Econometrica Supplementary Material

SUPPLEMENT TO "A DISTRIBUTIONAL FRAMEWORK FOR MATCHED EMPLOYER EMPLOYEE DATA"

(*Econometrica*, Vol. 87, No. 3, May 2019, 699–739)

STÉPHANE BONHOMME
Department of Economics, University of Chicago

THIBAUT LAMADON
Department of Economics, University of Chicago

ELENA MANRESA
Department of Economics, New York University

This appendix contains details on estimation and computation in Section S1, an exercise on data simulated from a theoretical sorting model in Section S2, and various extensions in Section S3. Tables and figures may be found at the end of the document.

S1. ESTIMATION AND COMPUTATION

IN THIS SECTION, we provide details on our computation procedure, and on the estimation of persistence parameters in the dynamic model.

S1.1. *Exploration of the Likelihood Function*

Given the presence of local optima in our finite mixture model, the choice of initial conditions and exploration strategy is important. Here, we describe how we explore the likelihood function to obtain our baseline estimates of log-earnings in the static model, based on job movers. Our estimates—and estimates within the bootstrap—are based on 50 starting values.

To obtain a starting value, we first draw L wages from a Gaussian distribution with mean equal the mean of log-earnings in period 1 and standard deviation equal twice the standard deviation of log-earnings in the same period. Using the EM algorithm, holding mean log-earnings fixed across firms, we then compute estimates of proportions of worker types and type and class-specific log-earnings variances. We use these estimates as starting values for another preliminary estimation where mean log-earnings are held constant across firms, and estimated jointly with log-earnings variances and type proportions. We then use the estimates obtained with this second estimation as initial conditions for our estimation based on job movers. The resulting parameter estimates are then used in the final estimation step based on job stayers.

Graph Connectedness

Our identification results emphasize the importance of connectedness, through the presence of connecting cycles for each worker type. The mobility patterns of workers define a graph across firm classes. A measure of connectedness of a graph is the small-

Stéphane Bonhomme: sbonhomme@uchicago.edu
Thibaut Lamadon: lamadon@uchicago.edu
Elena Manresa: em1849@nyu.edu

 https://doi.org/10.3982/ECTA15722

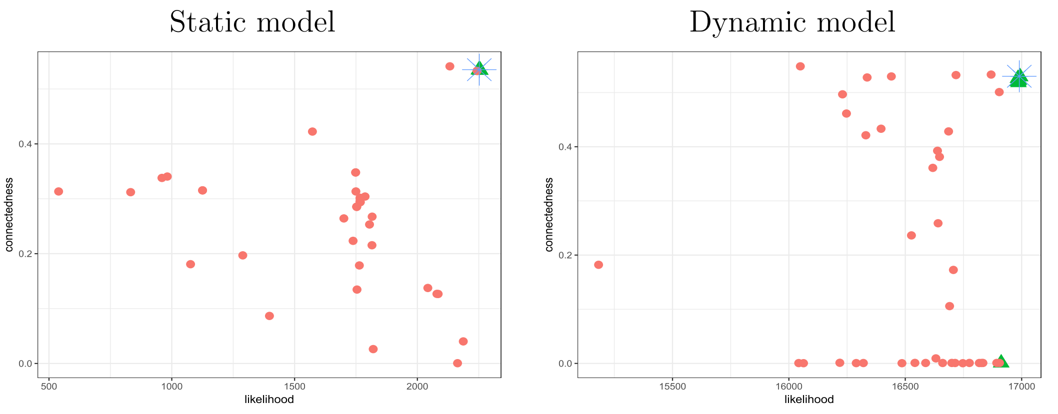

FIGURE S1.—Likelihood and connectedness of local optima. Notes: The dots show the likelihood values (x-axis) and connectedness measures (y-axis) corresponding to all local optima of the part of the likelihood function corresponding to job movers, obtained when starting the algorithm at each of the 50 starting values. The triangles show the ten best likelihood values. The stars show our selected values.

est nonzero eigenvalue of its normalized Laplacian, as recently studied in Jochmans and Weidner (2017).[1] We observed that the local optima of the likelihood function tended to vary substantially in terms of their connectedness, some of the solutions having types with very low connectedness. To discriminate between estimates that have similar likelihood values, we favor estimates with higher connectedness. Our main estimates are based on the best connected solution out of the ones yielding the 10 highest likelihood values. This strategy mainly improves stability across bootstrap repetitions and has little impact on the main estimates.

In the left panel of Figure S1, we plot the likelihood value against the connectedness measure for the static model. In this case, the solutions yielding the highest likelihood values (depicted as triangles in the figure) coincide with the one showing highest connectedness (the star). In the right panel, we show the same relationship for the dynamic model. In this case, there is more uncertainty about the exact location of the highest likelihood value. We see that our solution (the star) not only has high likelihood but also high connectedness. In both cases, the solutions using the maximum likelihood estimates are very similar to the ones we report.

S1.2. *Estimation of $\rho_{4|3}$ and $\rho_{1|2}$ in the Dynamic Model*

Consider the dynamic model under the specification described in Section 4.2. Note that the unconditional means of log-earnings of job stayers of type α in class k are: $\mu_{1k\alpha} + \rho_{1|2}\mu^s_{2k\alpha}$, $\mu^s_{2k\alpha}$, $\mu^s_{3k\alpha}$, and $\mu_{4k\alpha} + \rho_{4|3}\mu^s_{3k\alpha}$, respectively. We make the following assumption, for all worker types α, α' and all firm classes k:

$$\begin{aligned}\mu^s_{2k\alpha'} - \mu^s_{2k\alpha} &= \mu_{1k\alpha'} + \rho_{1|2}\mu^s_{2k\alpha'} - \left(\mu_{1k\alpha} + \rho_{1|2}\mu^s_{2k\alpha}\right) \\ &= \mu^s_{3k\alpha'} - \mu^s_{3k\alpha} \\ &= \mu_{4k\alpha'} + \rho_{4|3}\mu^s_{3k\alpha'} - \left(\mu_{4k\alpha} + \rho_{4|3}\mu^s_{3k\alpha}\right). \qquad \text{(S1)}\end{aligned}$$

[1]Empirically, we measure connectedness as the minimum, across all worker types, of the smallest nonzero eigenvalues of the normalized Laplacians of the type-specific graphs (weighted by number of movers), where the graphs are at the firm-class level.

While (S1) imposes that the effect of worker heterogeneity on mean log-earnings is constant over time *within firm*, it allows for unrestricted interactions between firm classes and time.

When (S1) holds, the persistence parameters $\rho_{4|3}$ and $\rho_{1|2}$ can be estimated using simple covariance restrictions, as we now explain. The four periods' log-earnings of a job stayer of type α in class k can be written as

$$\begin{aligned}
Y_{i1} &= c_{1k} + (1-\rho_{1|2})\mu^s_{2k\alpha} + \rho_{1|2}Y_{i2} + \nu_{i1},\\
Y_{i2} &= c_{2k} + \mu^s_{2k\alpha} + \nu_{i2},\\
Y_{i3} &= c_{3k} + \mu^s_{2k\alpha} + \nu_{i3},\\
Y_{i4} &= c_{4k} + (1-\rho_{4|3})\mu^s_{2k\alpha} + \rho_{4|3}Y_{i3} + \nu_{i4},
\end{aligned}$$

where ν_{i1} is independent of $(\nu_{i2}, \nu_{i3}, \nu_{i4})$, ν_{i4} is independent of $(\nu_{i3}, \nu_{i2}, \nu_{i1})$, and (taking as reference type $\alpha' = 1$): $c_{1k} = \mu_{1k1} + \rho_{1|2}\mu^s_{2k1} - \mu^s_{2k1}$, $c_{2k} = 0$, $c_{3k} = \mu^s_{3k1} - \mu^s_{2k1}$, and $c_{4k} = \mu_{4k1} + \rho_{4|3}\mu^s_{3k1} - \mu^s_{2k1}$.

The within-firm covariances between Y_{i1} and $Y_{i2} - Y_{i3}$ and between Y_{i4} and $Y_{i3} - Y_{i2}$ then deliver consistent estimators under standard rank conditions. As an example, the model implies the panel-IV restriction $\operatorname{Cov}(Y_{i4}, Y_{i3} - Y_{i2}|k) = \rho_{4|3}\operatorname{Cov}(Y_{i3}, Y_{i3} - Y_{i2}|k)$. Notice that here $\mu^s_{2k\alpha}$ plays the role of a "fixed effect" within firm class k. In practice, we combine those restrictions with all other covariance restrictions, hence also estimating the within-firm variances of the ν's and covariances of (ν_{i2}, ν_{i3}) (in particular, the parameter $\rho^s_{3|2}$). We estimate the parameters by minimum-distance with equal weights within firm classes, weighting each firm class according to the number of firms in the class.

S1.3. *Model Simulation*

The simulation for the parametric bootstrap is conditional on firm classes and the mobility links between firms and workers, including the size of firms. We describe the simulation algorithm for the static model.

1. (Job stayers) For each firm in a class, we draw independently the latent types of job stayers in the firm according to the distribution of types in the class. Log-earnings are then independently drawn across workers from the corresponding conditional distribution.
2. (Job movers) For each pair of firms in periods 1 and 2 in given classes, we draw the latent types of job movers between those firms according to their distribution in the pair of classes. We draw log-earnings in periods 1 and 2 according to their conditional distribution. In the static model, we draw log-earnings independently across periods.

S2. ESTIMATION USING DATA FROM A THEORETICAL MODEL

In this section, we consider a variation of the model of Shimer and Smith (2000) with on-the-job search. Relative to the main text, we modify some of the notation.

S2.1. *Model*

Environment

The economy is composed of a uniform measure of workers indexed by x with unit mass and a uniform measure of jobs indexed by y with mass $\bar{V}$. A match (x, y) produces output $f(x, y)$ and separates exogenously at rate δ. Workers are employed or unemployed. We denote $u(x)$ the measure of unemployed, $h(x, y)$ the measure of matches, and $v(y)$ the measure of vacancies. We let $U = \int u(x)\,\mathrm{d}x$ the mass of unemployed, and $V = \int v(y)\,\mathrm{d}y$ the mass of vacancies. Unemployed workers meet vacancies at rate λ_0, and employed workers meet vacancies at rate λ_1. Vacancies meet unemployed workers at rate μ_0, and employed workers at rate μ_1. A firm cannot advertise for a job that is currently filled. Unemployed workers collect benefits $b(x)$, and vacancies have to pay a flow cost $c(y)$.

Timing

Each period is divided into four stages. In stage one, active matches collect output and pay wages. In stage two, active matches exogenously separate with probability δ. In stage three, vacant jobs can advertise and all workers can search. In stage four, workers and vacant jobs meet randomly and, upon meeting, the worker and the firm must decide whether or not to match based on expected surplus generated by the match. The wage is set by Nash bargaining, where α is the bargaining power of the worker. We assume that wages are continuously renegotiated with the value of unemployment; see Shimer (2006) for a discussion. Since workers and firms can search in the same period as job losses occur, it is convenient to introduce within-period distributions:

$$v_{1/2}(y) := \frac{\delta + (1-\delta)v(y)}{\delta + (1-\delta)V}, \qquad u_{1/2}(x) := \frac{\delta + (1-\delta)u(x)}{\delta + (1-\delta)U}, \qquad h_{1/2}(x, y) := \frac{h(x, y)}{1-U},$$

where each distribution integrates to 1 by construction.

Value Functions

We then write down the value functions for this model. Let $S(x, y)$ be the surplus of the match, $W_0(x)$ the value of unemployment, and $\Pi_0(y)$ the value of a vacancy. We have

$$rW_0(x) = (1+r)b(x) + \lambda_0 \int M(x, y)\alpha S(x, y)v_{1/2}(y)\,\mathrm{d}y, \tag{BE-W0}$$

and

$$\begin{aligned} r\Pi_0(y) = {}& \mu_0 \int M(x, y)(1-\alpha)S(x, y)u_{1/2}(x)\,\mathrm{d}x \\ & + \mu_1 \iint P\big(x, y', y\big)(1-\alpha)S(x, y)h_{1/2}\big(x, y'\big)\,\mathrm{d}y'\,\mathrm{d}x, \end{aligned} \tag{BE-P0}$$

where $M(x, y) := \mathbf{1}\{S(x, y) \geq 0\}$ is the matching decision, and $P(x, y', y)$ is 1 when $S(x, y) > S(x, y')$ (i.e., when y is preferred to y' by x), zero when $S(x, y) < S(x, y')$, and 1/2 when $S(x, y) = S(x, y')$.

We write the Bellman equation for a job y that currently employs a worker x at wage w:

$$\begin{aligned} (r+\delta)\Pi_1(x, y, w) = {}& (1+r)\big[f(x, y) - w + \delta\big(\Pi_0(y) + c(y)\big)\big] \\ & - (1-\delta)\lambda_1 q(x, y)(1-\alpha)S(x, y), \end{aligned}$$

where $q(x, y) = \int P(x, y, y')v_{1/2}(y')\,\mathrm{d}y'$ represents the total proportion of firms y' that can poach a worker x from firm y. We then turn to the Bellman equation for the employed worker:

$$
\begin{aligned}
&(r+\delta)W_1(x,y,w) \\
&\quad = (1+r)\big[w+\delta\big(W_0(x)-b(x)\big)\big] \\
&\qquad + (1-\delta)\lambda_1 \int P\big(x,y,y'\big)\big(\alpha S\big(x,y'\big)-\alpha S(x,y)\big)v_{1/2}\big(y'\big)\,\mathrm{d}y'.
\end{aligned} \tag{BE-W1}
$$

Finally, we derive the value of the surplus associated with the match (x, y), defined by $S := W_1 + \Pi_1 - W_0 - \Pi_0$:

$$
\begin{aligned}
(r+\delta)S(x,y) &= (1+r)\big[f(x,y)-\delta\big(b(x)-c(y)\big)\big] - r(1-\delta)\big(\Pi_0(y)+W_0(x)\big) \\
&\quad + (1-\delta)\lambda_1 \int P\big(x,y,y'\big)\big(\alpha S\big(x,y'\big)-S(x,y)\big)v_{1/2}\big(y'\big)\,\mathrm{d}y'.
\end{aligned} \tag{BE-S}
$$

Flow Equations

Last, we write the flow equation for the joint distribution of matches at the beginning of the period:

$$
\begin{aligned}
&\big(\delta+(1-\delta)\lambda_1 q(x,y)\big)h(x,y) \\
&\quad = \lambda_0\big(\delta+(1-\delta)U\big)u_{1/2}(x)v_{1/2}(y)M(x,y) \\
&\qquad + \lambda_1(1-\delta)(1-U)\int P\big(x,y',y\big)h_{1/2}\big(x,y'\big)\,\mathrm{d}y' v_{1/2}(y),
\end{aligned} \tag{EQ-H}
$$

where

$$
\begin{aligned}
\mu_0\big(\delta+(1-\delta)V\big) &= \lambda_0\big(\delta+(1-\delta)U\big), \quad \text{and} \\
\mu_1\big(\delta+(1-\delta)V\big) &= \lambda_1(1-\delta)(1-U)
\end{aligned} \tag{MC-S}
$$

are the total number of matches coming out of unemployment and coming from on-the-job transitions, respectively. The market clearing conditions on the labor market are given by

$$
\int h(x,y)\,\mathrm{d}x + v(y) = \bar{V} \quad \text{and} \quad \int h(x,y)\,\mathrm{d}y + u(x) = 1. \tag{MC-L}
$$

Equilibrium

For a set of primitives δ, λ_0, λ_1, $f(x, y)$, $b(x)$, $c(y)$, α, the stationary equilibrium is characterized by the values $S(x, y)$, $W_0(x)$, $\Pi_0(y)$ and the measure of matches $h(x, y)$ such that: (i) Bellman equations (BE-W0), (BE-P0), and (BE-S) are satisfied, (ii) h satisfies the flow equation (EQ-H), and (iii) the constraints (MC-S) and (MC-L) hold.

Wages

We then derive the wage function using equation (BE-W1), and using that Nash bargaining gives $W_1(x,y,w(x,y)) = \alpha S(x,y) + W_0(x)$:

$$\begin{aligned}(1+r)w(x,y) &\\ &= (r+\delta)\alpha S(x,y) + (1-\delta)rW_0(x)\\ &\quad - (1-\delta)\lambda_1 \int P(x,y,y')\big(\alpha S(x,y') - \alpha S(x,y)\big)v_{1/2}(y')\,\mathrm{d}y'.\end{aligned}$$

Mapping to Our Framework

From there, we can recover our static model's cross-sectional worker type proportions conditional on firm heterogeneity ($q_k(\alpha)$ in the body of the paper):

$$q_y(x) = \frac{h(x,y)}{1-v(y)},$$

and the type proportions for job movers ($p_{k'k}(\alpha)$ in the main text), which are given by

$$p_{yy'}(x) = \frac{\big(\delta\lambda_0 + (1-\delta)\lambda_1 \mathbf{1}\{S(x,y') > S(x,y)\}\big)h(x,y)M(x,y')}{\int \big(\delta\lambda_0 + (1-\delta)\lambda_1 \mathbf{1}\{S(\tilde{x},y') > S(\tilde{x},y)\}\big)h(\tilde{x},y)M(\tilde{x},y')\,\mathrm{d}\tilde{x}}.$$

Last, we assume that the wage is measured with a multiplicative independent measurement error:

$$\tilde{w} = w(x,y)\exp(\varepsilon),$$

from which we can derive the log-wage distributions ($F_{k\alpha}$ in the main text).

Without On-the-Job Search ($\lambda_1 = \mu_1 = 0$)

Let us first consider the case without on-the-job search. Equation (EQ-H) gives

$$\delta h(x,y) = \lambda_0\big(\delta + (1-\delta)U\big)u_{1/2}(x)v_{1/2}(y)M(x,y).$$

Hence

$$p_{yy'}(x) = \frac{M(x,y)M(x,y')u_{1/2}(x)}{\int M(\tilde{x},y)M(\tilde{x},y')u_{1/2}(\tilde{x})\,\mathrm{d}\tilde{x}}. \tag{PX-YY$'$}$$

These probabilities are symmetric in (y, y'). In the context of Theorem 1, this means that Assumption 3(i) is not satisfied, as $a(\alpha) = 1$ for all α. This is the setup considered in Shimer and Smith (2000). Symmetry occurs because, in that case, all job changes are associated with an intermediate unemployment spell, where all information about the previous firm disappears. Empirically, the majority of job changes occur via job-to-job transitions. Moreover, in Figure S6, we find evidence against the particular symmetry of equation (PX-YY$'$).

In the left graph of Figure S2, we show the wage functions under a particular parameterization of the model of Shimer and Smith (2000), without on-the-job search. We use the parameterization under positive assortative matching that we describe below, with the only difference that we shut down the on-the-job search channel. In the right graph,

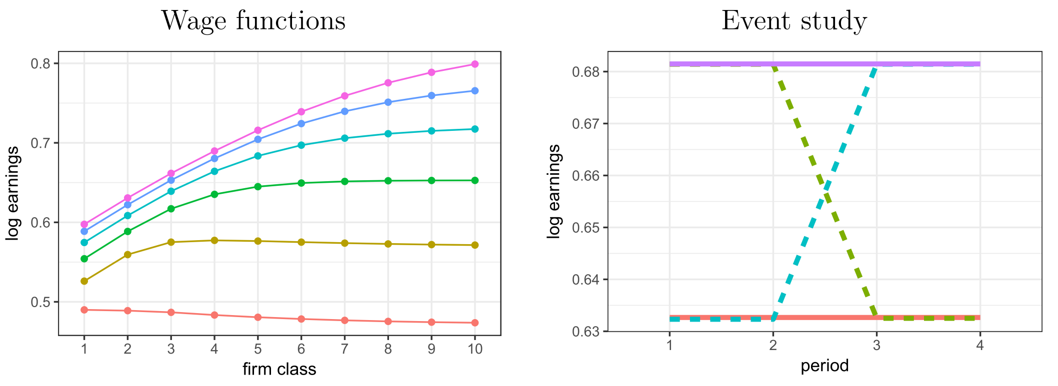

FIGURE S2.—Event study graph in the Shimer–Smith model in the presence of complementarities. Notes: The sample is generated according to the model of Shimer and Smith (2000), without on-the-job search. The parameter values imply positive assortative matching. In the left graph, we show log-wage functions for each worker type (y-axis), by firm class (x-axis). In the right graph, we show mean log-wages of workers moving between firms within classes 4 and 10 (solid), and moving between firms between classes 4 and 10 (dashed), between periods 2 and 3.

we show the mean log-wages of workers moving between firms within classes 4 and 10 (solid), and moving between firms between classes 4 and 10 (dashed). We see that, on the event-study graph, the wage changes are exactly symmetric around the move, yet wages are clearly non-additive in worker types and firm classes.

S2.2. *Simulation, Estimation, and Results*

We pick two parameterizations of the model associated with positive assortative matching (PAM) and negative assortative matching (NAM) in equilibrium. We set $b(x) = b = 0.3$, $c(y) = c = 0$, and $\bar{V} = 2$. We solve the model at a yearly frequency, and we set $\delta = 0.02$, $\lambda_0 = 0.4$, and $\lambda_1 = 0.1$. The production function is CES: $f(x, y) = a + (\nu x^\rho + (1-\nu)y^\rho)^{1/\rho}$, where we set $\nu = 0.5$ and $a = 0.7$. The relative variance of measurement error is set to 10%. Finally, we set either $\rho = -3$ (PAM) or $\rho = 3$ (NAM).

We simulate a sample of 500,000 individuals working in 5000 firms, in an economy with $K = 10$ firm classes and $L = 6$ worker types. In the left graph of Figure S3, we show means and quantiles of log-wages in the simulated samples. We see that, while mean log-wages are monotonic in firm productivity under PAM, they are non-monotonic under NAM. However, as we will see, there is sufficient variation in wage distributions to separate firm classes. In the middle graph of Figure S3, we report the wage functions for the different worker types. We see clear differences between PAM and NAM. Last, in the right graph, we show the wage functions as estimated by our static model. In estimation, we use the same procedure as on the Swedish data, with $K = 10$ and $L = 6$. In particular, we estimate firm classes using k-means clustering on empirical cdfs of log-wages evaluated at 20 grid points. The estimates seem to capture nonlinearities in log-wages remarkably well. Note that the ordering of firm classes on the x-axis is arbitrary, since the ranking of firms in terms of productivity (i.e., y) is not identified using wage information only. However, the variance decompositions and reallocation results we report below are not affected by this labeling indeterminacy.

In the first four rows of Table SI, we next report the results of variance decompositions using the samples generated according to the theoretical model, and based on estimates

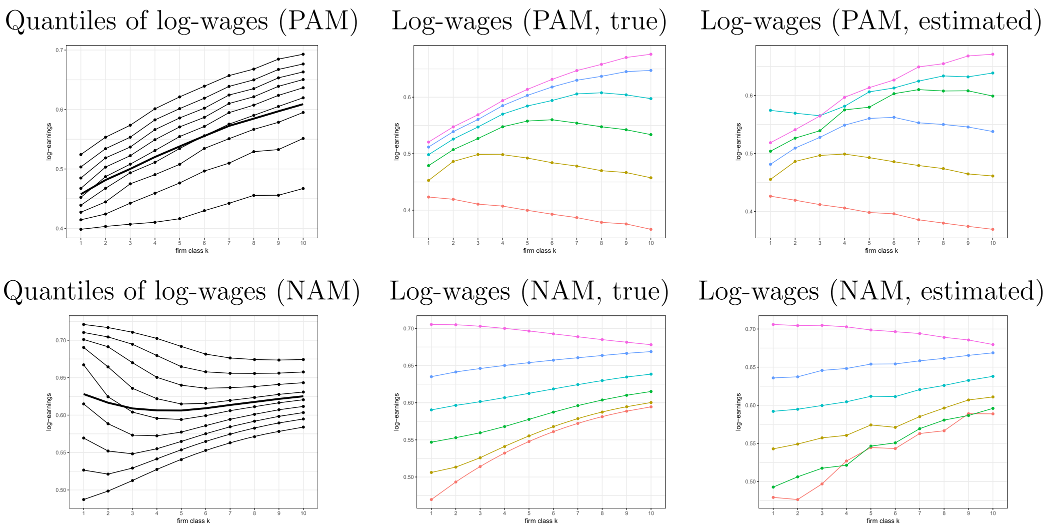

FIGURE S3.—True and estimated wage functions, and quantiles of log-wages, in a sample simulated according to the model of Shimer and Smith (2000) with on-the-job search. Notes: In the left graphs, we show deciles of log-wages (with measurement error) by firm class. The thick lines correspond to mean log-wages. In the middle graphs, we show log-wages (without measurement error), by worker type and firm class. In the right graphs, we show estimates from our static model. In the top panel, parameter values imply positive assortative matching, while in the bottom panel, parameter values imply negative assortative matching.

TABLE SI

VARIANCE DECOMPOSITION AND REALLOCATION EXERCISE ON A SAMPLE SIMULATED ACCORDING TO THE MODEL OF SHIMER AND SMITH (2000) WITH ON-THE-JOB SEARCH[a]

	Variance Decomposition ($\times$100)				
	$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha,\psi)$
			PAM		
True	64.5	4.5	12.8	18.1	37.5
Estimated	64.9	4.2	12.7	18.3	38.4
			NAM		
True	90.3	10.4	−16.8	16.1	−27.4
Estimated	89.5	9.2	−14.5	15.8	−25.3
	Reallocation Exercise ($\times$100)				
	Mean	Median	10%-Quantile	90%-Quantile	Variance
			PAM		
True	−1.1	−0.7	−1.5	−1.2	−0.1
Estimated	−1.2	−0.8	−1.6	−2.1	−0.1
			NAM		
True	−0.5	−0.5	−0.3	−0.7	0.0
Estimated	−0.4	−0.7	0.3	−0.8	0.0

[a]Notes: Variance decomposition and reallocation effects based on data generated from the theoretical sorting model of Shimer and Smith (2000) with on-the-job search. Parameter values imply either positive (in the PAM case) or negative (in the NAM case) assortative matching. See the notes to Table II.

from our static model. We see that the decomposition is very well reproduced under both PAM and NAM. In the last four rows, we report the results of an exercise where we randomly reallocate workers to firms. The results are again well reproduced by our static model. Overall, this set of results shows that our method is able to accurately recover the link between wages and worker-firm heterogeneity in simulated economies that feature positive or negative assortative matching.

S3. EXTENSIONS

In this section, we describe several extensions of our approach.

S3.1. *Model-Based Reclassification of Firms*

Here, we describe the model-based reclassification that we use as a robustness check. We describe the method for the static model, the strategy being very similar in the dynamic case.

Consider the log-likelihood functions in equations (13) and (14). Starting from an initial firm classification $\widehat{k}$, we first estimate $\widehat{\theta}_p$, $\widehat{\theta}_f$, and $\widehat{\theta}_{f^m}$ by maximizing (13), where the $\widehat{k}$'s correspond to an initial classification of firms. We then estimate $\widehat{\theta}_q$ by maximizing (14). With those estimates at hand, we next reclassify firm j, for all $j = 1, \dots, J$, as follows:

$$\begin{aligned}\widetilde{k}(j) = \underset{k=1,\dots,K}{\operatorname{argmax}} & \sum_{i=1}^{N} \mathbf{1}\{j_{i1} = j\} \ln\left(\sum_{\alpha=1}^{L} q_k(\alpha; \widehat{\theta}_q) f_{k\alpha}(Y_{i1}; \widehat{\theta}_f)\right) \\ & + \sum_{i=1}^{N_m} \mathbf{1}\{j_{i1} = j\} \sum_{k'=1}^{K} \mathbf{1}\{\widehat{k}_{i2} = k'\} \ln\left(\sum_{\alpha=1}^{L} p_{kk'}(\alpha; \widehat{\theta}_p) f_{k\alpha}(Y_{i1}; \widehat{\theta}_f) f^m_{k'\alpha}(Y_{i2}; \widehat{\theta}_{f^m})\right) \\ & + \sum_{i=1}^{N_m} \mathbf{1}\{j_{i2} = j\} \sum_{k'=1}^{K} \mathbf{1}\{\widehat{k}_{i1} = k'\} \ln\left(\sum_{\alpha=1}^{L} p_{k'k}(\alpha; \widehat{\theta}_p) f_{k'\alpha}(Y_{i1}; \widehat{\theta}_f) f^m_{k\alpha}(Y_{i2}; \widehat{\theta}_{f^m})\right).\end{aligned}$$

To iterate this approach further, we re-estimate the θ parameters given the new $\widetilde{k}$ classification, reclassify firms into classes based on the above formula (modified in order to account for the new parameter values), and so on.

S3.2. *A Two-Sided Random-Effects Approach*

Here, we describe a model with two-sided heterogeneity that we estimate using a random-effects strategy. This exercise, which is close in spirit to Beffy, Kamionka, Kramarz, and Robert (2003) and Abowd, McKinney, and Schmutte (2018), requires us to specify a probabilistic structure for the mobility patterns of workers between firms. In contrast, the two-step approach we develop in the paper is conditional on the network of worker and firm links, so a mobility model is not needed.

The model is as follows.

- Worker types α_i and firm classes k_j are drawn independently from multinomial distributions. We denote the probabilities as p^w_α and p^f_k.

- The initial class of the firm where i works is drawn from a multinomial model:

$$\Pr(k_{i1}=k|\alpha_i=\alpha,k_1,\dots,k_J)=\frac{\exp(q_{\alpha k})}{\sum_{k'=1}^{K}\exp(q_{\alpha k'})}. \tag{S2}$$

- The firm identifier in period 1 is drawn from a multinomial model:

$$\Pr(j_{i1}=j|\alpha_i=\alpha,k_{i1}=k,k_1,\dots,k_J)=1\{k_j=k\}\frac{\exp(s_j)}{\sum_{j'\in J_k}\exp(s_{j'})}, \tag{S3}$$

where s_j denotes the logarithm of the size of firm j in period 1, and J_k denotes the set of firms of class k.

- The wages Y_{i1} in period 1 are drawn conditional on α_i and k_{i1}, exactly as in our baseline static model.
- The probability that i changes firm is specified as

$$\Pr(m_{i1}=1|j_{i1}=j,\alpha_i=\alpha,k_1,\dots,k_J)=r_{\alpha k_j}.$$

- After a job move, the firm class of i in period 2 is drawn from a discrete distribution with probabilities

$$\Pr\big(k_{i2}=k'|j_{i1}=j,m_{i1}=1,\alpha_i=\alpha,k_1,\dots,k_J\big)=\frac{\exp(p_{\alpha k_j k'})}{\sum_{k''=1}^{K}\exp(p_{\alpha k_j k''})}.$$

- The firm identifier in period 2 is then drawn according to

$$\begin{aligned}&\Pr\big(j_{i2}=j'|j_{i1}=j,k_{i2}=k',m_{i1}=1,\alpha_i=\alpha,k_1,\dots,k_J\big)\\&\quad=\mathbf{1}\big\{j'\neq j\big\}\mathbf{1}\big\{k_{j'}=k'\big\}\frac{\exp(s_{j'})}{\sum_{j''\in J_{k'}\setminus j}\exp(s_{j''})},\end{aligned}$$

where $J_k\setminus j$ denotes the set of firms of class k, without firm j.

- The wages Y_{i2} in period 2 are then drawn conditional on α_i and k_{i2}, exactly as in our baseline model.

This model contains a large number of parameters. Moreover, due to the fact that workers and firms interact in a network, the likelihood function does not factor in a simple way. Specifically, as is often the case in network models, the likelihood involves an intractable sum over $k_1,\dots,k_J$. To see this, let $Z_i=(j_{i1},Y_{i1},j_{i2},Y_{i2})$ be the observed data for individual i (abstracting from firm size for simplicity). Let f be a generic notation for a density or probability function. The likelihood function takes the following form:

$$\begin{aligned}&\sum_{k_1,\dots,k_J}\sum_{\alpha_1,\dots,\alpha_N}f(k_1,\dots,k_J,\alpha_1,\dots,\alpha_N,Z_1,\dots,Z_N)\\&\quad=\sum_{k_1,\dots,k_J}\left(\prod_{j=1}^{J}p^f_{k_j}\right)\prod_{i=1}^{N}\sum_{\alpha_i=1}^{L}f(\alpha_i,Z_i|k_1,\dots,k_J),\end{aligned}$$

where (recall that $m_{i1} = \mathbf{1}\{j_{i2} \neq j_{i1}\}$)

$$\begin{aligned} &f(\alpha_i, j_{i1}, Y_{i1}, j_{i2}, Y_{i2} | k_1, \ldots, k_J) \\ &\quad = p^w_{\alpha_i} \frac{\exp(q_{\alpha_i k_{j_{i1}}})}{\sum_{k'=1}^{K} \exp(q_{\alpha_i k'})} \frac{\exp(s_{j_{i1}})}{\sum_{j' \in J_{k_{j_{i1}}}} \exp(s_{j'})} f(Y_{i1} | \alpha_i, k_{j_{i1}}) [1 - r_{\alpha_i k_{j_{i1}}}]^{1-m_{i1}} \\ &\qquad \times \left[r_{\alpha_i k_{j_{i1}}} \frac{\exp(p_{\alpha_i k_{j_{i1}} k_{j_{i2}}})}{\sum_{k'} \exp(p_{\alpha_i k_{j_{i1}} k'})} \frac{\exp(s_{j_{i2}})}{\sum_{j'' \in J_{k_{j_{i2}}} \setminus j_{i1}} \exp(s_{j''})} f(Y_{i2} | \alpha_i, k_{j_{i2}}) \right]^{m_{i1}}. \end{aligned}$$

Full-information maximum likelihood (FIML) is conceptually attractive, since it treats both worker and firm heterogeneity as stochastic, and it is expected to be consistent under correct specification, irrespective of the size of the firms. However, due to the complexity of the likelihood function, FIML estimation is computationally challenging. The key difficulty is to sample the firm classes, since firms are not independent of each other.

To understand the link between this two-sided random-effects approach and our baseline two-step approach, it is useful to note that the *conditional* likelihood function, given the firm classes $k_1, \ldots, k_J$ and the firm identifiers $j_{11}, j_{12}, \ldots, j_{N1}, j_{N2}$, is, denoting $Y_i = (Y_{i1}, Y_{i2})$,

$$\begin{aligned} &\sum_{\alpha_1, \ldots, \alpha_N} f(\alpha_1, \ldots, \alpha_N, Y_1, \ldots, Y_N | k_1, \ldots, k_J, j_{11}, \ldots, j_{N2}) \\ &\quad = \prod_{i=1}^{N} \sum_{\alpha_i=1}^{L} f(\alpha_i, Y_i | k_1, \ldots, k_J, j_{11}, \ldots, j_{N2}), \end{aligned}$$

which takes the standard, separable form of a likelihood function with random effects in single-agent panel data models. Hence, conditioning on the estimated firm classes, and the whole set of firm identifiers, results in a drastic simplification of the problem. This highlights a substantial computational advantage of our two-step approach, relative to the two-sided random-effects approach that we study in this subsection.

Estimation

The key step in the algorithm is to draw firm classes k_j. For this, we use information from both job stayers and job movers. The posterior probability of class k_j, given all other classes $k_{-j} = \{k_1, \ldots, k_{j-1}, k_{j+1}, \ldots, k_J\}$ and the data, is

$$\begin{aligned} &\Pr(k_j = k | Z_1, \ldots, Z_N, k_{-j}) \\ &\quad = p^f_k \times \prod_{i \in I_j} \sum_{\alpha=1}^{L} p^w_\alpha \frac{\exp(q_{\alpha k})}{\sum_{k'} \exp(q_{\alpha k'})} \cdot \frac{\exp(s_j)}{\exp(s_j) + \sum_{j' \in J_k \setminus j} \exp(s_{j'})} f(Y_{i1} | \alpha, k) [1 - r_{\alpha k}]^{1-m_{i1}} \\ &\qquad \times \left[r_{\alpha k} \frac{\exp(p_{\alpha k k_{j_{i2}}})}{\sum_{k'} \exp(p_{\alpha k k'})} \cdot \frac{\exp(s_{j_{i2}})}{\sum_{j' \in J_{k_{j_{i2}}} \setminus j} \exp(s_{j'})} f(Y_{i2} | \alpha, k_{j_{i2}}) \right]^{m_{i1}} \end{aligned}$$

$$
\times \prod_{j'\setminus j}\prod_{i\in I_{j'}}\sum_{\alpha=1}^{L} f(\alpha)\frac{\exp(q_{\alpha k_{j'}})}{\sum_{k'}\exp(q_{\alpha k'})}
$$

$$
\times \frac{\exp(s_{j'})}{1\{k_{j'}=k\}\exp(s_j)+\sum_{j''\in J_{k_{j'}}\setminus j}\exp(s_{j''})} f(Y_{i1}|\alpha,k_{j'})[1-r_{\alpha k_{j'}}]^{1-m_{i1}}
$$

$$
\times \left[r_{\alpha k_{j_{i1}}}\frac{\exp(p_{\alpha k_{j_{i1}} k})}{\sum_{k'}\exp(p_{\alpha k_{j_{i1}} k'})}\cdot\frac{\exp(s_j)}{\exp(s_j)+\sum_{j''\in J_k\setminus\{j,j'\}}\exp(s_{j''})} f(Y_{i2}|\alpha,k)\right]^{m_{i1}\cdot 1[j_{i2}=j]}
$$

$$
\times \left[r_{\alpha k_{j_{i1}}}\frac{\exp(p_{\alpha k_{j_{i1}} k_{j_{i2}}})}{\sum_{k'}\exp(p_{\alpha k_{j_{i1}} k'})}\right.
$$

$$
\left.\times \frac{\exp(s_{j_{i2}})}{1\{k_{j_{i2}}=k\}\exp(s_j)+\sum_{j''\in J_{k_{j_{i2}}}\setminus\{j,j'\}}\exp(s_{j''})} f(Y_{i2}|\alpha,k_{j_{i2}})\right]^{m_{i1}\cdot(1-1\{j_{i2}=j\})},
$$

where I_j denotes the set of workers in firm j in the first period.

Given the expression of the posterior probabilities for the class of a firm given the classes of all other firms, we use a stochastic EM algorithm to estimate the parameters of the model. The update of the model's parameters conditional on the firm classes is almost identical to the likelihood estimation step of our baseline two-step estimation. Here, however, we need to recover the latent firm classes $k_1, \dots, k_J$ in each of the E-steps. Using Gibbs sampling, we draw the k_j's one at a time, updating 200 firms at random in every E-step. We then update the other parameters of the model in a standard M-step, where we compute the posterior probabilities of worker types. Unlike the posterior distribution of firm classes, this sub-problem separates across individuals since it is conditional on the classification of firms. This inner EM algorithm is almost identical to the update rule in our baseline two-step method. The main difference is for the parameters $q_{\alpha k}$ and p_k^f, which we update using empirical frequencies of the firm classes based on the simulated draws.

Results

We fix $K = 10$ and $L = 6$, and report results based on two sets of initial parameter values: parameters obtained from a classification based on deciles of the AKM fixed effects, and parameter values based on the estimates of our static model. Since the AKM starting values are computed on connected components of a graph, we focus on the largest connected set, and for comparability we focus on the same sample for both sets of starting values. The largest connected component contains 389,451 job stayers, 17,205 job movers, and 8794 unique firms. In Figure S4, we report the main results from this exercise. Every step in the estimation algorithm is costly, since a very large number of firm classes needs to be updated every time. Producing each curve on those graphs required 10 days of computing time on a desktop computer using 15 cores. In addition, there is evidence in the figure that the Markov chains have not yet converged, particularly for the correlation between worker and firm effects.

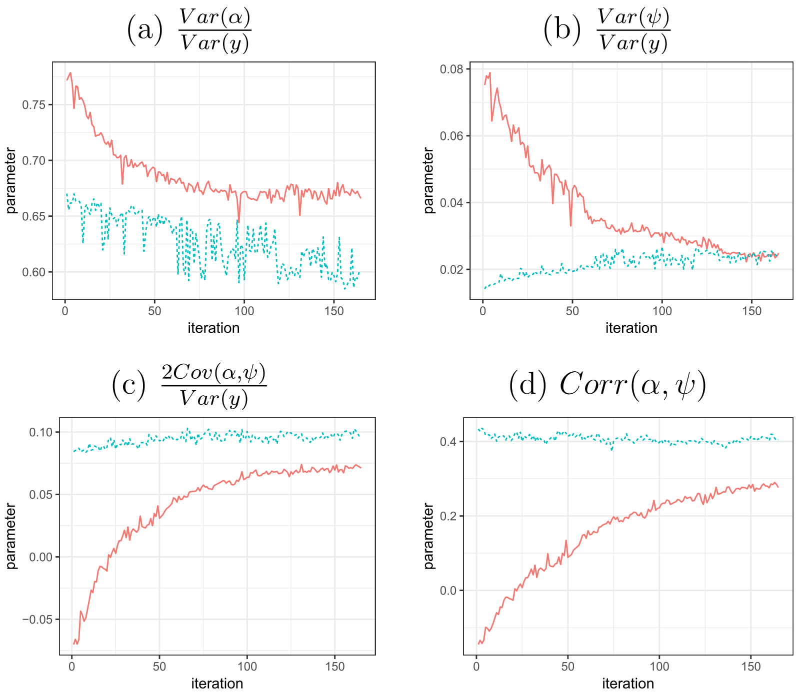

FIGURE S4.—Variance decomposition estimates based on two-sided random-effects estimation. Notes: Static model with two-sided heterogeneity, estimated using a two-sided random-effects estimator. The figure shows the different components of the variance decomposition (on the y-axis) as a function of the iterations of the stochastic EM algorithm (on the x-axis). In solid line, we report estimates using an AKM-based classification to produce starting parameter values. In dotted line, we show estimates using our k-means classification as a starting value.

Nevertheless, the results in Figure S4 are informative. We see evidence of convergence between the two sets of starting values. For example, the correlations between worker and firm effects, and the variances of firm effects, get closer to each other as the number of iterations increases. While the correlation coefficient is actually negative when starting from the AKM-based classification, it increases steadily over the course of the estimation. In the last iteration, both estimates indicate similar values for the variance components of log-earnings, particularly a small percentage explained by firms net of worker composition, around 3%, and a large correlation between worker and firm effects, between 30% and 40%. These magnitudes agree quite well with the baseline two-step estimates reported in the paper. Hence, though based on a different estimation approach, this exercise confirms the empirical results we obtained based on our two-step method. In addition, it illustrates that our one-sided random-effects approach with an initial classification of firms has computational advantages relative to a full random-effects approach, in a network context where firms and workers are linked to each other in complex ways.

S3.3. *Combining Our Approach With AKM*

Our approach is based on discretizing firm heterogeneity. This might result in a loss of information, since firms might be heterogeneous within classes. At the same time, the discrete classification might alleviate the incidental parameter biases due to the fact that some firms have very few job movers in the sample. In this subsection and the next, we study these issues in two ways. We start by giving firms that have many job movers their own firm classes, and by only applying the k-means classification to the firms that have few job movers. We use $K = 10$ groups in the classification, and focus on the variance of the firm component in an additive specification of the static model.

In Table SII, we show, in columns, estimation results based on different thresholds for the minimum number of job movers needed in order to give a firm its own class. For example, in the first column, we show results where firms with at least one mover are given their own class; hence, the corresponding estimator is the AKM fixed-effects estimator. In the last column we show the estimator where no firm is assigned its own class, and our two-step classification method is applied to all firms; hence, this last column is identical to our baseline two-step estimator with 10 classes on this sample (under additivity). The results show a high sensitivity to the number of job movers per firm. For example, we see a large jump in the variance of the firm effect when focusing on firms with at least two job movers instead of only one job mover: the variance drops from 32% to 14%. This instability suggests the presence of large incidental parameter biases in this sample.

In order to better understand the bias of the AKM fixed-effects estimator, we report the results of two existing bias-correction methods: the trace correction of Andrews, Gill, Schank, and Upward (2008), with a degree-of-freedom correction to estimate the variance of idiosyncratic shocks, and the recently proposed "leave-out" estimator of Kline, Saggio, and Sølvsten (2018), which is motivated by the fact that the correction of Andrews et al. (2008) does not allow for heteroscedasticity. This estimator is computed on a smaller sample, the "leave-out connected set." Both corrections tend to reduce the variance of firm effects significantly, supporting the conjecture that the AKM estimator is substantially biased in this sample. Moreover, compared to the AKM variance estimate of 32%, the estimates from the two corrections on the Swedish sample including all firms in the first column give 11% (Andrews et al. (2008), reported on the top panel) and 4.6% (Kline, Saggio, and Sølvsten (2018), reported on the bottom panel), respectively. This last number is larger than, but quantitatively close to, our baseline estimate.

S3.4. *Accounting for Within-Class Firm Heterogeneity*

Here, we extend our two-step approach to allow for within-class firm heterogeneity, in a static additive model of log-earnings. Consider the following model in first differences:

$$\Delta Y = A\psi + \Delta\varepsilon,$$

where $\Delta Y = Y_2 - Y_1$, $\Delta\varepsilon = \varepsilon_2 - \varepsilon_1$, and we have abstracted from i indices for simplicity. ψ is the vector of firm effects, and $A = A_2 - A_1$ is the difference of design matrices that map firms to workers; that is, $[A_t]_{ij} = \mathbf{1}\{j_{it} = j\}$ for $t = 1, 2$.

We assume that firm heterogeneity ψ is drawn conditional on A, according to a distribution with mean and variance matrix:

$$\mathbb{E}(\psi|A) = \mu, \qquad \mathrm{Var}(\psi|A) = D_\omega,$$

where D_ω is diagonal with $[D_\omega]_{jj} = [\sigma_\omega]_{\widehat{k}(j)}$, and μ is constant with firm classes. Here, the firm classes $\widehat{k}(j)$ are the ones estimated using the k-means approach. In turn, $\Delta\varepsilon$ has variance matrix $A_1 D_\varepsilon A_1' + A_2 D_\varepsilon A_2'$, where D_ε is diagonal with elements $[D_\varepsilon]_{jj} = [\sigma_\varepsilon]_{\widehat{k}(j)}$. If D_ω was equal to zero, the model would be identical to a simple additive version of our

TABLE SII

COMBINING OUR ESTIMATOR WITH AKM[a]

Min Number of Movers	1	2	3	4	5	10	25	50	100	150	Inf
n Firms With Own Class	8794	4709	2741	1813	1325	532	177	74	32	15	0
						Connected set					
n Firms	8794	8794	8794	8794	8794	8794	8794	8794	8794	8794	8794
Var(ψ), Abowd, Kramarz, and Margolis (1999)	0.0382	0.0169	0.0112	0.0086	0.0073	0.0050	0.0026	0.0021	0.0020	0.0015	0.0013
Var(ψ), Share	32.0%	14.1%	9.4%	7.2%	6.1%	4.2%	2.2%	1.8%	1.7%	1.3%	1.1%
Var(ψ) Corrected, Andrews et al. (2008)	0.0133	0.0063	0.0050	0.0044	0.0042	0.0037	0.0022	0.0019	0.0019	0.0015	0.0012
Var(ψ) Corrected, Share	11.2%	5.3%	4.2%	3.7%	3.5%	3.1%	1.8%	1.6%	1.6%	1.2%	1.0%
						Leave Out Connected Set					
n Firms	3439	8714	8778	8790	8790	8792	8794	8794	8794	8794	8794
Var(ψ), Abowd, Kramarz, and Margolis (1999)	0.0205	0.0166	0.0112	0.0086	0.0073	0.0050	0.0026	0.0021	0.0020	0.0015	0.0013
Var(ψ), Share	17.2%	13.9%	9.4%	7.2%	6.1%	4.2%	2.2%	1.8%	1.7%	1.3%	1.1%
Var(ψ) Corrected, Kline, Saggio, and Sølvsten (2018)	0.0054	0.0049	0.0045	0.0043	0.0042	0.0036	0.0022	0.0019	0.0020	0.0015	0.0012
Var(ψ) Corrected, Share	4.6%	4.1%	3.8%	3.6%	3.5%	3.0%	1.8%	1.6%	1.6%	1.2%	1.0%

[a]Notes: Estimates of a static additive model, on 2002–2004. We report estimates of the variance of firm effects. In the top panel, we show uncorrected variances and corrected estimates based on Andrews et al. (2008). In the bottom panel, we show uncorrected variances and corrected estimates based on Kline, Saggio, and Sølvsten (2018), on a subsample. Each column corresponds to a particular number of firms that are given their own class, the remaining firms being clustered into ten groups using k-means.

static model with class-specific heterogeneity. Permitting $D_\omega \neq 0$ allows for within-class firm heterogeneity.

We are interested in estimating the $K \times 1$ parameter vectors σ_ω and σ_ε. Let us denote as C the firm-class selection matrix with elements $C_{jk} = 1\{\widehat{k}(j) = k\}$. We have that $\mu = C\tilde{\mu}$, for a $K \times 1$ vector $\tilde{\mu}$. The vector $\tilde{\mu}$ is identified from mean restrictions. Moreover, variance restrictions imply

$$\mathbb{E}\big[\Delta Y \Delta Y'\big] - AC\tilde{\mu}\tilde{\mu}'C'A' = AD_\omega A' + A_1 D_\varepsilon A_1' + A_2 D_\varepsilon A_2'.$$

Hence, using standard matrix notation, we obtain that

$$\operatorname{vec}\big(\mathbb{E}\big[\Delta Y \Delta Y'\big] - AC\tilde{\mu}\tilde{\mu}'C'A'\big) = (A \otimes A)\operatorname{vec}(D_\omega) + (A_1 \otimes A_1 + A_2 \otimes A_2)\operatorname{vec}(D_\varepsilon),$$

which is a linear system of equations with $2K$ parameters. For a suitable selection matrix P, we thus have

$$\operatorname{vec}\big(\mathbb{E}\big[\Delta Y \Delta Y'\big] - AC\tilde{\mu}\tilde{\mu}'C'A'\big) = (A \otimes A)P\sigma_\omega + (A_1 \otimes A_1 + A_2 \otimes A_2)P\sigma_\varepsilon.$$

We estimate σ_ω and σ_ε from empirical counterparts of these restrictions. The matrix $\mathbb{E}[\Delta Y \Delta Y'] - AC\tilde{\mu}\tilde{\mu}'C'A'$ is constructed using all pairs of products of workers. Many such cross-products are in fact irrelevant. Indeed, if workers i and i' are such that $j_{i1} \neq j_{i'1}$, $j_{i2} \neq j_{i'2}$, $j_{i2} \neq j_{i'1}$, and $j_{i1} \neq j_{i'2}$, then the corresponding elements will be zero. To simplify computation we directly focus on relevant pairs, and in particular on pairs such that $j_{i1} = j_{i'1}$ and $j_{i2} \neq j_{i'2}$, and pairs such that $j_{i1} \neq j_{i'1}$ and $j_{i2} = j_{i'2}$.

Results

In Figure S5, we report the estimates of the between-and-within-class variances of firm effects, estimated according to the model with within-class firm heterogeneity (solid line). In the same graph, we also plot the between-class firm variance (dotted line). We see

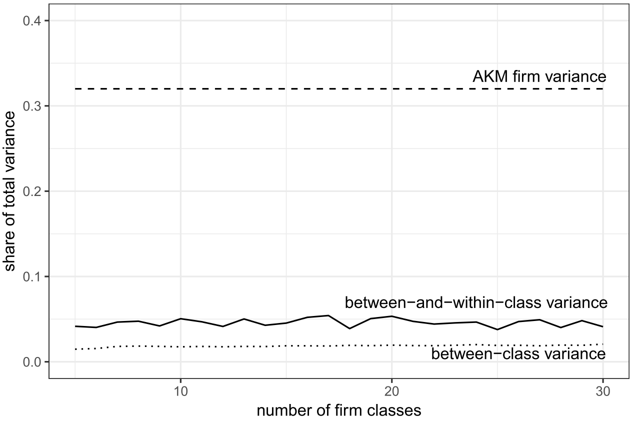

FIGURE S5.—Share of log-earnings variance explained by firm effects, static additive model. Notes: Estimates of a static additive model, on 2002–2004. The figure shows the share of total log-earnings variance explained by between-class variance (in dotted), the share explained by between-and-within-class firm variance (in solid), and the share explained by firms according to the AKM fixed-effects estimates (in dashed). We report the number of classes K on the x-axis. We describe our method that allows for within-class firm heterogeneity in Section S3.

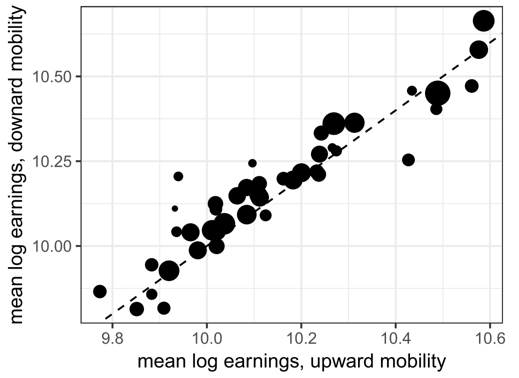

FIGURE S6.—Earnings of job movers. Notes: The figure shows average log-earnings in 2002 and 2004 of workers who moved from firm class k to firm class k' (on the x-axis), and of workers who moved from firm class k' to firm class k (on the y-axis), where $k < k'$. The size of the dots is proportional to the number of job movers in the cells.

that allowing for within-class dispersion increases the contribution of firm effects to the overall log-earnings variance by another 2%, for a total firm component of approximately 4%. As a comparison, we plot in dashed line the share of variance explained by the AKM fixed effects, which amounts to 32%. While suggesting that within-class dispersion—which we abstract from in our approach—may contribute additional firm heterogeneity, these results confirm the overall evidence of a small contribution of firm effects to log-earnings dispersion in our data.

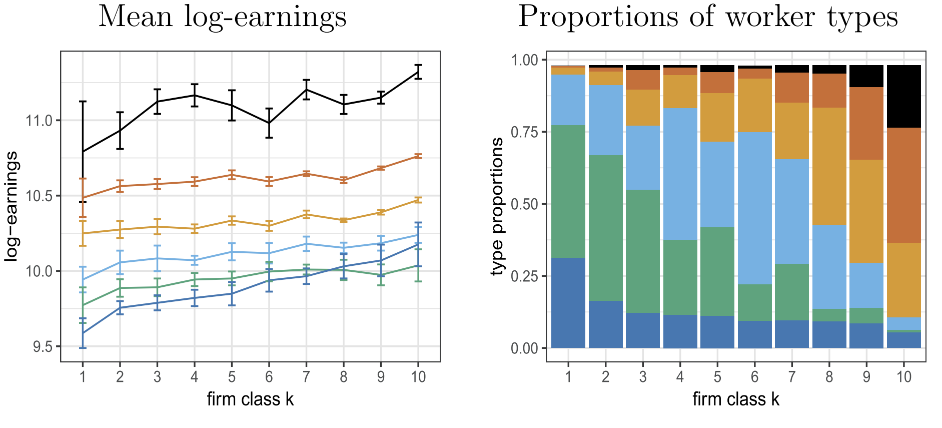

FIGURE S7.—Parameter estimates of the dynamic model. Notes: Estimates of the dynamic model, on 2001–2005. See the notes to Figure 2.

TABLE SIII

DATA DESCRIPTION, BY ESTIMATED FIRM CLASSES, IN THE SAMPLE USED TO ESTIMATE THE DYNAMIC MODEL[a]

Class:	1	2	3	4	5	6	7	8	9	10	All
Number of Workers	13,535	39,048	45,293	50,435	63,433	45,175	81,878	34,098	49,187	20,675	442,757
Number of Firms	5017	5565	3390	4695	2836	3763	2316	3300	3160	2886	36,928
Mean Firm Reported Size	13.37	21.24	50.00	28.05	70.57	30.10	89.69	30.36	63.28	54.43	39.67
Number of Firms ≥ 10 (Actual Size)	125	814	938	871	948	693	773	470	632	369	6633
Number of Firms ≥ 50 (Actual Size)	2	78	155	137	214	110	190	81	134	58	1159
% High School Drop Out	30.4%	28.9%	26.8%	27.6%	23.8%	25.4%	19.1%	18.5%	9.2%	3.9%	21.5%
% High School Graduates	60.6%	62.9%	61.0%	63.4%	59.9%	63.4%	58.6%	57.4%	40.6%	29.0%	57.0%
% Some College	9.0%	8.2%	12.2%	9.0%	16.3%	11.3%	22.2%	24.1%	50.2%	67.1%	21.4%
% Workers Younger Than 30	18.9%	15.8%	16.8%	15.4%	15.5%	13.2%	12.4%	10.5%	11.6%	11.0%	13.9%
% Workers Between 31 and 50	57.6%	56.6%	57.8%	58.3%	58.1%	59.2%	60.6%	59.8%	62.0%	64.5%	59.4%
% Workers Older Than 51	23.5%	27.6%	25.4%	26.3%	26.4%	27.6%	27.0%	29.8%	26.4%	24.5%	26.7%
% Workers in Manufacturing	29.7%	44.5%	51.1%	53.4%	54.9%	50.7%	58.1%	47.0%	40.1%	11.4%	48.5%
% Workers in Services	32.1%	29.0%	17.7%	19.9%	15.9%	15.7%	9.8%	23.9%	37.7%	65.1%	22.4%
% Workers in Retail and Trade	27.1%	17.2%	27.8%	9.4%	25.1%	8.8%	9.7%	10.9%	16.8%	22.2%	16.3%
% Workers in Construction	11.1%	9.4%	3.3%	17.2%	4.2%	24.7%	22.4%	18.1%	5.5%	1.2%	12.8%
Mean Log-Earnings	9.75	9.94	10.04	10.06	10.15	10.15	10.24	10.30	10.45	10.73	10.19
Variance of Log-Earnings	0.073	0.044	0.085	0.043	0.086	0.040	0.080	0.056	0.105	0.152	0.113
Skewness of Log-Earnings	−1.568	−0.505	0.562	0.131	0.720	0.265	0.638	0.801	0.537	1.202	0.754
Kurtosis of Log-Earnings	9.483	15.061	7.944	18.123	7.985	15.611	8.999	12.144	6.409	7.157	7.489
Between-Firm Variance of Log-Earnings	0.0337	0.0033	0.0033	0.0016	0.0022	0.0013	0.0019	0.0023	0.0063	0.0436	0.0441
Mean Log-Value-Added per Worker	12.45	12.61	12.72	12.69	12.85	12.75	12.94	12.85	13.01	13.18	12.76

[a]Notes: The table corresponds to males fully employed in the same firm in 2001–2002 and 2004–2005, for firms that are continuously present in the sample. The "actual size" is the number of workers per firm in our sample. All numbers in the table correspond to 2002.

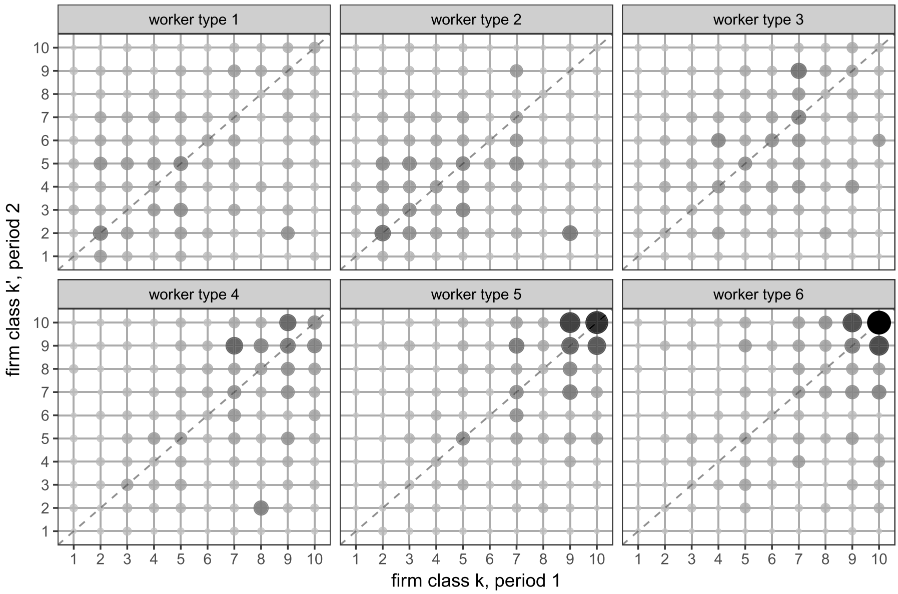

FIGURE S8.—Mobility across firm classes, dynamic model. Notes: Estimates of the dynamic model, on 2001–2005. Estimated joint probability of firm classes in 2002 (on the x-axis) and 2004 (on the y-axis) for job movers. The size of the dots is proportional to the number of job movers in the cells.

TABLE SIV

NUMBER OF JOB MOVERS BETWEEN FIRM CLASSES[a]

	Firm Class in Period 2									
	1	2	3	4	5	6	7	8	9	10
Firm Class in Period 1										
1	76	120	95	87	94	50	63	42	29	10
2	129	348	352	237	241	164	147	76	62	18
3	128	292	417	349	399	203	217	146	126	49
4	59	318	304	356	249	303	210	102	68	24
5	60	190	502	294	424	235	271	198	139	64
6	39	115	154	267	172	230	275	128	79	32
7	48	158	204	253	355	355	363	331	457	100
8	14	315	145	110	243	157	348	461	609	258
9	11	77	114	187	217	195	323	384	368	402
10	12	21	83	39	114	27	161	229	313	369

[a]Notes: The table corresponds to males fully employed in the same firm in 2002 and 2004, for firms that are continuously present in the sample. We report the number of job movers from firm class k in 2002 (on the vertical axis) to firm class k' in 2004 (on the horizontal axis).

TABLE SV

VARIANCE DECOMPOSITION AND REALLOCATION EXERCISE, STATIC MODEL, MONTE CARLO EXERCISE[a]

	Variance Decomposition ($\times 100$)				
	$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha,\psi)$
True value	60.03	2.56	12.17	25.24	49.13
Monte Carlo mean	61.2	2.1	11.7	25.1	52.2
Monte Carlo 2.5%-quantile	59.6	1.7	10.8	23.9	50.4
Monte Carlo 97.5%-quantile	62.6	2.3	12.4	26.2	53.8
AKM mean	120.9	37.6	−62.6	4.1	−46.3
	Reallocation Exercise ($\times 100$)				
	Mean	Median	10%-Quantile	90%-Quantile	Variance
True value	0.50	0.58	2.60	−1.24	−1.12
Monte Carlo mean	0.63	0.60	2.77	−0.92	−1.07
Monte Carlo 2.5%-quantile	0.43	0.36	2.42	−1.44	−1.26
Monte Carlo 97.5%-quantile	0.81	0.83	3.23	−0.27	−0.84

[a]Notes: See the notes to Table II. We use the parameter estimates of the static model on 2002–2004 as true parameter values and simulate samples based on them. We show mean and percentiles of the Monte Carlo distribution, computed using 200 replications. In the fifth row of the table, we also show the Monte Carlo means of AKM estimates.

TABLE SVI

VARIANCE DECOMPOSITION AND REALLOCATION EXERCISE, DYNAMIC MODEL, MONTE CARLO EXERCISE[a]

	Variance Decomposition ($\times 100$)				
	$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha,\psi)$
True value	60.27	4.24	13.40	22.09	41.90
Monte Carlo mean	60.3	4.1	13.9	21.7	44.3
Monte Carlo 2.5%-quantile	56.1	3.6	13.1	20.9	36.6
Monte Carlo 97.5%-quantile	61.7	6.3	14.5	23.8	46.6
	Reallocation Exercise ($\times 100$)				
	Mean	Median	10%-Quantile	90%-Quantile	Variance
True value	0.26	0.80	2.57	−3.24	−1.05
Monte Carlo mean	0.20	0.78	2.46	−3.05	−0.70
Monte Carlo 2.5%-quantile	−0.18	0.43	0.56	−3.84	−1.57
Monte Carlo 97.5%-quantile	0.41	1.15	3.00	−1.84	2.43

[a]Notes: See the notes to Table II. We use the parameter estimates of the dynamic model on 2001–2005 as true parameter values and simulate samples based on them. We show mean and percentiles of the Monte Carlo distribution, computed using 200 replications.

TABLE SVII

VARIANCE DECOMPOSITION ($\times 100$), STATIC MODEL, RECLASSIFYING FIRMS[a]

	$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha, \psi)$
			A. Starting Using AKM Deciles		
Initial Value	84.53	12.28	−16.67	19.86	−25.87
1 Iteration	67.94	7.50	2.57	21.98	5.72
10 Iterations	63.01	2.80	8.03	26.17	30.26
			B. Starting Using Value Added Deciles		
Initial Value	70.18	0.49	3.11	26.22	26.38
1 Iteration	62.71	2.98	10.15	24.16	37.11
10 Iterations	57.54	4.49	10.98	27.00	34.14
			C. Starting Using Poaching Rank Deciles		
Initial Value	74.11	0.16	0.89	24.85	13.05
1 Iteration	61.51	2.09	9.65	26.76	42.60
10 Iterations	61.22	3.01	8.64	27.14	31.82
			D. Starting Using Share of Movers Deciles		
Initial Value	71.92	0.06	0.17	27.84	4.13
1 Iteration	65.90	2.20	7.91	23.98	32.80
10 Iterations	61.07	4.56	11.33	23.04	33.95

[a]Notes: Estimates of the static model, on 2002–2004. On each panel we show the results of the model-based reclassification described in Section S3, starting from a different classification of firms: deciles of the firm effects estimates of Abowd, Kramarz, and Margolis (1999), deciles of log value added per worker, deciles of the poaching rank measure of Bagger and Lentz (2014), and deciles of the firm-specific shares of job movers.

REFERENCES

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333. [15,21]

ABOWD, J. M., K. L. MCKINNEY, AND I. M. SCHMUTTE (2018): "Modeling Endogenous Mobility in Earnings Determination," *Journal of Business & Economic Statistics*, 1–14. [9]

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?" *J. R. Stat. Soc. Ser. A Stat. Soc.*, 171, 673–697. [14, 15]

BAGGER, J., AND R. LENTZ (2014): "An Empirical Model of Wage Dispersion With Sorting," Report, University of Wisconsin. [21]

BEFFY, M., T. KAMIONKA, F. KRAMARZ, AND C. ROBERT (2003): "Job Mobility and Wages With Worker and Firm Heterogeneity," Tech. rep., Crest working paper. [9]

BONHOMME, S., AND E. MANRESA (2015): "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 83, 1147–1184.

HAHN, J., AND H. R. MOON (2010): "Panel Data Models With Finite Number of Multiple Equilibria," *Econometric Theory*, 26, 863–881.

JOCHMANS, K., AND M. WEIDNER (2017): "Fixed-Effect Regressions on Network Data," Unpublished Manuscript. [2]

KLINE, P., R. SAGGIO, AND M. SØLVSTEN (2018): "Leave-out Estimation of Variance Components," ArXiv e-prints. [14,15]

SCHWARZ, G. (1978): "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461–464.

SHIMER, R. (2006): "On-the-Job Search and Strategic Bargaining," *Eur. Econ. Rev.*, 50, 811–830. [4]

SHIMER, R., AND L. SMITH (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–369. [3,6-8]

Co-editor Aviv Nevo handled this manuscript.

Manuscript received 25 September, 2017; final version accepted 21 December, 2018; available online 18 February, 2019.