464	14,888	0.414	39.6%	12.2%	5.5%	-	6.2%	1.1%	13.5%	-	15.0%
	6	leave-out	1,689	52,484	13,968	0.416	38.8%	9.5%	5.5%	5.8%	5.9%	5.9%	13.0%	12.5%	14.6%
US	3	connected	1,241	36,826	4,252	0.436	38.2%	16.3%	4.1%	-	5.3%	-12.0%	11.7%	-	12.5%
	3	leave-out	670	33,031	3,645	0.440	37.6%	10.4%	4.3%	4.5%	5.0%	-0.8%	11.0%	10.6%	12.1%
Austria	6	connected	206	3,396	1,123	0.187	45.5%	18.7%	15.3%	-	11.7%	4.7%	10.5%	-	19.6%
	6	leave-out	140	3,240	1,055	0.182	43.7%	15.5%	12.7%	12.9%	11.1%	8.7%	13.5%	13.0%	18.9%
Austria	3	connected	117	2,845	387	0.183	43.7%	19.7%	12.1%	-	10.1%	-5.3%	9.3%	-	17.5%
	3	leave-out	68	2,604	336	0.178	41.8%	15.0%	10.7%	13.9%	9.2%	1.5%	9.7%	3.2%	16.2%
Italy	6	connected	92	1,111	379	0.167	46.1%	23.1%	17.5%	-	12.7%	-1.3%	8.7%	-	20.0%
	6	leave-out	61	1,034	346	0.168	44.8%	19.3%	15.8%	15.7%	12.3%	4.7%	11.1%	11.2%	19.3%
Italy	3	connected	54	864	148	0.176	44.9%	24.1%	15.7%	-	11.0%	-8.4%	7.7%	-	17.7%
	3	leave-out	30	755	121	0.181	43.5%	18.5%	14.6%	10.9%	10.2%	1.3%	8.8%	16.1%	17.2%
Norway	6	connected	114	1,286	556	0.239	47.2%	24.4%	13.9%	-	11.8%	-7.7%	11.3%	-	16.8%
	6	leave-out	78	1,199	519	0.236	45.8%	19.2%	12.5%	12.3%	11.0%	0.8%	12.6%	12.8%	16.3%
Norway	3	connected	63	986	203	0.229	44.5%	37.8%	14.9%	-	11.5%	-41.3%	2.5%	-	12.2%
	3	leave-out	37	856	175	0.227	42.6%	24.2%	12.1%	10.2%	10.3%	-16.7%	6.3%	10.1%	11.3%
Sweden	6	connected	63	1,921	608	0.164	31.6%	14.6%	8.2%	-	5.0%	-8.1%	3.9%	-	10.3%
	6	leave-out	52	1,850	596	0.164	30.9%	11.6%	7.8%	7.1%	4.7%	-3.2%	3.7%	5.0%	10.0%
Sweden	3	connected	42	1,497	237	0.161	31.3%	23.6%	11.6%	-	4.6%	-28.5%	-5.4%	-	9.0%
	3	leave-out	29	1,377	221	0.161	30.2%	15.5%	8.9%	7.4%	4.3%	-14.1%	-1.3%	1.5%	8.1%

Notes: In this table, we provide FE, FE-HO, FE-HE, and CRE estimates for the main results shown in the paper.

Table D4: Sample comparison for the US

	T	set	firms	workers	movers	var(y)	btw firm	$Var(\psi)$	$2 \times Cov(\alpha, \psi)$
US (baseline results)	6	connected	2,568	55,464	14,888	0.414	39.6%	12.2%	1.1%
Varying earnings threshold									
US (earnings threshold at $\$7,500$, 50% of min. wage)	6	connected	3,201	66,474	22,140	0.601	42.5%	13.7%	5.1%
US (earnings threshold at $\$3,750$, 25% of min. wage)	6	connected	3,489	72,228	27,325	0.823	42.9%	14.6%	6.4%
US (earnings threshold at $\$1,875$ 12.5% of min. wage)	6	connected	3,624	75,412	30,587	1.056	42.6%	15.2%	6.6%
Varying minimum firm size									
US (firm size ≥ 10)	6	connected	748	48,810	10,704	0.425	38.0%	8.2%	6.0%
US (firm size ≥ 20)	6	connected	360	43,140	8,361	0.434	38.1%	6.6%	8.4%
US (firm size ≥ 30)	6	connected	227	39,581	7,015	0.440	38.3%	5.9%	9.4%

Notes: In this table, we provide FE estimates for a range of sample construction rules. This focuses on 6 years of data and the connected set. See Appendix C for more details.

Table D5: Sample comparison for Italy

	sample information							$Var(\psi)$			$2 \times Cov(\alpha, \psi)$		
	T	set	firms	workers	movers	var(y)	btw firm	FE	FE-HO	FE-HE	FE	FE-HO	FE-HE
Italy (baseline results)	6	leave-out	61	1,034	346	0.168	44.8%	19.3%	15.8%	15.7%	4.7%	11.1%	11.2%
Italy (baseline sample selection)	3	leave-out	30	755	121	0.181	43.5%	18.5%	14.6%	10.9%	1.3%	8.8%	16.1%
Italy (sample construction of Kline et al. 2020)	2	leave-out	42	985	164	0.206	46.1%	19.4%	14.8%	14.4%	6.1%	14.9%	15.7%

Notes: In this table, we provide FE, FE-HO and FE-HE estimates for an alternative sample construction based on Kline et al. (2020) for comparison. We report results for the leave-out set. See Appendix C for more details.