Online Supplement: Firm and Worker Effect Estimation

Thibaut Lamadon[*] Magne Mogstad[†] Bradley Setzler[‡]

December 2020

In this supplemental appendix, we provide a more detailed analysis of how we identify and estimate the model of earnings inequality by Lamadon et al. (2020, hereafter LMS). We build up to the main specification, equation (14) of LMS, in three steps: First, we discuss the additive model of firm and worker fixed effects proposed by Abowd et al. (1999, AKM hereafter), providing an assessment of the biases that arise due to a limited number of movers. Second, we consider the non-additive extension proposed by Bonhomme et al. (2019, BLM hereafter) in which firm and worker effects interact, providing a number of checks on the reliability of the estimates of interaction parameters. These checks include a comparison between our estimates and the interaction effects that arise due to observed worker heterogeneity and a comparison against data on hourly wages instead of annual earnings. Third, we consider the time-varying firm effects extension proposed by LMS in which firm productivity shocks pass through to the earnings of workers.

To review, LMS assume the economy is composed of a large number of workers indexed by i and a large set of firms indexed by $j = 1, ..., J$. Each worker is employed by a firm at time t, $j(i,t)$, and each firm belongs to a market, $r(j)$. Let J_r denote the set of firms in market r. The model in equation (14) of LMS is as follows:

$$\mathbb{E}\left[\underbrace{w_{it} - \overbrace{\left(\frac{1}{1+\lambda\beta}\bar{y}_{rt} + \frac{\rho_r}{\rho_r + \lambda\beta}(y_{jt} - \bar{y}_{rt})\right)}^{\psi^a_{jt}}}_{w^a_{it}} \middle| j(i,t){=}j \in J_r\right] = \theta_j x_i + \psi_j$$

The observables are log earnings w_{it}, log value added y_{jt}, the market level mean of log earnings, $w_{rt} \equiv \mathbb{E}\left[w_{it} | j(i,t) \in J_r\right]$, and the market level mean of log value added, $y_{rt} \equiv \mathbb{E}\left[y_{it} | j(i,t) \in J_r\right]$. The parameters $(\beta, \rho_r, \alpha_r, \lambda)$ govern the pass through of value added shocks from firms to workers and are recovered in Section 5.3 of LMS. This supplemental appendix discusses how we identify and estimate worker fixed effects x_i, firm-worker complementarities θ_j, firm fixed effects ψ_j, and the pass through of productivity shocks from firms to workers, $\psi_{jt} - \psi_j$. See LMS for further details on the model and the derivation of LMS equation (14).

[*]Department of Economics, University of Chicago. E-mail: lamadon@uchicago.edu

[†]Department of Economics, University of Chicago; Statistics Norway; NBER; IFS. E-mail: magne.mogstad@gmail.com

[‡]Becker Friedman Institute for Economics, University of Chicago. E-mail: bradley.setzler@gmail.com

1 Two-way Fixed Effects and Limited Mobility Bias

To begin with, we consider a special case in which $\theta_j = 1$ and $\gamma_r = \Upsilon = 0$. The first restriction imposes a log additive structure on the earnings that worker i can expect to receive from working in firm j. Under this functional form, the worker fixed effect captures the (time-invariant) portable component of earnings ability, whereas the firm fixed effect can be interpreted as a firm-specific relative pay premium. The second restriction assumes there is no pass through of firm or market level shocks. As a result, the firm effects on earnings do not vary over time. By invoking these two restrictions, our statistical model of earnings reduces to the two-way (worker and firm) fixed effect model of AKM.

Under the restrictions $\theta_j = 1$ and $\gamma_r = \Upsilon = 0$, the variance of log earnings can be written:

$$Var(\log W_{it}) = \underbrace{Var(x_i + \mathcal{X}'_{it}b)}_{\text{Worker component}} + \underbrace{Var(\psi_{j(i,t)})}_{\text{Firm component}} + \underbrace{2Cov(x_i + \mathcal{X}'_{it}b, \psi_{j(i,t)})}_{\text{Sorting component}} + \underbrace{Var(v_{it})}_{\text{Residual}} \quad \text{(A.1)}$$

where the worker and firm components tell us how much of the variation in log earnings can be attributed to heterogeneity in worker and firm effects, respectively. The third component captures the contribution to earnings inequality from the sorting of workers to firms. The goal is to quantify these three components to draw inference about the determinants of earnings inequality in the U.S. economy. The decomposition includes both workers who move between firms and non-movers. However, the firm and worker effects are only separately identified within a connected set of firms that are linked by worker mobility. Consistent with previous work, we therefore restrict our sample of workers (including movers and non-movers) to those who work at a firm in the largest connected set in each time interval (2001-2008 and 2008-2015). In the U.S., this set covers more than 90 percent of the workers (see Appendix Table A.1).

1.1 Limited mobility bias

Even if the above restrictions hold, it is challenging to draw inference about the inequality contribution from firm effects and worker sorting. A key challenge is the incidental parameter bias caused by the large number of firm-specific parameters that are solely identified from workers who move across firms. The analysis of Andrews et al. (2008) suggests this limited mobility bias can be substantial. With few movers per firm, the firm component is biased upwards while the sorting component is biased downwards, with the size of the bias depending inversely on the degree of worker mobility among firms.

To get a better sense of the scope for limited mobility bias in the U.S. data, we would ideally apply the AKM estimator to alternative samples of workers and firms that are comparable except for the number of movers per firm. Figure 1 presents the results from such an analysis, suggesting that the variance of firm effects declines monotonically as the number of movers per firm increases. To construct this figure, we consider a subsample of firms with reasonably many movers; that is, at least 15 movers per firm over the period 2001-2008. Applying AKM to this subsample gives an estimate of the variance of firm effects of 6.7 percent. Next, we

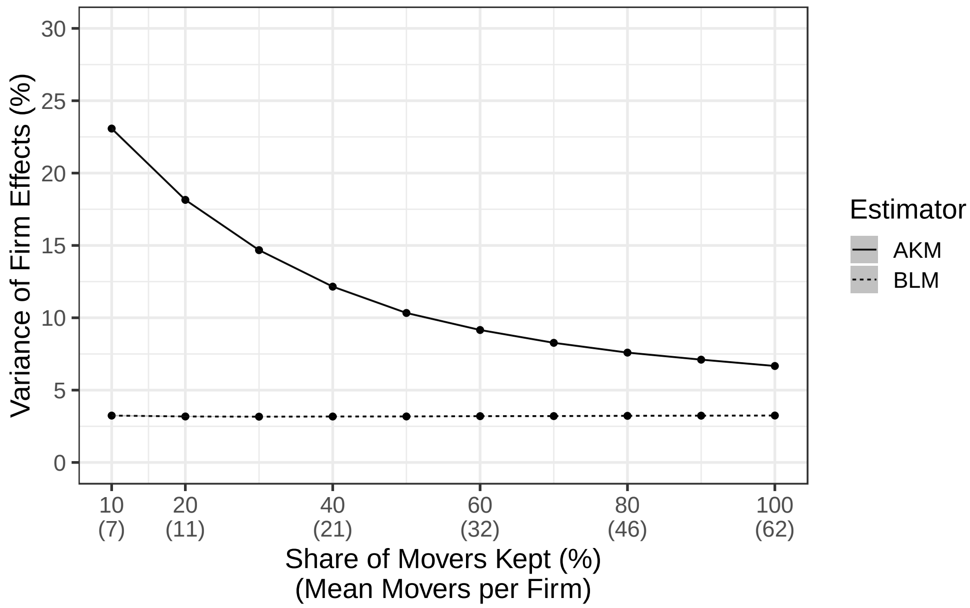


Figure 1: Empirical Characterization of Limited Mobility Bias

Notes: In this figure, we consider the subset of firms with at least 15 movers. We randomly remove movers within each firm and re-estimate the variance of firm effects using the AKM and BLM estimators. For each estimator, we repeat this procedure several times, and then take averages of the variance estimates across these repetitions. The procedure allows us to keep the connected set of firms approximately the same and examine the bias that results from having fewer movers available in estimation.

remove movers randomly within firms (keeping the connected set of firms approximately the same) before re-estimating the AKM model. The solid line displays the AKM estimates of the variance of firm effects after randomly removing movers. 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 several times as large (23 percent) if we only keep ten percent of the movers within each firm (on average, 7 movers per firm) as compared to what we obtained if we keep all the movers per firm (at a minimum 15 and, on average, 62 movers per firm). By way of comparison, there are around 18 movers per firm in the full estimation sample (which roughly corresponds to the number of movers per firm when randomly removing 40% of movers).

Until recently, the procedures for addressing limited mobility bias required strong and questionable assumptions about the covariance structure of the time-varying errors (see e.g. the discussion in Card et al., 2018). To address this shortcoming, BLM and Kline et al. (2020) propose approaches to address limited mobility bias that rely on a different or weaker set of assumptions.[1] The first approach reduces the firm heterogeneity to a finite number of types. BLM show how this approach can be used to alleviate the biases arising from low mobility rates.

[1] Borovickova and Shimer (2017) propose a different approach that redefines firm types as the average firm wage rather than the wage premium, and relies on independence restrictions to recover the variance decomposition for this alternative definition.

The second approach uses a version of the Jackknife method. Kline et al. (2020) show how this approach allows one to relax the homoskedasticity assumption in the bias correction procedure proposed by Andrews et al. (2008). Since it is computationally challenging to apply Andrews et al. (2008) and Kline et al. (2020) to very large data sets (as one needs to compute the trace of the inverse of the mobility matrix), our main analysis is based on the approach of BLM. As a robustness check, however, we below use a subset of the U.S. states to assess the sensitivity of the results to the choice of procedure for addressing limited mobility bias.

In Figure 1, the dotted line shows estimates of the variance of firm effects based on the procedure of BLM that addresses limited mobility bias. Firms are first classified into groups based on the empirical earnings distribution using the k-means clustering algorithm. The k-means classification groups together firms whose earnings distribution is most similar.[2] Then, in a second step, the worker effects and firm effects are estimated, restricting ψ_j to be the same for all firms of a given type. While the specification of BLM in Figure 1 assumes there exists 10 firm types, Appendix Figure A.1 shows the BLM estimates do not materially change if we instead allow for 20, 30, 40 or 50 firm types. Consistent with limited mobility bias, the BLM estimates are noticeably smaller than the standard AKM estimates in the samples with few movers. As expected, the AKM estimates become more similar to the BLM estimates when there is a large number of movers per firm, and thus, limited mobility bias should be small.

1.2 Estimates of the two-way fixed effects model

While the analysis in Figure 1 is useful to illustrate the scope for limited mobility bias, it does not offer estimates of firm effects for the entire connected set. In Table 1, we present results from the variance decomposition in equation (A.1) based on data for all firms and workers in the connected set (which includes both movers and non-movers). This table reports estimates of the worker, firm and sorting components as defined in equation (A.1).

Consider first Panel A of Table 1 where we present estimates from the AKM estimator for two different time periods (2001-2008 and 2008-2015) as well as pooled estimates where we combine the data from these time periods. The results show that the worker, firm and sorting components change little over time. Therefore, we focus attention on the pooled estimates. These results suggest that the firm effects explain around 9 percent of the variation in log earning, whereas worker sorting accounts for 5 percent. The correlation between firm effects and worker effects is only 0.1.

Next, consider Panel B of Table 1 where we report the BLM estimates. As discussed above, a possible advantage of the BLM estimator is that it addresses limited mobility bias. Once we correct for such bias we find that firm effects are very small in the U.S. labor market, accounting for only 3 percent of the variation in log earnings. Instead, a larger part of the earnings variation is explained by worker sorting. The correlation between firm effects and worker effects exceeds 0.4 once we correct for limited mobility bias. This finding suggests that sorting of better workers

[2] Here, we follow BLM. Concretely, we use a weighted k-means algorithm with 100 randomly generated starting values. We use the firms' empirical distributions of log-earnings on a grid of 10 percentiles of the overall log-earnings distribution.

Years:		2001-2008	2008-2015	Pooled
Panel A.			**AKM Estimation**	
Share explained by:				
i) Worker Effects	$Var(x_i)$	75%	75%	75%
ii) Firm Effects	$Var(\psi_{j(i)})$	9%	9%	9%
iii) Sorting	$2Cov(x_i, \psi_{j(i)})$	5%	6%	5%
Sorting Correlation:	$Cor(x_i, \psi_{j(i)})$	0.09	0.11	0.10
Panel B.			**BLM Estimation**	
Share explained by:				
i) Worker Effects	$Var(x_i)$	72%	72%	72%
ii) Firm Effects	$Var(\psi_{j(i)})$	3%	3%	3%
iii) Sorting	$2Cov(x_i, \psi_{j(i)})$	13%	14%	14%
Sorting Correlation:	$Cor(x_i, \psi_{j(i)})$	0.43	0.46	0.44

Table 1: Decomposition results using AKM and BLM

Notes: This table presents the decomposition of log earnings variation using the AKM and BLM estimators for two time intervals. The analysis uses both movers and non-movers.

to better firms is an empirically important feature of the U.S. labor market.

Our findings of small firm effects and strong sorting are at odds with recent work from the U.S. (Sorkin, 2018; Song et al., 2018) as well as many studies from other developed countries (Card et al., 2018). We argue the reason is that the existing literature do not properly address the concern over limited mobility bias.[3] This raises questions such as: How do our results from the AKM estimator compare to those reported in the existing literature? Are the bias-corrected estimates sensitive to the procedure used?

To examine the first question of how our AKM results compare to existing work, consider Song et al. (2018, using SSA data from all U.S. states) and Sorkin (2018, using LEHD data for a subset of states). Both studies apply the AKM estimator, finding that firm effects explain 9 percent (Song et al., 2018) and 14 percent (Sorkin, 2018) of the variation in log earnings. By comparison, our AKM estimates suggest that firm effects explain 9 percent of the variation in log earnings, which matches the estimate of Song et al. (2018) but is smaller than the estimate of Sorkin (2018). We highlight some differences in our sample restrictions. We only include workers with earnings above the full-time minimum wage threshold. By comparison, Song et al. (2018) and Sorkin (2018) include individuals who work part time as long as their annual

[3] Sorkin (2018) and Song et al. (2018) point out that limited mobility may bias their AKM estimates. In an attempt to investigate this issue, Sorkin (2018) also performs a few checks, including restricting the sample to large firms and splitting the sample in half on the basis of workers (which lets him compare results from two separate samples). Limited mobility bias, however, is about having few movers per firm, not small firms. Furthermore, the checks he performs involve significant changes in the composition of firms and workers in the estimation sample, in part because the connected set changes. Thus, it is not clear what, if any, conclusions one may draw about limited mobility bias from these checks.

earnings exceed 25 percent of the full-time minimum wage threshold. In Appendix Figure A.2, we investigate what happens if we use their earnings cutoff. The AKM estimates then suggest that firm effects explain 11 percent of the variation of log earnings, which is between the 9 percent estimate of Song et al. (2018) and the 14 percent estimate of Sorkin (2018). This figure also shows estimates for a range of alternative earnings cutoffs. While the total log earnings variance increases substantially as the earnings cutoff decreases, the share of variation explained by AKM firm effects is relatively stable.

To investigate the second question of the sensitivity of the bias-corrected estimates, we restrict attention to workers and firms from a set of smaller states. This is necessary because it is computationally challenging to apply the methods of Andrews et al. (2008) and Kline et al. (2020) to the entire U.S. data. Appendix Figure A.3 compares the results from these alternative procedures for correcting for limited mobility bias to the estimates from AKM and BLM. The conclusion is clear: Limited mobility bias leads to upward bias in the AKM estimate of the firm component and downward bias in the AKM estimate of the worker sorting component. On average across the states we consider, AKM suggests that firm effects explain more than 11 percent of the variation in log earnings. By contrast, the methods of Andrews et al. (2008) and Kline et al. (2020) suggest that firm effects explain about 5 percent, whereas the BLM method produces an estimate of firm effects around 2-3 percent. For additional discussion of limited mobility bias and a comparison of alternative estimation methods across countries and states, see Bonhomme et al. (2020).

1.3 Bias in worker effect variance estimates with measurement errors

LMS follow BLM in making a discrete heterogeneity assumption to recover firm-specific parameters ψ_j. However, for the worker-specific parameters x_i, LMS follow an approach closer to AKM than BLM. Like AKM, exogenous mobility delivers identification of x_i using that $x_i = \mathbb{E}\left[w_{it} - \psi_{j(i,t)} \middle| i\right]$. The sample counterpart, $\hat{x}_i \equiv \frac{1}{T}\sum_{t=1}^{T}\left(w_{it} - \psi_{j(i,t)}\right)$, is an unbiased but inconsistent estimator of x_i, and the variance of $\hat{x}_i$ will typically be an upward-biased estimator for $\mathrm{Var}(x_i)$. To formalize this bias, note that $w_{it} - \psi_{j(i,t)} = x_i + v_{it}$, so $\hat{x}_i = x_i + \frac{1}{T}\sum_{t=1}^{T} v_{it}$. To fix ideas, suppose $v_{it} = u_{it} + \varsigma_{it}$, where u_{it} is an i.i.d. transitory shock and $\varsigma_{it} = \sum_{t' \le t} \mu_{it}$ is a random walk process with i.i.d. shocks μ_{it}. While u_{it} is serially uncorrelated, ς_{it} is strongly serially correlated. Denote the variance of μ_{it} by σ_μ^2 and the variance of u_{it} by σ_u^2. Thus, $\mathrm{Var}(\hat{x}_i) = \mathrm{Var}(x_i) + \frac{1}{T}\sigma_u^2 + \frac{(T-1)(T-2)}{T^2}\sigma_\mu^2$. LMS Online Appendix Table A.4 suggests that $\sigma_\mu^2 \approx 0.01$ and $\sigma_u^2 \approx 0.01$. Thus, with $T = 8$, $\mathrm{Var}(\hat{x}_i) - \mathrm{Var}(x_i) \approx 0.008$, while with $T = 4$, $\mathrm{Var}(\hat{x}_i) - \mathrm{Var}(x_i) \approx 0.006$. Since we estimate that the variance of x_i is about 0.31, while the upward-bias is less than 0.01, this suggests a relatively minor role for bias in our estimate of $\mathrm{Var}(x_i)$.

2 Non-additivity and Complementarities

The assumption that $\theta_j = 1$ implies that all workers who move from firm j to j' will experience an earnings change of $\psi_{j'}$ - ψ_j, no matter their quality x_i. The absence of interactions between worker and firm effects rules out strong complementaries in production, as in Shimer and Smith (2000) and Eeckhout and Kircher (2011). In this section, we consider relaxing this assumption to allow for interactions between worker and firm effects, while maintaining the assumption that $\gamma_r = \Upsilon = 0$ from the previous section.

2.1 Informal assessment of non-additivity

An informal way to assess this log additive structure is to perform an event study of the earnings changes experienced by workers moving between different types of firms. Card et al. (2013) and Card et al. (2018) use matched employer-employee data from Germany and Portugal to perform such event study analyses of the earnings changes experienced by workers moving between different types of firms. In Appendix Figure A.4, we perform the same exercise, but this time for our U.S. data. This analysis uses the movers sample. As in Card et al. (2013) and Card et al. (2018), we define firm groups based on the average pay of coworkers.

The results from the event study mirror those reported in Card et al. (2013) and Card et al. (2018). 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. Additionally, the gains and losses for movers in opposite directions between any two groups of firms are relatively symmetric. By comparison, earnings do not change materially when workers move between firms with similarly paid coworkers. Another relevant finding from the event study is that the earnings profiles of the various groups are all relatively stable in the years before and after a job move. This lends support to Assumption 1.c in LMS, as it suggests that worker mobility does not seem to depend strongly on the trends in earnings beforehand or afterwards. Lastly, it is interesting to observe that the gains and losses for movers seem to be permanent. In contrast, in a large class of search models with job ladders, moves to firms that currently pay less is rationalized by arguing that these firms will pay more in the future.

Although the event study results are consistent with the log additive functional form, we cannot rule out interaction effects between worker and firm effects. Indeed, Bonhomme et al. (2019) point out that even if the functional form is non-additive, the gains and losses may look symmetric if workers making upward moves are of similar quality as those making downward moves. More generally, the degree of asymmetry one observes in the event study depends both on the magnitudes of any interaction effects and on the extent to which workers making upward moves differ in quality from those making downward moves. We explore asymmetry in the movers event study using our estimated model below. Thus, the event study analysis needs to be interpreted with caution.

2.2 Method to estimate firm-worker interactions

To obtain an actual estimate of the importance of interactions between worker and firm effects, we follow BLM in using the following model of earnings:

$$w_{it} = \underbrace{\theta_{j(i,t)} \cdot x_i}_{\text{interaction}} + \psi_{j(i,t)} + v_{it} \tag{A.2}$$

which reduces to AKM when θ_j is the same for all firms. Under Assumptions 1.a-1.c in LMS as well as $\gamma_r = \Upsilon = 0$, we obtain:

$$\begin{aligned}
\mathbb{E}[w_{it+1}|j_2 \to j_1] - \mathbb{E}[w_{it}|j_1 \to j_2] &= \theta_{j_1} \left(\mathbb{E}\left[x_i|j_2 \to j_1\right] - \mathbb{E}\left[x_i|j_1 \to j_2\right]\right) \\
\mathbb{E}[w_{it+1}|j_1 \to j_2] - \mathbb{E}[w_{it}|j_2 \to j_1] &= -\theta_{j_2} \left(\mathbb{E}\left[x_i|j_2 \to j_1\right] - \mathbb{E}\left[x_i|j_1 \to j_2\right]\right)
\end{aligned}$$

where $j_1 \to j_2$ ($j_2 \to j_1$) is an indicator for a worker moving from firm 1 to 2 (firm 2 to 1). As long as the workers moving from 1 to 2 are not exactly the same as those moving from 1 to 2, the right hand side of these equalities are non-zero and we can recover θ_{j1}/θ_{j2} from the moment condition:

$$\frac{\mathbb{E}[w_{it+1}|j_2 \to j_1] - \mathbb{E}[w_{it}|j_1 \to j_2]}{\mathbb{E}[w_{it}|j_2 \to j_1] - \mathbb{E}[w_{it+1}|j_1 \to j_2]} = \frac{\theta_{j1}}{\theta_{j2}} \tag{A.3}$$

Thus, provided that the composition of movers differs across firms, it is possible to identify θ_j (up to scale) for every firm. To take equation (A.2) to the data, however, it is useful to reduce the number of parameters to estimate. As above, we follow BLM in classifying firms to ten types according to the empirical earnings distribution within firms. Then, in a second step, θ_j is estimated, restricting θ_j to be the same for all firms of a given type. Finally, in a third step, the worker effects and firm effects are estimated, restricting ψ_j to be the same for all firms of a given type.

In order to characterize the contribution of firm-worker interactions to earnings inequality, note that we can re-arranging equation (A.2) as,

$$w_{it} = \underbrace{\bar{\theta}(x_i - \bar{x})}_{\tilde{x}_i} + \underbrace{\left(\psi_{j(i,t)} + \theta_{j(i,t)}\bar{x}\right)}_{\tilde{\psi}_{j(i,t)}} + \underbrace{(\theta_{j(i,t)} - \bar{\theta})(x_i - \bar{x})}_{\varrho_{ij(i,t)}} + v_{it} \tag{A.4}$$

where $\bar{\theta} \equiv \mathbb{E}\left[\theta_{j(i,t)}\right]$ and $\bar{x} \equiv \mathbb{E}\left[x_i\right]$. This equation decomposes the earnings of worker i in period t into three distinct components: $\tilde{x}_i$ gives the direct effect of the quality of worker i (evaluated at the average firm), $\tilde{\psi}_{j(i,t)}$ represents the direct effect of firm j (evaluated at the average worker), and $\varrho_{ij(i,t)}$ captures the interaction effect between firm j and worker i quality. Using equation (A.4), we obtain a new variance decomposition of log earnings:

$$\begin{aligned}
Var(w_{it}) =& Var\left[\tilde{x}_i\right] + Var\left[\tilde{\psi}_{j(i,t)}\right] + 2Cov\left[\tilde{x}_i, \tilde{\psi}_{j(i,t)}\right] \\
&+ Var\left[\varrho_{ij(i,t)}\right] + 2Cov\left[\tilde{x}_i + \tilde{\psi}_{j(i,t)}, \varrho_{ij(i,t)}\right] + Var\left[v_{it}\right]
\end{aligned} \tag{A.5}$$

		Model Specification			
		(1)	(2)	(3)	(4)
Share explained by:					
i) Worker Quality	$Var(x_i)$	72.4%	70.4%	73.5%	71.6%
ii) Firm Effects	$Var(\psi_{j(i)})$	3.2%	4.3%	3.0%	4.3%
iii) Sorting	$2Cov(x_i, \psi_{j(i)})$	12.9%	13.1%	12.8%	13.0%
iv) Interactions	$Var(\varrho_{ij})$		3.0%		3.4%
	$+2Cov(x_i + \psi_{j(i)}, \varrho_{ij})$		-1.8%		-2.5%
v) Time-varying Effects	$Var(\psi_{j(i),t} - \psi_{j(i)})$			1.2%	1.2%
	$+2Cov(x_i, \psi_{j(i),t} - \psi_{j(i)})$			-0.9%	-0.9%
Sorting Correlation:	$Cor(x_i, \psi_{j(i)})$	0.43	0.38	0.43	0.37
Variance Explained:	R^2	0.89	0.89	0.90	0.90
Specification:					
Firm-Worker Interactions		✗	✓	✗	✓
Time-varying Firm Effects		✗	✗	✓	✓

Table 2: Comparison of BLM Specifications

Notes: This table presents the decomposition of log earnings variation into firm and worker effects using the BLM estimator for four specifications: baseline, allowing for worker effects to interact with firm effects ("Firm-Worker Interactions"), allowing for a time-varying component in the firm effects due to the pass through of value added shocks ("Time-varying Firm Effects"), and allowing for both interactions between firm and worker effects and time-varying firm effects. The analysis uses both movers and non-movers.

The first three components are informative about the inequality contribution from worker effects, firm effects and worker sorting, net of interaction effects. The next two components are informative about the inequality contribution from interaction effects, as measured by the dispersion of $\varrho_{ij(i,t)}$ across firms and the extent to $\varrho_{ij(i,t)}$ is larger in firms with high wages. If $\theta_j = \bar{\theta}$ for every firm j, then these two components would be zero, and the decomposition in (A.5) reduces to the standard AKM decomposition.

2.3 Estimates from the model with firm-worker interactions

In Table 3 under "No Time-varying Effects", we provide the estimates of (ψ_k, θ_k) from the model with firm-worker interactions in equation (A.2). As discussed above, we follow BLM in classifying firms to ten types according to the empirical earnings distribution within firms. The firm groups $k = 1, 2, ..., 10$ are in ascending order by mean earnings.

The numerical values of ψ_k range from 0.0 for group $k = 1$ (this is the normalization), to about 0.5 for group $k = 6$, and a maximum of about 1.0 for group $k = 10$. The numerical values of θ_k range from 1.0 for group $k = 1$ (this is the normalization), to about 1.4 for group $k = 6$, and a maximum of about 2.2 for group $k = 10$. We see that θ_j is nearly monotonic in ψ_k, i.e., firms with greater fixed effects also tend to have greater production complementarities. In order to better understand how (ψ_k, θ_k) impact earnings, Table 3 also provides the predicted values of log earnings, $\mathbb{E}[w|x, k] = \psi_k + \theta_k \cdot x_q$, for each group k and considering various quantiles x_q in the distribution of x. For example, these predicted values indicate that moving from the lowest

	No Time-varying Effects							With Time-varying Effects						
					$\mathbb{E}[w\|x,k]$ by x quantile							$\mathbb{E}[w\|x,k]$ by x quantile		
Cluster k	N_k	$\mathbb{E}\log W$	ψ_k	θ_k	20th	50th	80th	N_k	$\mathbb{E}\log W$	ψ_k	θ_k	20th	50th	80th
1	11.16	10.10	0.00	1.00	-0.68	-0.43	-0.18	11.06	10.10	0.00	1.00	-0.67	-0.43	-0.18
2	24.88	10.28	0.17	1.12	-0.59	-0.31	-0.03	24.71	10.28	0.17	1.13	-0.58	-0.31	-0.03
3	32.82	10.44	0.38	1.39	-0.57	-0.22	0.13	32.70	10.43	0.39	1.41	-0.56	-0.21	0.14
4	32.59	10.56	0.45	1.42	-0.52	-0.16	0.19	32.67	10.56	0.47	1.45	-0.50	-0.15	0.21
5	23.19	10.70	0.51	1.43	-0.46	-0.10	0.25	23.41	10.70	0.54	1.47	-0.45	-0.09	0.27
6	21.21	10.75	0.65	1.81	-0.59	-0.14	0.32	21.05	10.75	0.68	1.88	-0.58	-0.12	0.34
7	28.39	10.87	0.67	1.69	-0.48	-0.06	0.36	28.45	10.87	0.69	1.74	-0.47	-0.05	0.38
8	24.26	11.07	0.77	1.89	-0.51	-0.04	0.43	24.38	11.07	0.80	1.94	-0.51	-0.03	0.44
9	18.28	11.30	0.89	2.02	-0.49	0.01	0.52	18.49	11.30	0.91	2.08	-0.48	0.03	0.54
10	9.02	11.63	1.01	2.24	-0.52	0.04	0.60	8.88	11.64	1.03	2.31	-0.52	0.05	0.62

Table 3: Parameter Estimates with Firm-Worker Interactions

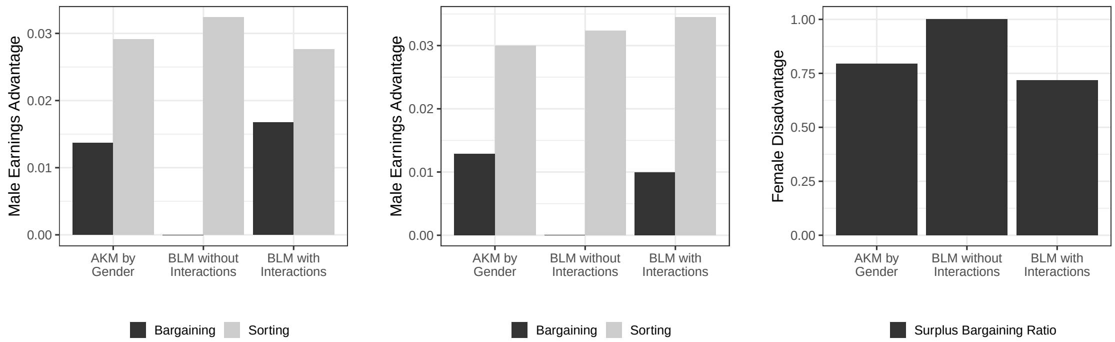
Notes: In this table, we describe the estimated parameters and wage predictions from the BLM specification with firm-worker interactions. $\log W$ refers to the mean of (non-residual) observed log earnings within a group k. The model prediction for (residual) log earnings is given by $\mathbb{E}\left[w|x = x_q, k = k\right] = \psi_k + \theta_k \cdot x_q$, for each group $k = 1, 2, ..., 10$ and considering various quantiles x_q in the distribution of x. N_k refers to the number of worker-years (in millions) observed in the cluster during the 2001-2008 time interval.

to the highest type of firm increases earnings by 17, 47 and 80 percentage points for individuals at the 20, 50 and 80 percentile in the worker quality distribution.

The evidence of nonlinearities raises several questions. To what extent do interaction effects bias the estimates from the log additive model? Are nonlinearities empirically important as a source of earnings inequality? In Table 2, we investigate these questions by extending the AKM decomposition to incorporate the contribution from interactions between worker and firm effects. The results from the decomposition in equation (A.5) are presented in column (2) of Table 2. Our estimates suggest the dispersion of interaction effects across firms explains three percent of the earnings inequality. However, the total contribution to earnings inequality from nonlinearities is muted by the interaction effects being larger in firms with higher paid workers. We also find that omitting interaction effects causes a downward bias in the firm effects and an upward bias in the worker effects.

2.4 Comparison to AKM interacted with gender

Our specification with firm-worker interactions allows the same firm to offer different firm effects for different unobserved worker ability types. Relatedly, a recent literature has extended the AKM model to allow for different observable types of workers to have different firm effects within the same firm. In particular, Card et al. (2016) estimate AKM firm effects separately for men and women in Portugal, restricting the firm effects for both men and women to be zero on average among a set of “low surplus” firms. This allows a given firm to have distinct firm effects for men and women, while also anchoring the level of the firm effects so that men and women can be directly compared. Our model of interactions will account for different firm effects between men and women within firms to the extent that men and women have different worker effects within firms. Thus, to the extent that differences in unobserved worker ability capture differences between observable worker types, our specification may be consistent with



(a) Decomposition (A.6) (b) Decomposition (A.7) (c) π^F/π^M

Figure 2: Gender Bargaining and Sorting in the US

Notes: In this figure, we present estimates from equations (A.6)-(A.7), as well as the bargaining ratio π^F/π^M. We compare results based on actual earnings observed in the data to results based on predicted earnings when using only parameters $\left(x_i, \psi_{j(i)}, \theta_{j(i)}\right)$ from the model with firm-worker interactions.

AKM specifications that allow for different firm effects by gender.

In order to investigate this possibility, we follow Card et al. (2016) in separately estimating AKM for females and males, except we apply it to the US data. We use a similar strategy to impose the mean zero restriction at low surplus firms. In particular, we classify low surplus firms as those below the lower kink point in a nonparametric regression of log earnings per worker on log value added per worker; Appendix Figure A.5 shows where we chose the kink point. Denoting the firm effects for females and males by $\psi^F_{j(i)}, \psi^M_{j(i)}$, respectively, we perform the following decompositions proposed by Card et al. (2016), where we use their terminology to label components:

$$\mathbb{E}\left[\psi^M_{j(i)}|M\right] - \mathbb{E}\left[\psi^M_{j(i)}|F\right] = \underbrace{\mathbb{E}\left[\psi^M_{j(i)} - \psi^F_{j(i)}|M\right]}_{\text{Bargaining (male)}} + \underbrace{\mathbb{E}\left[\psi^F_{j(i)}|M\right] - \mathbb{E}\left[\psi^F_{j(i)}|F\right]}_{\text{Sorting (female)}} \tag{A.6}$$

$$= \underbrace{\mathbb{E}\left[\psi^M_{j(i)} - \psi^F_{j(i)}|F\right]}_{\text{Bargaining (female)}} + \underbrace{\mathbb{E}\left[\psi^M_{j(i)}|M\right] - \mathbb{E}\left[\psi^M_{j(i)}|F\right]}_{\text{Sorting (male)}} \tag{A.7}$$

where conditioning on F or M means restricting to female or male workers, respectively. We also follow Card et al. (2016) in regressing $\psi^F_{j(i)}$ and $\psi^M_{j(i)}$ on net surplus. Net surplus is defined as log value added per worker after low surplus firms are set to have zero log value added per worker. Denote these regression coefficients by π^F and π^M. The ratio π^F_j/π^M_j is, according to Card et al. (2016), a measure of female bargaining power. It is less than one when females have less bargaining power than males.

Figure 2 displays the results from these decompositions. We compare the estimates from equations (A.6)-(A.7), as well as the so called bargaining ratio, π^F/π^M. Our first conclusion is that the bargaining and sorting estimates for the US are relatively similar to those estimated by

Card et al. (2016) for Portugal.[4] Next, we compare the estimates by gender to those implied by our estimated model using the BLM approach, both with and without allowing for interaction effects with unobserved worker ability. In the absence of interaction effects, men and women have the same premium within a firm when using the BLM approach, so the bargaining component is zero and the bargaining ratio π_j^F/π_j^M is one. Our model of interactions will account for different firm effects between men and women within firms to the extent that men and women have different worker effects within firms. Importantly, our estimator with interaction effects does not use any information on whether an individual is male or female. When allowing for interactions, we are able to approximately recover the bargaining effect as well as the bargaining ratio from the BLM estimator. Thus, our second conclusion is that our empirical specification of interaction effects on worker ability appears to well represent differences in firm premiums offered to men and women within firms.

2.5 Comparing Hourly Wages and Annual Earnings

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 estimation, one must then take a stand on the proper measure of wages or earnings. One may be concerned that the estimates of firm-worker interactions are particularly sensitive to the measure chosen. Norwegian administrative data is an exception, as we have accurate measures of hours worked in this data set and can thus construct administrative measures of hourly wages (see Appendix A for a description of this data). Appendix Figure A.2 compares results from the model with firm-worker interactions when using annual earnings (subfigure a) and hourly wages (subfigure b) in Norway. Using hourly wages does not alter the conclusion that there are substantial complementarities between firms and workers such that high quality workers earn disproportionately greater earnings or wages at high wage firms.

2.6 Evidence from Discrete Worker Types

In our preferred specification, worker ability is allowed to have a continuous support while firms are assumed to have a discrete support. While the number of worker types is unrestricted, the interactions between worker ability and firm types are restricted to have a linear functional form (see equation A.2). We now consider an alternative estimator proposed by BLM in which worker types are assumed to also be discrete while the interaction effects between worker types and firm types are unrestricted. We maintain the same 10 firm types from the main results, while restricting workers to have 5 types.

Appendix Figure A.6 provides the predicted log earnings from this estimator with discrete worker types. We find that the main specification with interactions provides a good approximation to the estimation with discrete worker types, which supports our choice of main specification.

[4] For example, in decompositions (A.6)-(A.7), we find that bargaining and sorting components are in the range 0.01-0.02 and 0.03-0.04, respectively, while their estimates are in the range 0.00-0.02 and 0.03-0.05, respectively.

In particular, we find evidence of complementarities between high-type workers and high-type firms for both estimation methods.

3 Pass through of shocks and time-varying types

The assumption that $\gamma_r = \varUpsilon = 0$ restricts firm effects to be constant over time. However, the significant pass through rates estimated by LMS imply that firm effects actually evolve over time as employers experience changes in the value added at the firm or market level. To capture this, we now let $\gamma_r \neq 0$, $\varUpsilon \neq 0$, and for γ_r to vary across r, and propose an adjustment to the AKM model which allows us to isolate the time-invariant component of the firm effects. The estimates of (γ_r, Υ) are provided in Section 5.2 of LMS.

Our approach proceeds in two steps. First, we construct an adjusted earnings measure by removing the time-varying firm and market specific component of earnings. To do so, we use the firm and market level value added multiplied by the estimated passthrough coefficients at the firm and market level. Second, we recover the time-invariant firm and worker effects by applying the methods of AKM or BLM to the adjusted measure of earnings. Consider the following adjusted two-way specification for earnings of workers across firms:

$$\mathbb{E}[w_{it} - \Upsilon \bar{y}_{r(i,t),t} - \gamma_r(y_{j(i,t),t} - \bar{y}_{r(i,t),t})|j(i,t) = j \in J_r] = x_i + \psi_j.$$

The left-hand side removes the earnings dynamics due to passthrough of firm-specific shocks, $\gamma_r(y_{j(i,t),t} - \bar{y}_{r(i,t),t})$, and market shocks, $\varUpsilon \bar{y}_{r(i,t),t}$. What remains is the worker effect x_i and the time-invariant firm effect ψ_j, which can be estimated by applying AKM or BLM to the adjusted earnings measure.

In column (3) of Table 2, we extend the BLM decomposition of the variance of log earnings in equation (A.1) to incorporate the contribution from time-invariant and time-varying firm effects. We find that time-varying firm effects explain little if any of the variation in log earnings, and that the importance of firm effects and worker sorting do not change materially if we take the pass through of firm shocks into account. Comparing the results in column (4), which is the main specification from equation (14) of LMS, to those presented in column (2) shows that time variation also has little to no explanatory power when accounting for nonlinearities. Furthermore, in Table 3, we see that time-invariant firm effects and interaction parameters are very similar regardless of whether or not we account for time-varying effects.

References

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

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.

Barth, E., A. Bryson, J. C. Davis, and R. Freeman (2016): "It's where you work: Increases in earnings dispersion across establishments and individuals in the US," *Journal of Labor Economics*, 34, S67–S97.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020): "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?" Working Paper w27368, National Bureau of Economic Research.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87, 699–739.

Borovickova, K. and R. Shimer (2017): "High Wage Workers Work for High Wage Firms," Working Paper 24074, National Bureau of Economic Research.

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.

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

Kline, P., R. Saggio, and M. Sølvsten (2020): "Leave-out estimation of variance components," *Econometrica*, 88, 1859–1898.

Lamadon, T., M. Mogstad, and B. Setzler (2020): "Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market," Working Paper w25954, National Bureau of Economic Research.

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 (2018): "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.

A Description of the Norwegian Data

The Norwegian 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 annual earnings at the primary employer as our main outcome of interest. Because the Norwegian data also provides hours worked per day, we can construct the average hourly wage. We supplement the employer-employee data with a measure of value added, which we define as the difference in sales and non-wage operating costs as reported to the Norwegian tax authority by the firm.

To harmonize the Norwegian data with our sample from the US, we follow Bonhomme et al. (2020) by applying 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. Recall that, since we do not observe hours worked in US data, or a formal measure of full-time employment, we defined 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 Norwegian sample to workers with annual earnings above 32.5% of the national average.

B Additional Tables and Figures

Sample:	Full Sample	≥ 2 Movers	Connected Set
Workers in 2001-2008:			
Worker-Years (Millions)	245.0	227.8	227.4
	(100.0%)	(93.0%)	(92.8%)
Unique Workers (Millions)	66.2	61.8	61.7
	(100.0%)	(93.3%)	(93.2%)
Workers in 2008-2015:			
Worker-Years (Millions)	232.9	212.4	211.9
	(100.0%)	(91.2%)	(91.0%)
Unique Workers (Millions)	64.0	58.8	58.6
	(100.0%)	(91.9%)	(91.7%)

Table A.1: Floor on Number of Movers and the Connected Set

Notes: This table demonstrates the fraction of workers kept in the sample in the AKM and BLM analysis when imposing that a firm must have at least two movers and must belong to the connected set of firms.

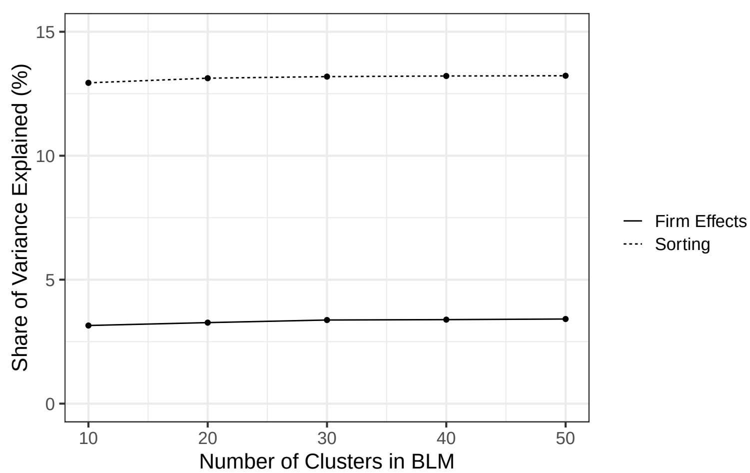
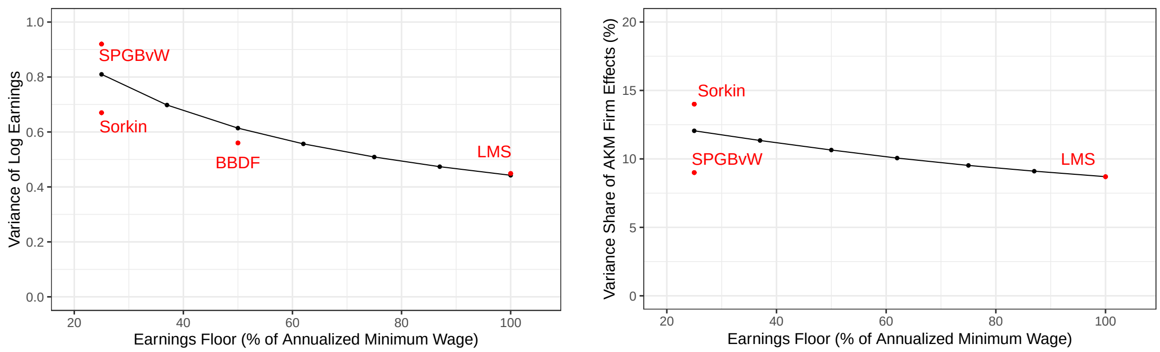


Figure A.1: BLM Decomposition by Number of Clusters

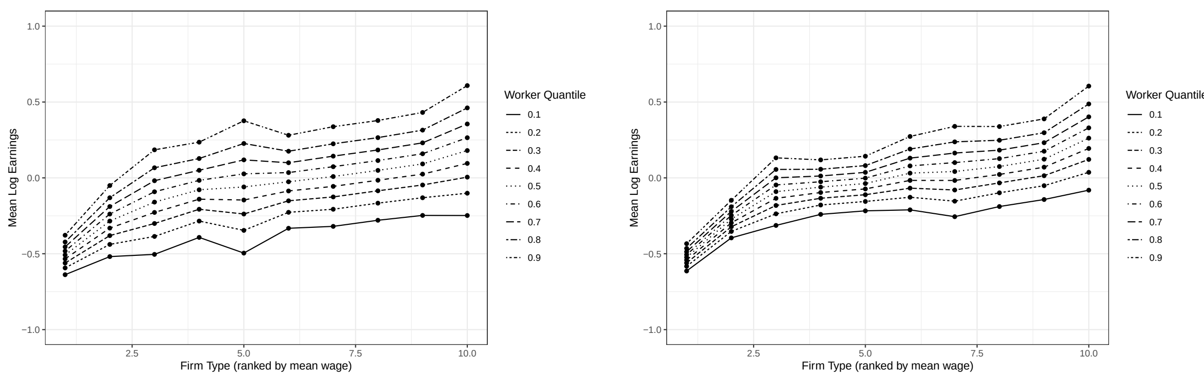
Notes: In this figure, we estimate the BLM decomposition for different numbers of firm clusters.



(a) Total Variance of Log Earnings (b) AKM Estimates of Firm Component

Figure A.2: Earnings Variance and AKM Estimates of Firm Component by Earnings Floor

Notes: In this figure, we report estimates of the variance of log earnings (subfigure a) and AKM estimates of the firm component (subfigure b) when imposing different FTE wage floors. Literature abbreviations are BBDF for Barth et al. (2016), SPGBvW for Song et al. (2018), and Sorkin for Sorkin (2018).



(a) Annual Earnings (b) Hourly Wage

Table A.2: Predicted Earnings or Wages with Firm-Worker Interactions in Norway

Notes: In this table, we describe the estimated parameters and wage predictions from the BLM specification with firm-worker interactions. The prediction is given by $\mathbb{E}\left[w|x,k\right] = \psi_k + \theta_k \cdot x$.

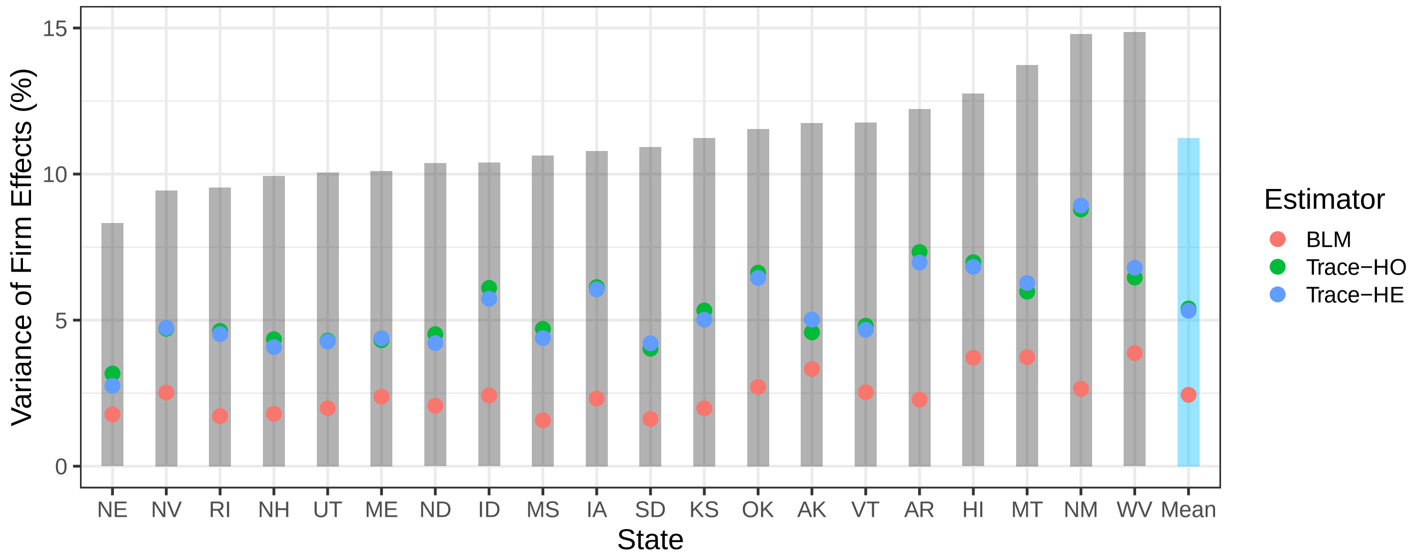


Figure A.3: Comparison of Estimators for a Subset of the Smaller States in the U.S.

Notes: This figure considers the 2001-2008 sample of workers for a number of smaller U.S. states. It compares four estimators: Abowd et al. (1999) (AKM, in the shaded bars), the Andrews et al. (2008) estimator (Trace-HO), the Kline et al. (2020) estimator (Trace-HE), and the Bonhomme et al. (2019) estimator (BLM). All estimation is performed on the leave-one-out connected set. We also report the average of the estimated variances of firm effects across the states.

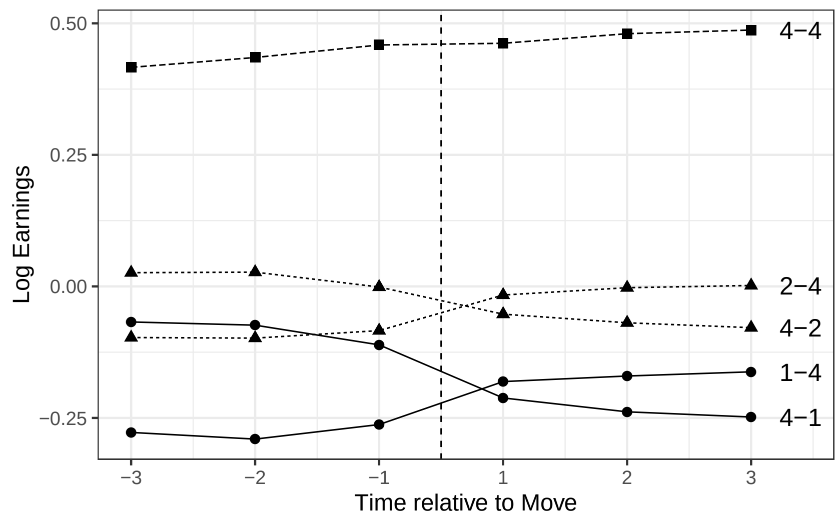


Figure A.4: Event Study of Changes in Earnings when Workers Move Between Firms

Notes: In this figure, we classify firms into four equally sized groups based on the mean earnings of non-movers in the firm (with 1 and 4 being the group with the lowest and highest mean earnings, respectively). We then compute mean log earnings for the workers that move between these groups of firms in the years before and after the move. Note that the employer differs between event times -1 and 1, but we do not know exactly when the change in employer occurred. Thus, to avoid concerns over workers exiting and entering employment during these years, one might prefer to compare earnings in event years -2 and 2.

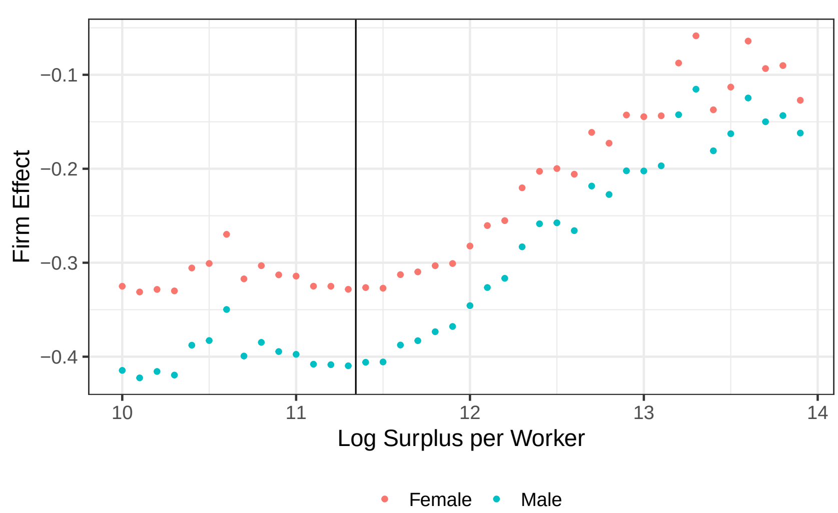


Figure A.5: Choosing the restriction for gender-specific AKM firm effects

Notes: This figure plots AKM firm effects for females and males against log value added per worker. For the purposes of making this plot, firm effects are normalized to zero for both men and women at the largest firm in the sample. The solid vertical line indicates our chosen "kink point" at which both lines become upward sloping.

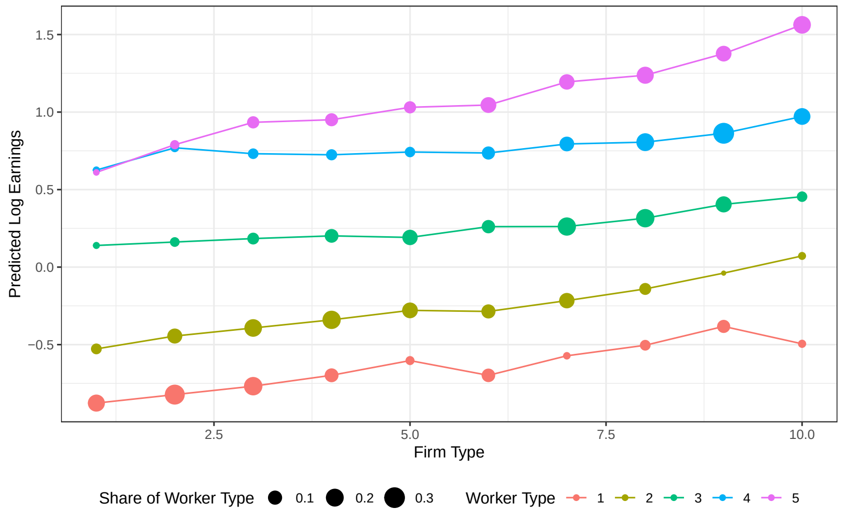


Figure A.6: BLM Estimator with Discrete Worker Types

Notes: This figure plots predicted log earnings when using the BLM estimator with 5 discrete worker types. "Share of Worker Type" refers to the distribution of a given worker type across firm types, that is, it sums to 1.0 along a Worker Type line.