Online Appendix

Identifying Sorting in Practice

by Cristian Bartolucci, Francesco Devicienti and Ignacio Monzón

Contents

1. Robustness Checks

Our identification strategy relies on the availability of random draws from the distribution of partners and wages. We focus on matches mediated by an unemployment spell to guarantee independence. If workers search on-the job, partners' types and wages after an unemployment spell are not necessarily random draws from steady state. In Section 1.1, we present evidence suggesting that this concern does not drive our results. Moreover, information spillovers from former coworkers may generate correlation between employer types, even conditioning by worker type. In Section 1.2, we challenge our assumption of conditional independence between two subsequent partners.

We have maintained the simplifying assumption that job destruction is only exogenous. This is a common assumption in models describing the labor market and allows us to interpret firm types before unemployment as random draws from the distribution of partners. However, some firms may be more likely to layoff workers than others. Similarly, some workers may be more likely to be fired than others. These potential sources of selection bias are analyzed and discarded in Sections 1.3 and 1.4 respectively.[1]

1.1 Identification of Sorting with On-the-Job Search

We take advantage of the longitudinal dimension of our data set to make inference on the strength and sign of sorting when draws out of unemployment are not necessarily from steady state. Intuitively, workers may be less selective from unemployment if they can continue to search while on the job. Over time workers change jobs, and eventually the effect of the unemployment spell fades away.[2] Let $p_{i,t}$ be the type of the employer t periods after the beginning of the unemployment spell. As t grows, the distribution of employers converges to the steady state distribution.

We measure the strength of sorting η through the correlation of independent draws of employer types. Let p_i^{PREV} denote the employer type before unemployment. The corre-

[1]For all robustness checks we order firms by economic profits (aggregated and per-worker). Results using different measures of profits do not differ significantly.

[2]When the process that drives transitions (both in and out of unemployment, and job-to-job) is ergodic, the distribution of partners converges to its steady state.

lation $\eta_t^2 \equiv \text{cov}\left(p_i^{PREV}, p_{i,t}\right)/\sigma_p^2$ converges to the correlation between two independent draws from the steady state distribution as t grows. Panels (A) and (B) in Figure A1 present estimates of η_t^2 over time. For low values of t the correlation η_t^2 increases as t grows. However, after approximately a year η_t^2 becomes stable around the values from Section 4.1 in the paper. The fact that η_t^2 is low when t is small suggests that after an unemployment spell, workers are less discriminating in terms of which firms they accept, and therefore there is weaker sorting.

We follow a similar line of reasoning to identify the sign of sorting. We study whether better workers move to better firms (as in Panel (E) of Table 2 in the paper). We focus on the correlation γ_t between wages w_i^{PREV} before unemployment and the type $p_{i,t}$ of the employer t periods after the unemployment spell starts. The correlation γ_t converges to the correlation between random draws from the steady state distribution of wages and partner types as t grows. Panels (C) and (D) in Figure A1 present estimates of γ_t. The sign of sorting is always positive (and γ_t is significantly different from zero for all t except for $t = 12$, with firms ordered by average profits).

1.2 Serially Correlated Transitory Component in Employer's type

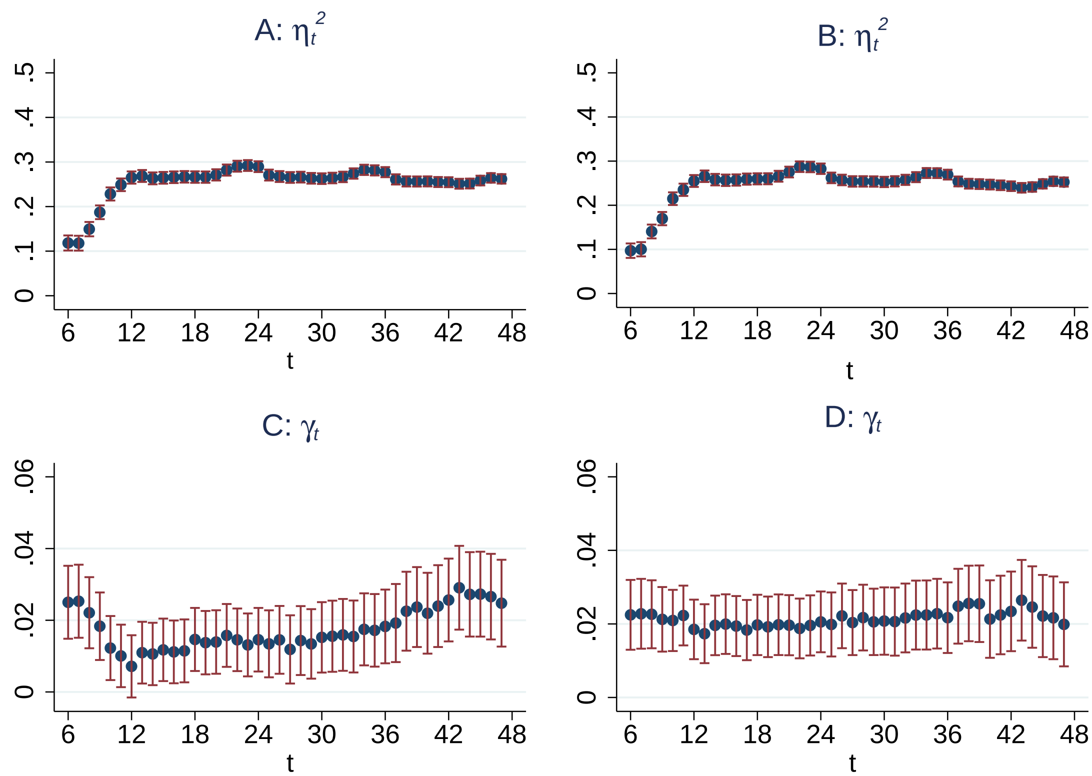
Employers' types before and after an unemployment spell are not necessarily independent. Workers may find new jobs exploiting networks of former fellow workers (see Cingano and Rosolia [2012]). Then, employer types before and after an unemployment spell may be correlated, even conditioning on worker type. We challenge our assumption of conditional independence allowing for a serially correlated transitory component in the variation of employer types. Let worker i's employer type be given by

$$p_{ij} = \phi_i + \omega_{ij}, \tag{1}$$

where $\omega_{ij} = \zeta\omega_{i,j-1} + \nu_{ij}$ and ν_{it} is white noise.

We use information on individuals with at least two transitions. If only two partners are observed, unobserved heterogeneity (σ_ϕ) and individual dynamics ζ cannot be distinguished (see Arellano [2003]). The model is just identified for those workers for whom

Figure A1: Convergence of η_t^2 and γ_t



Notes: This figure presents estimates for the strength and the sign of sorting. The subindex t indicates the number of months elapsed between the beginning of unemployment and the period in which the worker is observed working on a firm of type p_t. Panels (A) and (B) present estimates of η_t^2 for t ranging between 6 and 48. Panels (C) and (D) present estimates of γ_t for t ranging between 6 and 48. Red bars indicate 95% confidence intervals. We rank firms in terms of average profit in Panels (A) and (C). We rank firms in terms of average profit per worker in Panels (B) and (D). For each t, η_t and γ_t are estimated on a sample of workers that are observed at least for 48 months after an unemployment spell.

we observe three partners. We use standard panel data techniques to analyze the covariance structure of models with dynamic error components. Consider a sample of workers observed in three consecutive jobs. We observe employer types $(p_{ij}^{PREV}, p_{ij}, p_{ij}^{POST})$. As described in Arellano [2003], a model with a heterogeneous permanent component and an AR(1) transitory component is summarized by the three following variance restrictions:

$$\mathrm{var}(p_{ij}^{PREV}) = \mathrm{var}(p_{ij}) = \mathrm{var}(p_{ij}^{POST}) = \sigma_\phi^2 + \sigma_\omega^2$$

Table A1: Estimates of strength of sorting η with serially correlated transitory component in partner types

	Excluding Job-to-Job Transitions		Including Job-to-Job Transitions	
	Economic profits per worker (1)	Economic profits total (2)	Economic profits per worker (3)	Economic profits total (4)
η	0.623	0.614	0.563	0.588
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
ζ	0.054	0.052	0.025	0.005
	(< 0.001)	(< 0.001)	(0.002)	(0.475)

Notes: This table presents estimates for the strength of sorting η when employer types may be correlated, even conditional on the worker type. Number of individuals in columns (1) and (2) is $20,747$. Number of individuals in columns (3) and (4) is $28,223$. P-values - in parentheses - are obtained by bootstrap based on $1,000$ re-samples. Bootstrap samples from the pool of workers.

$$\mathrm{cov}(p_{ij}^{PREV}, p_{ij}) = \mathrm{cov}(p_{ij}, p_{ij}^{POST}) = \sigma_\phi^2 + \zeta\sigma_\omega^2$$
$$\mathrm{cov}(p_{ij}^{PREV}, p_{ij}^{POST}) = \sigma_\phi^2 + \zeta^2\sigma_\omega^2$$

We can recover σ_ϕ^2, ζ and σ_v^2 from the variance of partners, the covariance between the current partner and previous partner, and the covariance between the previous partner and the next partner. Therefore we can construct $\eta = \sqrt{\frac{\sigma_\phi^2}{\sigma_\phi^2+\sigma_v^2}}$.

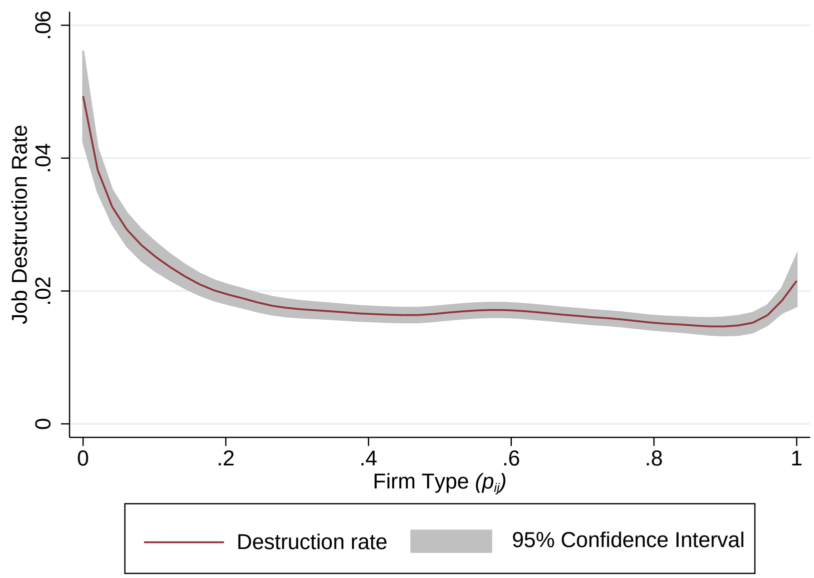
Columns (1) and (2) of Table A1 present our results. Estimated η are slightly larger than those presented in Table 1 in the paper. The estimated ζ suggests that the serial correlation in the transitory component in partner's types is weak but significantly different than zero. Since we allow for serial correlation in the transitory component of the employer type, we can also exploit information contained in job-to-job transitions. Estimates of the strength of sorting including job-to-job transitions are presented in columns (3) and (4) of Table A1, and are in line with those presented in Table 1 in the paper.

1.3 Selection of Firms

We next show that job destruction is associated with firm type, incorporate this in our estimates of the strength and sign of sorting, and show that results do not change significantly. Firms that are more likely to layoff workers appear more often as previous employers in a sample of workers who transit unemployment. Our identification strategy relies on observing random draws from the steady state distribution of partners, so we must account for the over-representation of firms with higher destruction rates.

We calculate the monthly firm-specific destruction rate as the fraction of employees observed in unemployment in the following month. Although firm-specific destruction rates are significantly heterogeneous, we find a clear pattern between destruction rates and firm type. Figure A2 presents a non-parametric regression of the firm-specific destruction rate on the firm type. Firms of worse types are more likely to layoff workers than those of higher types.

Figure A2: Destruction Rate by Firm Type



Notes: This figure presents estimates of the destruction rate as a function of firm type. Firms are ordered by economic profits per worker. Kernel non parametric regression based on 178, 219 observations. Kernel type is Epanechnikov.

Table A2 shows estimates of the strength (η) and sign (γ) of sorting weighting each observation by the inverse of the destruction rate corresponding to the previous employer. The estimated strength of sorting η is similar to that reported in Section 4 in the paper (the range of η in Table A2 is between 0.49 and 0.54, compared to a range between 0.51 and 0.53 in Table 1 in the paper). The sign of sorting γ is positive and significant in all specifications.

Table A2: Strength η and Sign γ of Sorting, Weighted by Firm-specific Destruction Rates

	Economic Profits per worker		Total Economic Profits	
η	0.491	(< 0.001)	0.536	(< 0.001)
γ	0.033	(0.009)	0.058	(0.008)

Notes: This table presents estimates for the strength and for the sign of sorting when the destruction rate may be firm-specific. Observations are weighted by the inverse of the destruction rate estimated for each firm. There are 178,219 observations for the estimation of the strength of sorting η. There are 120,426 observations for the estimation of the sign of sorting. P-values in parentheses.

1.4 Selection of Workers

Workers who are laid off are potentially different from those who do not transit unemployment. Our results are consistent for the group of workers who transit unemployment. However, that group is a non-random sample of workers. Their sorting pattern may be different from that of other workers.

We consider firms that layoff their complete workforce. In this case, all workers are forced to leave the firm, irrespective of their characteristics.[3] In our data it is possible to identify 710 firms that closed their business during the 1995-2001 time period, involving 15,255 workers. We obtain estimates of the sign and strength of sorting for this subsample (see Table A3). Despite this dramatic reduction in sample size, the results are once again indicative of positive sorting, with a similar strength to that of our baseline analysis. We find η close to 50% and γ positive and statistically significant.

1.5 Strength of Sorting: Robustness with Respect to Choice of Sample

Table A4 compares estimates of the strength of sorting (η) obtained with the narrower sample of transitions with "clean wages" (column (3) of Table 5 in the paper) and with the original sample of transitions (column (2) of Table 5 in the paper). The results, displayed respectively in rows (1) and (2) of Table A4, are very similar.

[3] Cingano and Rosolia [2012] use a similar strategy to identify the strength of information spillovers on workers' unemployment duration.

Table A3: Strength η and Sign γ of Sorting Estimated from Firm Closures

	Economic Profits per worker		Total Economic Profits	
η	0.496	(< 0.001)	0.478	(< 0.001)
γ	0.016	(0.009)	0.017	(0.008)

Notes: This table presents estimates for the strength and for the sign of sorting only from observations from firm closures. γ is the coefficient of a regression of the new employer's firm type on the worker's wage percentile in the current firm. The regression used to obtain γ includes controls for worker's age, age squared, tenure, tenure squared, time dummies and indicators for females, foreign-born workers, blue collar, white collars and managerial occupations. Firm fixed-effects and time fixed-effects are also included. P-values in parentheses. Number of observations is $15,255$.

Table A4: Estimates of the strength of sorting η

	Economic profits		GOS		AP		ROE
	per worker	total	per worker	total	per worker	total	
(1)	0.540	0.517	0.540	0.527	0.521	0.531	0.541
(2)	0.526	0.520	0.527	0.523	0.510	0.530	0.523

Notes: This table compares the estimates for the strength of sorting η from two possible samples. Row (1) refers to the narrower sample of transitions with "clean wages" (column (3) of Table 5 in the paper). Row (2) refers to the original sample of transitions (column (2) of Table 5 in the paper). Number of observations is $120,426$ in the first row and $178,219$ in the second row. All estimates are statistically significant: p-values < 0.001.

2. The Cost of Mismatch. Details

We follow Atakan [2006] and Eeckhout and Kircher [2011]. Consider the model with explicit and constant search costs and job scarcity as introduced in Section 5.2 of Eeckhout and Kircher [2011]. There is a discrete time, stationary economy, populated by risk neutral and infinitely lived firms and workers. Firms are a collection of jobs characterized by their productivity $p \in [0, 1]$. Each firm has N jobs. Worker types are denoted by $\varepsilon \in [0, 1]$. For simplicity, assume symmetry in the distribution of jobs and workers. $G(\cdot)$ denotes the stationary distribution of unmatched types.

Unmatched workers and jobs meet a potential partner every period. There is no on-the-job search in this simple model; hence, movements of workers between firms feature an interim unemployment spell. When two unmatched agents meet, they immediately observe each other's type. They match only if they both agree. As in Eeckhout and Kircher [2011], matched workers and jobs disappear from the market until there is an exogenous destruction of the match, at which point they return to the market. Each period a match is destroyed with probability δ.

The match (p, ε) produces $f(p, \varepsilon)$ per period, with $f_p > 0$ and $f_\varepsilon > 0$. We assume that the output of a firm p is the sum of the output of its matched jobs. A worker ε employed by a firm p receives $w(p, \varepsilon)$ and the firm receives $\pi(p, \varepsilon)$. Since payoffs exhaust match output, $f(p, \varepsilon) = w(p, \varepsilon) + \pi(p, \varepsilon)$. Unemployed workers and vacancies pay a constant cost equal to c if they reject the potential partner that they have met. The value $v(\varepsilon)$ of being unemployed for a worker of type ε is given by

$$v(\varepsilon) = \int_{p \in M(\varepsilon)} w(p, \varepsilon) dG(p) + \int_{p \notin M(\varepsilon)} [v(\varepsilon) - c] dG(p). \tag{2}$$

where $M(\varepsilon)$ is the set of acceptable partners for a worker of type ε. The value of a vacant job for a firm p is defined equally due to symmetry. Payoffs are determined by symmetric Nash Bargaining.

This model provides a convenient framework to describe our methodology. The following lemma provides a simple characterization of the monotonicity of payoffs and value functions.

Lemma 1 *Values $v(\varepsilon)$ are increasing in ε. Wages $w(p,\varepsilon)$ are increasing in ε, given p.*

Proof. Take two workers $\tilde{\varepsilon}$ and ε, with $\tilde{\varepsilon} > \varepsilon$. As wages are set by Nash-Bargaining, $w(p,\varepsilon) = \frac{1}{2}S(p,\varepsilon) + v(\varepsilon) - c$. $S(p,\varepsilon)$ represents the gains from trade of the match (p,ε): $S(p,\varepsilon) = f(p,\varepsilon) - v(\varepsilon) - v(p) + 2c$. Replacing $w(p,\varepsilon)$ in (2) leads to

$$v(\varepsilon) = \int_{p\in M(\varepsilon)} \left[\frac{1}{2}S(p,\varepsilon) + v(\varepsilon) - c\right] dG(p) + \int_{p\notin M(\varepsilon)} [v(\varepsilon) - c] dG(p).$$

Rearranging, this shows that, as in Atakan [2006], the expected surplus is constant:

$$c = \int_{M(\varepsilon)} \frac{1}{2}S(p,\varepsilon)dG(p). \tag{3}$$

Next, assign ε's matching set $M(\varepsilon)$ to worker $\tilde{\varepsilon}$. Then:

$$v(\tilde{\varepsilon}) \geq \int_{p\in M(\varepsilon)} w(p,\tilde{\varepsilon})dG(p) + \int_{p\notin M(\varepsilon)} [v(\tilde{\varepsilon}) - c] dG(p).$$

Similar steps as before lead to $c \geq \int_{M(\varepsilon)} \frac{1}{2}S(p,\tilde{\varepsilon})dG(p)$. Together with equation (3), this implies

$$\begin{aligned}
\int_{M(\varepsilon)} \frac{1}{2}S(p,\varepsilon)dG(p) &\geq \int_{M(\varepsilon)} \frac{1}{2}S(p,\tilde{\varepsilon})dG(p) \\
\int_{M(\varepsilon)} \left(f(p,\varepsilon) - v(\varepsilon) - v(p) + 2c\right) dG(p) &\geq \int_{M(\varepsilon)} \left(f(p,\tilde{\varepsilon}) - v(\tilde{\varepsilon}) - v(p) + 2c\right) dG(p) \\
[v(\tilde{\varepsilon}) - v(\varepsilon)] \int_{M(\varepsilon)} dG(p) &\geq \int_{M(\varepsilon)} \left[f(p,\tilde{\varepsilon}) - f(p,\varepsilon)\right] dG(p) \geq 0
\end{aligned}$$

This establishes that $v(\tilde{\varepsilon}) - v(\varepsilon) > 0$ for all (non-trivial) cases $\int_{M(\varepsilon)} dG(p) \neq 0$. Next, one can write wages as $w(\varepsilon,p) = \frac{1}{2}\left[f(p,\varepsilon) + v(\varepsilon) - v(p)\right]$. As both $f(p,\varepsilon)$ and $v(\varepsilon)$ are increasing, then so is $w(\varepsilon,p)$. ■

Payoffs $w(p,\varepsilon)$ or $\pi(p,\varepsilon)$ from *one* match provide an imperfect ranking of workers as firms. Payoffs depend on the partner type, which is not deterministic due to frictions in the matching process. Instead, the value of the firm $v(p)$ depends only on its type. Lemma 1 shows that $v(p)$ strictly increases with type. This value represents the expected profits (expected payoff $\pi(p,\varepsilon)$ minus the expected search costs). Firm-level profits are the sum of profits per match, for every worker employed by the firm.

2.1 Calibration of the Model

When bringing the model to the data we use the following stylized production function:

$$f(p,\varepsilon) = \alpha\left(p + \varepsilon + \psi p \varepsilon\right)$$

where ψ captures the degree of complementarity between p and ε, and α is a scale parameter. The model is calibrated in yearly basis and in thousands of euros. We assume that there are 100 types uniformly distributed in $[0,1]$. The stylized nature of this model leads us to interpret our findings as an illustration.

The equilibrium of the model is fully characterized by the vector of parameters $\Theta = (\delta, c, \alpha, \psi)$. We estimate the model matching a list of empirical moments which provides information on Θ. We briefly describe identification below.

Since job destruction is exogenously driven by a Poisson process described by δ, employment duration follows an exponential distribution with mean $1/\delta$. Therefore δ is identified by the average duration in employment observed in the data (102 months). We follow Eeckhout and Kircher [2011] to estimate c, exploiting information on the expected wage, minimum acceptance wages and the probability of employment.[4] α scales the production function to match mean wages observed in the data to mean wages produced in the model by $\widehat{\Theta}$.

The degree of complementarity ψ is estimated matching the strength of sorting generated by the model to the obtained one from the data. Given the rest of primitives, the strength of sorting can be used to identify how complementary are workers and firm types in production. The intuition behind this reasoning is that agents only wait for their preferred partners if the complementarity is strong enough to compensate for the waiting cost. So keeping everything else equal, as complementarity increases, the acceptance sets become narrower, which leads to a larger strength of sorting.

[4] To calculate the difference between the minimum wage and expected wage, we consider workers with at least 4 employers. The difference in yearly wages is $4,402$ euros.

The model fits duration of employment, mean-wages and the strength of sorting perfectly for $\widehat{\Theta} = (0.118, 3.408, 2.496, 51.74)$. As described in Eeckhout and Kircher [2011] there is another equilibrium with the same moments with negative ψ. However, we only search among positive values of ψ because we found in the data that the sign of sorting is positive. $\widehat{\Theta}$ is used in Section 5.2 to calculate the cost of frictions.

3. Additional Results on the Sign and Strength of Sorting

3.1 Heterogeneity in Unemployment Risk

Expected wages may not reflect worker type if workers are heterogeneous in search ability. Workers of similar type ε but different search ability possess different outside options. Therefore, they may get different wages, even when hired by the same firm right after unemployment. In this section we re-estimate γ comparing coworkers who are similar in terms of search performance.

In order to control for search performance we exploit the full length of the VWH. We focus on the sub-sample of 1995-2001 movers who were active in the labor market before 1995. We reconstruct their labor market history going back to 1975. We estimate equation (5) in the paper with controls for workers' past labor market histories. Controls include number of past employment spells, number of past unemployment spells, average duration of past employment spells, and average duration of past unemployment spells. After controlling for these additional sources of heterogeneity, the estimates of γ remain positive and significant for all measures of profits (see Table A5).

3.2 On-the-Job Search, Endogenous Search Intensity and Renegotiation

Under renegotiation and endogenous search intensity, wages are not necessarily increasing on the worker type, not even within the firm (see Bagger and Lentz [2016]). In Lentz [2010] and Bagger and Lentz [2016], when a worker meets a potential employer, the current and poaching firms compete *à la* Bertrand for the worker, and the most productive firm wins. Were poaching and current firms identical, the worker would extract the full

Table A5: Sign of sorting controlling for Heterogeneity in Unemployment Risk

	Economic profits		GOS		AP		ROE
	p-w	total	p-w	total	p-w	total	
γ	0.035	0.055	0.030	0.049	0.023	0.034	0.026
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
Avg. Past Tenure	-0.032	-0.019	-0.031	-0.022	-0.029	-0.023	-0.016
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Avg. Unempl. Duration	0.017	0.044	0.017	0.049	0.026	0.042	0.006
	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)	(0.008)	(0.009)
Past Empl. Spells	0.194	0.077	0.142	0.028	-0.042	-0.085	-0.095
	(0.074)	(0.082)	(0.075)	(0.084)	(0.075)	(0.071)	(0.075)
Past Unempl. Spells	-0.253	-0.223	-0.184	-0.141	0.023	0.035	0.065
	(0.085)	(0.093)	(0.085)	(0.096)	(0.086)	(0.081)	(0.085)

Notes: This table presents estimates for the sign of sorting controlling for worker specific search ability. Controls for gender, age, age squared, migration status, tenure, tenure squared, year, and occupation are included in all regressions. Avg. Past Tenure is the average tenure in past employment spells ($\times$ 0.01). Avg. Unempl. Duration is average duration in past unemployment spells ($\times$ 0.01). Past Empl. Spells is the number of past employment spells ($\times$ 0.01). Past Unempl. Spells is the number of past unemployment spells ($\times$ 0.01). Sub-sample of 1995-2001 movers active in the labor market prior to 1995. Standard errors in parentheses. Number of observations is $89,007$.

rent, obtaining a wage equal to the productivity of the match. We can use this last implication to rank workers within the firm by their types. For job-to-job transitions between firms of similar quality, the wage proxies the productivity of the match. Therefore, we focus on the subsample of workers whose previous firm (not mediated by unemployment) is similar to the current firm. Of those, we pick coworkers who transition to a *third* firm through unemployment.

This exercise is demanding in terms of data. We select workers who transition at least three times. To illustrate this, denote some worker's employers chronologically as 1, 2 and 3. We order workers by their wages in firm 2. To do so we identify workers who move on-the-job to firm 2, coming from firms 1 similar to 2. Next, we order employers by their next firm 3. This sample trimming significantly reduces the number of valid observations to $16,870$.

We regress equation (6) in this subsample. We present the results in Table A6. Transitions between two similar firms are defined as transitions without interim unemployment spell where the difference in rankings between previous and next firm is less than 0.10 in Columns (1) and (3) or below 0.05 in Columns (2) and (4). In all cases we find that better workers are more likely to move to better firms.

Table A6: Sign of Sorting with On-the-job Search and Renegotiation

	(1)	(2)	(3)	(4)
	Aggregated Profits		Profits per Worker	
Difference in Ranking	< 0.10	< 0.05	< 0.10	< 0.05
Wage$\times\mathbb{1}$(Similar Firm)	0.037	0.028	0.037	0.029
	(0.013)	(0.011)	(0.012)	(0.012)
Wage$\times[1-\mathbb{1}$(Similar Firm)]	0.041	0.031	0.042	0.032
	(0.013)	(0.011)	(0.012)	(0.012)

Notes: This table presents estimates for the sign of sorting allowing for on-the-job search and renegotiation. Number of observations is $16,870$. Each column represents a single regression. In columns (1) and (2) firms are ordered in terms of economic profit, whereas in columns (3) and (4) they are ordered in terms of profit per worker. We include controls for gender, age, age squared, migration status, tenure, tenure squared, year, and occupation in all regressions. Firms fixed effects are included in every specification. All specifications consider workers who switch at least three times. $\mathbb{1}$(Similar firm) is an indicator that takes a value of one if the worker comes from a firm similar to the firm where the wage is observed. Standard errors in parentheses.

Table A7: Estimates of γ (Sign of Sorting) with Detailed Occupation Controls

Economic profits		GOS		AP		ROE
per worker	total	per worker	total	per worker	total	
0.014	0.021	0.010	0.020	0.011	0.015	0.015
(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.008)	(0.006)

Notes: This table presents estimates for the sign of sorting including occupation controls. The dependent variable is the type of the previous firm before an unemployment spell. Each column represents a single regression. Controls for gender, age, age squared, migration status, tenure, tenure squared, year and occupation are included in all regressions. Firm×occupation fixed effects are included in every specification. Average number of workers in each firm×occupation cluster is 26.16. Standard errors in parentheses. Number of observations is $114,871$.

3.3 Detailed Occupation Categories

Ideally, our test for the sign of sorting should be based on transitions of workers with identical jobs in the new firm. We include controls for broad occupation categories (blue collar, white collar and managers) in our results on Table 2 in the paper. However, one concern with these estimates is that our occupational controls may be too broad. Detailed occupation classification such as ISCO is missing in our data set. In spite of this, we use detailed information on "job ladders" (*livelli di inquadramento*) within each industry-wide collective national contract to construct a more precise classification of jobs within firms.[5] Although job ladders cannot be compared across different firms (they may be applying different national contracts), ladder codes can be used to identify workers doing similar jobs *within the firm*. Hence, we construct dummies identifying workers under the same national contract and job ladder code. We run our test including these more refined controls for occupation, instead of the firm dummies as in the previous tables. Table A7 shows the estimates of γ in specifications that absorb a large number of effects by firm×(national contract)×(job ladder) for different measures of profits. The results corroborate the presence of positive sorting: γ is positive and significant in every specification.

[5] Each national contract specifies its own job-ladder and stipulates minimum wages for each of level of the ladder (typically, there are 7 to 9 levels in each job ladder).

Table A8: Alternative specification for testing the sign of sorting.

	Profits per worker	Profits in level
γ^*	0.019	0.029
	(0.003)	(0.004)

Notes: This table presents estimates for the sign of sorting using equation (4). Profits are measured in terms of economic profits. The sample is the one from column (3) in Table 5 in the paper. Each regression includes firm fixed effects. The list of control variables is the same as in Table 2. Standard errors clustered by firm in parenthesis.

3.4 Sign of sorting: An Alternative Specification

In this section we report the results of an alternative specification to identify the sign of sorting. This alternative test relies on a specification where wages on the current job are a function of the firm type of the previous job. The intuition behind this strategy is the following: consider two coworkers, if the one who makes more money (likely a better worker) comes more often from a better firm, better workers are more often matched to better firms. We estimate the following regression:

$$w_{ij} = \zeta + \gamma^* p_{ij}^{PREV} + x_{ij}'\beta_j + \widetilde{v}_{ij}, \tag{4}$$

Estimates of γ^* are presented in Table A8, using our main analysis sample. We also find evidence of positive sorting.

4. Methodologies based on Abowd et al. [1999]. Details

In this section we report results using the strategy presented in Abowd et al. [1999]. We first estimate the following wage equation:

$$w_{i,j,t} = x_{i,j,t}'\beta + \vartheta_i + \xi_j + z_{i,j,t}, \tag{5}$$

where $x_{i,j,t}$ are observable and time-varying characteristics of the worker and the firm, ϑ_i is worker i fixed effect and ξ_j is firm j fixed effect. The dependent variable is the worker's daily wage and the time-varying controls include a quadratic in the worker's age, a quadratic in the worker's tenure with the current employer, indicators for white-collar and managerial occupations, indicators for five firm-size classes and year effects. In Abowd et al.'s methodology only workers and firm fixed effects within connected groups are identified.[6] In our analysis we have considered only the largest connected group, which contains more than 95% of the sample observations. We present the results in in Table A9. As reported by many other replications of the classical result by Abowd et al., we too find evidence of a small negative correlation (-0.02) between the worker fixed effects and the firm fixed effects. Moreover, this correlation is statistically significant in our data set.

Table A9: AKM Approach: OLS Estimates of Equation (5)

$y = \log(wages)$	Coefficient	Std-Dev.
Age	0.0486	(0.00018)
Age^2	-0.0004	(2.34×10^{-6})
Tenure	0.0006	(0.000013)
Tenure^2	-1.43×10^{-6}	(5.90×10^{-8})
White-Collar	0.0510	(0.000734)
Manager	0.2879	(0.003016)
Std. dev. of person effects ϑ_i (across person year obs) = 0.320		
Std. dev. of firm effects ξ_j (across person year obs) = 0.151		
Correlation$(\xi_j, \vartheta_i) = -0.0216$ with p-value < 0.0001		

Notes: This table presents OLS estimates from the Abowd et al. [1999]'s approach. Number of observations is $2,672,812$. Number of workers is $778,388$. Number of firms is $11,985$. Year dummies and dummies for firm size (5 categories) included. The sample covers the period 1995-2001.

Abowd et al.'s strategy may not be informative of sorting by underlying types because of several reasons. First, mean wages provide a noisy ranking of workers. If i.i.d., this noise biases the correlation toward zero. However, as described in Section 3 of the paper, the noise in the estimate of worker i's mean wages is potentially correlated with worker i's employers' types. If so, the direction of the bias is undetermined.

Second, there is also noise in the ranking of firms. The estimated covariance can be

[6] Connected workers and firms are those linked by workers' mobility. This restriction implies the loss of a small proportion of observations and is usually applied to simplify estimations. See Abowd, Creecy, and Kramarz [2002] for a discussion.

biased if the noise in the estimated ranking of workers and the noise in the estimated ranking of firms are correlated. Andrews, Gill, Schank, and Upward [2008] and Abowd, Kramarz, Lengerman, and Perez-Duarte [2004] find that, although the bias can be considerable, it is not sufficiently large to remove the negative correlation in data sets from Germany, France and the United States. Andrews, Gill, Schank, and Upward [2012] instead report that the estimated correlation becomes positive for larger samples with more inter-firm mobility. Our data are, however, already based on the population of workers and firms in our regional labor market.

Third, as pointed out by Gautier and Teulings [2006], Eeckhout and Kircher [2011] and Lopes de Melo [Forthcoming], firm fixed effects estimated from wage equations do not necessarily reflect the firms' underlying types. Whenever wages are non-monotone in firm type, the linear model in Abowd et al. [1999] is fundamentally misspecified. Gautier and Teulings [2006] present evidence of wages non-monotonicity in the firm type for five European countries and the United States.

References

ABOWD, J. M., R. H. CREECY, AND F. KRAMARZ (2002): "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau.

ABOWD, J. M., F. KRAMARZ, P. LENGERMAN, AND S. PEREZ-DUARTE (2004): "Are Good Workers Employed by Good Firms? A Test of a Simple Assortative Matching Model for France and the United States," Working Paper.

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–334.

ANDREWS, M., L. GILL, T. SCHANK, AND R. UPWARD (2012): "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias," *Economics Letters*, 117, 824 – 827.

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.

ARELLANO, M. (2003): "Panel data econometrics," *OUP Catalogue*.

ATAKAN, A. E. (2006): "Assortative Matching with Explicit Search Costs," *Econometrica*, 74, 667–680.

BAGGER, J. AND R. LENTZ (2016): "An Empirical Model of Wage Dispersion with Sorting," Working Paper.

CINGANO, F. AND A. ROSOLIA (2012): "People I know: job search and social networks," *Journal of Labor Economics*, 30, 291–332.

EECKHOUT, J. AND P. KIRCHER (2011): "Identifying sorting - in theory," *The Review of Economic Studies*, 78, 872–906.

GAUTIER, P. A. AND C. N. TEULINGS (2006): "How large are search frictions?" *Journal of the European Economic Association*, 4, 1193–1225.

LENTZ, R. (2010): "Sorting by search intensity," *Journal of Economic Theory*, 145, 1436 – 1452.

LOPES DE MELO, R. (Forthcoming): "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy*.