Paired movers

Isaac Sorkin Lucas Warwar *

March 2026

Abstract

Consider a pair of workers who both move from firm A to firm B. In a model with additively separable firm and worker effects, the pay gap between them at the origin firm should perfectly predict the gap at the destination firm (a forecast coefficient of 1). Using data from Brazil and Italy, we find forecast coefficients of 0.73 and 0.78, rejecting additive separability and indicating the presence of match effects. Under exogenous mobility, these estimates imply that the variance of match effects is about 40% of the variance of firm effects. In a model with endogenous mobility, these estimates imply a larger role for match effects. In simulated data, we show that standard specification tests are often biased or underpowered, and that extrapolating earnings patterns from observed to unobserved job offers can be misleading when match effects are important.

*Sorkin: Stanford University and NBER, sorkin@stanford.edu. Warwar: Stanford University, warwar@stanford.edu. Thanks to numerous seminar and conference participants, Jediphi Cabal, Sydnee Caldwell, Jon Cohen, Michael Dinerstein, Liran Einav, Paul Goldsmith-Pinkham, Hamish Low and Chris Moser for helpful comments and conversations. We thank Diogo Britto for assistance with the Brazilian data. We thank Katarina Borovickova for generously sharing code. The data used in this study were generously provided by the Fondazione Rodolfo De Benedetti and originally developed by the Economics Department of the Università Ca Foscari Venezia under the supervision of Giuseppe Tattara. The production of this article was supported by research funding from Stanford University.

A large literature is based on the Abowd, Kramarz, and Margolis (1999) (AKM) wage equation where the (log) wage can be written as the sum of a person effect, a firm effect, and an independent error term (e.g., Card, Heining, and Kline (2013), Card, Cardoso, and Kline (2016), Sorkin (2018), and Gerard et al. (2021)). A causal interpretation of the firm effect is justified either by the assumption of exogenous mobility or of additive separability.

This equation is controversial because it conflicts with many canonical models of wage determination and mobility (see, e.g., Gibbons et al. (2005), Eeckhout and Kircher (2011), and Borovičková and Shimer (2024)). Nonetheless, three pieces of evidence are supportive of this specification. First, Card, Heining, and Kline (2013) introduce what we call the event study test for exogenous mobility, and show that for movers across firms the change in the firm effects does a good job of predicting changes in earnings. Second, Card, Heining, and Kline (2013) look at the change in adjusted-R^2 from an additively separable model (AKM) to a non-additively separable model and find that the change is small, suggesting a minimal role for match effects. Finally, Bonhomme, Lamadon, and Manresa (2019) develop a way of estimating interactions in the wage equation and find minimal deviations from additive separability.

In this paper, we revisit this evidence by exploring a novel feature of the data: paired movers. We consider a pair of workers (*paired movers*) who both move from firm A to B. Additive separability implies that on average the gap between their earnings should depend only on the difference in their person effects, which should be stable across firms. We first test this implication and document meaningful deviations from additive separability. We then use paired movers to develop an estimator for the variance of match effects under exogenous mobility. Finally, we estimate a model of endogenous mobility and use a combination of the model and the paired movers to revisit each of the three pieces of evidence we described in the last paragraph. In each case, we find that these are less compelling than previously thought.

We begin by explaining how additive separability is stronger than exogenous mobility. Exogenous mobility says that workers can move on the basis of the firm effects, the worker effects, and non-wage related factors. Workers cannot move on the basis of match effects or the error term. However, there can still be match effects. Additive separability says that there are not match effects in the wage equation. To see the distinction, suppose that there are match effects in the wage equation but there is random assignment of workers to firms. In this case, we can consistently estimate the average treatment effects of firms because the match effects average out. So exogenous mobility is satisfied. But additive separability is not satisfied because there are match effects. Thus, rejecting additive separability does not imply that exogenous mobility fails. Nonetheless, we present additional evidence that challenges the exogenous mobility assumption.

We then introduce the paired movers and use them test an implication of additive separability. When additive separability holds, the regression of the gap between the paired movers at firm B on the gap at firm A is just a regression of the difference in worker effects on itself and so the forecast coefficient should be 1.

Assuming exogenous mobility, we develop an estimator for the variance of match effects. The estimator combines the forecast coefficient with the variance of the earnings gap between the paired movers.

We use administrative data from two countries: Italy and Brazil. For the Italian data, we use the Veneto Worker History data. For the Brazilian data, we use the RAIS data. Because the Brazilian dataset is much larger (and less noisy), we focus on the results from Brazil.

We show that the main data generating process for the paired movers is simply chance. The most important difference between paired movers and all movers is that paired movers start from—and end up at—much larger firms. Given mobility patterns, the number of paired movers is close to what we would expect by chance. While the paired movers are a narrow subset of the data (5% of movers), it is roughly as narrow as we would expect based on mobility patterns.

In our main results, we document substantial deviations from additive separability. To develop a reasonable benchmark, we show that among stayers at the same firm, the forecast coefficient is 0.96. Among all firm switchers, the coefficient is 0.73. Under the assumption of exogenous mobility, our estimator implies that the variance of match effects is about 40% of the Kline, Saggio, and Sølvsten (2020) (KSS)-corrected variance of firm effects in the full sample.

We then develop subgroup analyses which suggest violations of exogenous mobility. We find lower coefficients for moves across contexts than within contexts, where we define context based either on sector cluster. Moreover, workers are disproportionately likely to remain in the same context, rather than switch. This pattern occurs in models of endogenous mobility where the presence of match effects implies that much mobility is not realized.

To develop an estimate of the variance of match effects taking into account endogenous mobility, we estimate a random job search model where workers are heterogeneous in productivity and move on the basis of both a firm effect and a match effect. Quantitatively, the model finds that the variance of match effects in the steady state distribution is about 20% larger than that implied by our exogenous mobility estimator, or about half of the variance of firm effects in the full sample.

In the model, the presence of match effects means that the average wage changes experienced by workers who accept offers are a poor predictor of the wage changes of other potential movers. Specifically, we forecast the earnings changes of the rejected offers using the earnings changes of the accepted offers and find that there is a large negative intercept and a slope of around 0.8.

Turning to the first piece of evidence supporting the AKM specification, this model generates an event study coefficient of 0.99. Thus, the data from the model passes the event study test, which is surprising given that it features endogenous mobility. Splitting the data, we find a coefficient of 0.57 on employer to employer (EE) transitions (about 40% of the data) and 0.99 on employer to unemployment to employment (EUE) transitions. The pooled data passes the event study test because the EE moves see bigger increases in firms effects and in earnings than the EUE moves, which is a generic feature of endogenous mobility. As a result, the sample means differ, a mechanism analogous to Simpson's paradox. This mechanism is powerful: even when the share of EE moves is over 90% and the event study coefficients in the EE and EUE subsamples are both below 0.8, the event study coefficient is still above 0.9. This between-sample pooling mechanism offers a new explanation for why the event study can be low-powered.

We show evidence of this between-sample effect in the data: in the Brazilian data, the event study coefficient is 1.07. But when we split moves by whether there was an increase or decrease in the firm effect, we find coefficients of 1.25 and 1.47, which are less supportive of the exogenous mobility assumption. Thus, the event study test is low-powered against alternative hypotheses that include an important role for endogenous mobility.

Turning to the second piece of evidence, we document two sources of downward bias in the comparison of the adjusted-R^2 from AKM and an uninteracted model: endogenous mobility and short panels. With endogenous mobility and short panels, this estimator is downward biased by almost an order of magnitude. Thus, the small implied variance of match effects from this estimator is consistent with a much larger role for match effects than previously thought.

Turning to the third piece of evidence, we implement the static Bonhomme, Lamadon, and Manresa (2019) estimator in our data. Typically, the BLM estimator finds some deviations from additive separability and so we want to translate their deviations into our metrics. We simulate data from their estimated model and then run the paired mover test in the simulated data. The BLM estimator is unable to accurately recover the paired mover coefficient. Monte Carlo evidence shows that this failure is because BLM struggles to reliably estimate mobility patterns. We thus conclude that our test points to a new feature of the data that has been missed by previous work.

Literature: The most closely related paper is Kantenga and Law (2016). Section 6 develops a statistical test for additive separability based on the same feature of the data that this paper exploits: two workers observed at the same two firms (albeit, not restricted to the same years) should have constant earnings gaps up to concerns about measurement error. This paper develops a more transparent approach to looking at this implication. We also use this feature of the data to develop an estimator for the variance of match effects.

Woodcock (2015) develops estimators of match effects in the context of the AKM model. One is the orthogonal match effects estimator which defines the match effect as the mean residual of the AKM model within firm-worker match. As Kline (2024) emphasizes, the estimation of many parameters in the AKM model relies on indirect contrasts where it is less transparent what feature of the data is driving results. The present paper builds from the minimal feature of the data that identifies match effects.

Kline (2024) uses Veneto data and Table 1 looks at the edges where there are two or more movers, which is the same feature of the data that this paper relies on. It shows that the AKM model captures 84% of the variance of edge effects generated by looking directly at the movers. This paper looks at different feature of these edges: whether the earnings gaps between workers are stable with the moves. In this sense, the exercise in this paper is reminiscent of a pre-processing step in Hagedorn, Law, and Manovskii (2017) where they find that the ranking of workers across firms is not stable and they develop an approach to average this out (see their section 4.2).

A number of "AKM" papers present evidence that is consistent with match effects.

Most directly, Beauregard et al. (2025) allows the worker's fixed effect to depend on whether they are matched to a union or non-union firm. A number of papers ask whether the change in the firm effects can explain the change in earnings for a select subset of workers. Sometimes firm effects do a poor job of explaining these changes. For example, Lachowska, Mas, and Woodbury (2020) find that the change in firm effects does not on average explain the earnings losses of displaced workers. Relative to this literature which relies on having a specific shock to mobility or a particular model of match effects, this paper presents an omnibus approach to testing for and quantifying the role of match effects.

Conceptually, this paper is related to Borovičková and Shimer (2024). They develop a model with endogenous mobility because of selective hiring. The paper develops analytical and numerical evidence that the model passes the event study test of Card, Heining, and Kline (2013), despite an economic structure that violates exogenous mobility. We show that the reason their model passes the event study test is the between-sample effect that we emphasized above. More broadly, while the point of Borovičková and Shimer (2024) is to show analytically that it is possible for an economic structure featuring match effects to deliver additively separable wage equations, our goal is to directly test this functional form restriction.

1 Additive separability versus exogenous mobility

The basis of much empirical work in labor (and applied) economics is the following two way fixed effects specification (Abowd, Kramarz, and Margolis (1999)):

$$y_{it} = \alpha_i + \psi_{\mathbf{j}(i,t)} + \beta X_{it} + \epsilon_{it}, \tag{1}$$

where y is log earnings, i is an individual, $\mathbf{j}(i, t)$ is the firm that employs worker i at time t, α is a worker fixed effect, ψ is a firm effect, X_{it} is a set of time-varying worker covariates, and ϵ is an error term. This equation looks like it imposes additive separability in that there is no term for interactions between firms (j) and workers (i). The conventional assumption, however, deviates from additive separability.

For expositional purposes, we depart from the conventional equation and explicitly write the wage equation to include match effects:

$$y_{it} = \alpha_i + \psi_{\mathbf{j}(i,t)} + \beta X_{it} + \mu_{ij} + \tilde{\epsilon}_{it}, \tag{2}$$

where $\mu_{ij} \equiv \mathbb{E}[\epsilon_{it}|i,j]$ so that μ_{ij} is defined as the mean of the error term over all draws of the error term, not just the periods in which i and j actually match and $\tilde{\epsilon}_{it} \equiv \epsilon_{it} - \mu_{ij}$. This definition of a match effect as a permanent component of the match parallels the way person and firm effects are defined, and parallels the definition in equation (3) in Woodcock (2015). Note that $\tilde{\epsilon}_{it}$ is mean zero, while the potential ϵ_{it} for a particular worker i at firm j may not be. We refer to the μ_{ij} as match effects in what follows, though the μ_{ij} are a statistical—and not an economic—object.

Once we decompose the error term into a permanent component (the match effects) and an idiosyncratic component, exogenous mobility says that on average workers do not select firms on the basis of either component (this assumption uses notation from Kline (2024)):

Assumption 1. *Assume that:*

$$\mathbb{E}[\tilde{\epsilon}_{it}|\mathbf{j}(i,s), X_{is} = x] = 0, \tag{3}$$

and

$$\mathbb{E}[\mu_{ij}|\mathbf{j}(i,s), X_{is} = x] = 0, \tag{4}$$

where $\mathbf{j}(i,s)$ *is the employer of worker* i *in time period* s*, for all workers* $i \in \{1, ..., N\} \equiv [N]$*, all time periods,* $(s,t) \in [T]^2$ *(where* $t \in \{1, ..., T\} \equiv [T]$*), and all possible firm assignments* $j \in [J]$ *(where* $j \in \{1, ..., J\} \equiv [J]$*).*

This assumption still allows workers to select firms on the basis of the ψ, and thus is consistent with arbitrary patterns of sorting because workers with different α can select firms (or firms can select workers) differently. This assumption also allows the μ_{ij} to be nonzero and thus allows treatment effects of firms to be heterogeneous.

Additive separability pairs the same assumption on the idiosyncratic component of the error term with a stronger statement about μ_{ij}:

Assumption 2. *Assume that:*

$$\mathbb{E}[\tilde{\epsilon}_{it}|\mathbf{j}(i,s), X_{is} = x] = 0, \tag{5}$$

and

$$\mu_{ij} = 0 \ \forall \ i, j, \tag{6}$$

where $\mathbf{j}(i,s)$ *is the employer of worker* i *in time period* s*, for all workers* $i \in \{1, ..., N\} \equiv [N]$*, all time periods,* $(s,t) \in [T]^2$ *(where* $t \in \{1, ..., T\} \equiv [T]$*), and all possible firm assignments* $j \in [J]$ *(where* $j \in \{1, ..., J\} \equiv [J]$*).*

Additive separability says that the permanent component of the match effects are all zero, whereas exogenous mobility says that they are mean zero.

Exogenous mobility is an assumption which allows the researcher to estimate the average treatment effect of the firm, while additive separability is an assumption which implies that the treatment effect of the firm is constant. Thus, exogenous mobility can hold without additive separability holding and specification tests for exogenous mobility do not directly test additive separability.

2 Paired movers

In this section, we introduce the paired movers and show that under exogenous mobility we can use the paired movers to estimate the variance of match effects.

2.1 The paired movers coefficient

We consider workers i and i' who both work at firm A and B. For expositional simplicity, we replace the time subscript with the firm subscript and drop the covariates, but in implementation we keep track of both time and covariates. Then the difference between worker i and i''s earnings at firm A is given by:

$$y_{iA} - y_{i'A} = (\alpha_i - \alpha_{i'}) + (\mu_{iA} - \mu_{i'A}) + (\tilde{\epsilon}_{iA} - \tilde{\epsilon}_{i'A}). \tag{7}$$

By differencing across workers within firm, the firm effect drops out. Additive separability implies that $\mathbb{E}[\tilde{\epsilon}_{iA} - \tilde{\epsilon}_{i'A} | \mathbf{j}(i,s) = A, \mathbf{j}(i',s) = A] = 0$ and $\mu_{iA} = \mu_{i'A} = 0$. Hence, $\mathbb{E}[y_{iA} - y_{i'A}] = (\alpha_i - \alpha_{i'})$. Analogous expressions apply at firm B.

To construct the coefficient, we consider whether the earnings *gap* at firm A predicts the earnings *gap* at firm B:

$$(y_{iB} - y_{i'B}) = \beta_0 + \beta_1 (y_{iA} - y_{i'A}) + \varepsilon_{i,i',A,B}. \tag{8}$$

Under the assumption of additive separability and assuming no error term, the probability limit of β_1 is 1 because this regression is effectively a regression of $\alpha_i - \alpha_{i'}$ on itself. Thus, deviations from 1 are indicative of deviations from additive separability which implies the presence of match effects.

2.2 Interpretation of paired movers coefficient

To interpret deviations from 1 in the forecast coefficient, we compute its probability limit. For clarity, it is helpful to define $\mathcal{P}$ to be the set of paired movers. Then the probability limit is given by:

$$\begin{aligned}
&\operatorname{plim} \hat{\beta}_1 = \\
&\frac{\operatorname{Var}(\alpha_i - \alpha_{i'} \mid (i,i') \in \mathcal{P}) + \operatorname{Cov}(\mu_{iA} - \mu_{i'A}, \mu_{iB} - \mu_{i'B} \mid (i,i') \in \mathcal{P})}{\operatorname{Var}(\alpha_i - \alpha_{i'} \mid (i,i') \in \mathcal{P}) + 2\operatorname{Cov}(\alpha_i - \alpha_{i'}, \mu_{iA} - \mu_{i'A} \mid (i,i') \in \mathcal{P}) + \operatorname{Var}(\mu_{iA} - \mu_{i'A} \mid (i,i') \in \mathcal{P})} \\
&+ \frac{\operatorname{Cov}(\alpha_i - \alpha_{i'}, \mu_{iB} - \mu_{i'B} \mid (i,i') \in \mathcal{P}) + \operatorname{Cov}(\alpha_i - \alpha_{i'}, \mu_{iA} - \mu_{i'A} \mid (i,i') \in \mathcal{P})}{\operatorname{Var}(\alpha_i - \alpha_{i'} \mid (i,i') \in \mathcal{P}) + 2\operatorname{Cov}(\alpha_i - \alpha_{i'}, \mu_{iA} - \mu_{i'A} \mid (i,i') \in \mathcal{P})) + \operatorname{Var}(\mu_{iA} - \mu_{i'A} \mid (i,i') \in \mathcal{P})}.
\end{aligned} \tag{9}$$

The covariance terms reflect different models of endogenous mobility. In a model of selective hiring where firms differ in their standards (which also arises in models of complementarities in the production function), firms would trade off α and μ: the total quality of the worker has to be high enough to justify the hire. In such a model, the $\operatorname{Cov}(\alpha_i - \alpha_{i'}, \mu_{iB} - \mu_{i'B} \mid (i,i') \in \mathcal{P})$ term would be negative. In a model of worker search, workers would trade off the μ and ψ in their acceptance decisions. Thus, the $\operatorname{Cov}(\mu_{iA} - \mu_{i'A}, \mu_{iB} - \mu_{i'B} \mid (i,i') \in \mathcal{P})$ would be positive.

2.3 A match effects estimator under exogenous mobility

Under exogenous mobility, the only source of selection is the across-firm sorting of workers. To capture the within-firm distribution of workers, define $\bar{\alpha}_j$ to be the mean of the person effect of individuals at firm j who are in the paired mover sample ($i \mid \exists i'$ s.t. $(i,i') \in \mathcal{P}$; for compactness, $\mathcal{P}$). Define $\tilde{\alpha}_{ij} \equiv \alpha_i - \bar{\alpha}_j$ to be the demeaned person effect, where the demeaning is *within* firm.

In Appendix A, we show that under exogenous mobility equation (9) simplifies to

$$\text{plim}\ \hat{\beta}_1 = \frac{\text{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P})}{\text{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P}) + \text{Var}(\tilde{\mu}_{ij} \mid \mathcal{P})}, \tag{10}$$

so the coefficient only depends on the variances of person and match effects.

The above expression gives us one moment and two unknowns. For a second moment, consider the gap between the paired movers at the origin firm j:

$$Z_{ii'j} \equiv y_i - y_{i'} = \alpha_i + \mu_{ij} + \tilde{\epsilon}_{ij} - \alpha_{i'} - \mu_{i'j} - \tilde{\epsilon}_{i'j},$$

where the firm effects difference out. Using identical reasoning to above, we have:

$$\text{Var}(Z) = 2\,\text{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P}) + 2\,\text{Var}(\tilde{\mu}_{ij} \mid \mathcal{P}) + 2\,\text{Var}(\tilde{\epsilon}_{ij} \mid \mathcal{P}), \tag{11}$$

so that the variance of the gap depends on the variances of the person and match effects, as well as the variance of noise.

We can combine equations (10) and (11) to solve for the average within-firm variance of match effects in closed form:

$$\text{Var}(\tilde{\mu}_{ij} \mid \mathcal{P}) = (1 - \hat{\beta}_1)\,\text{Var}(Z)\frac{\Gamma_Z}{2}, \tag{12}$$

where Γ_Z is the signal to noise ratio in Z. These equations show that the paired mover coefficient maps linearly into the variance of the match effects. A smaller paired mover coefficient, $\hat{\beta}_1$, implies a larger variance of match effects. And, similarly, a larger variance of the gap, $\text{Var}(Z)$, implies a larger variance of match effects. In Appendix A, we present an estimator for Γ_Z under the assumption that the noise is independent and identically distributed over time.

3 Data

3.1 Brazilian data

We use *Relação Anual de Informações Sociais* (RAIS), which is a matched employer-employee dataset. Each observation is a job spell and an employer is an establishment within a firm. There is data from 2006 to 2022. We include only the worker's dominant job, which is the job with the highest earnings in each year. We drop workers on temporary contracts, public servants, and those with zero earnings. We include workers aged 18 to 64. This is the sample of *all workers*.

We define a mover as a worker who changes employers between two consecutive years and separated from the previous employer. We exclude workers who change their dominant job but remain employed at the original firm. We also exclude workers who transfer within establishments of the same firm. This is the sample of *all movers*.

To construct the sample of *paired movers*, we make a few restrictions. A move is defined by origin firm, destination firm, and year. First, we focus on paired movers where the workers are present for two years in the origin firm. This restriction allows us to instrument the earnings at the origin firm. Second, we require that no more than 30% of the workers in the origin firm go to the destination firm, and that no more than 30% of employment in the destination firm came from the origin firm (Benedetto et al. (2007)). In many cases, there are more than two workers who leave a firm in a particular year. In this case, we randomly match workers and keep only one pair per worker per move.

We also construct a sample of *paired stayers*, where the origin and destination firm is the same. For computational reasons, we take a random sample of 10%.

We construct three other samples to estimate AKM models. We start from the sample of *all workers* defined above in 2015. Due to the large size of the data, we take a random sample of 25% of the workers and construct a panel of these workers from 2012 to 2017. Building on this first sample, the second sample includes only workers who change employers at least once between 2012 and 2017—that is, we drop the stayers. The third sample includes only workers in the second sample who are in the paired movers sample (and her paired move is between 2012 and 2017).

Earnings are the average log hourly wage in the year. Job contracts in Brazil are defined by monthly wages and weekly working hours. We compute hourly wages as the average monthly wage divided by an estimate of monthly hours (weekly hours $\times$ 30/7) (Gerard et al., 2021) and deflate values to 2019 BRL. Firm size is the number of workers employed on December 31st of each year.

Age is measured on December 31st of the year. Years of education is the highest level of education completed using the *Instituto Brasileiro de Geografia e Estatística* (IBGE) classification. Tenure is measured in months using the date of hire and the date of separation (or Dec. 31st). Occupation is classified using three-digit

Classificação Brasileira de Ocupações (CBO) codes. Sectors are defined using two-digit *Classificação Nacional de Atividades Econômicas* (CNAE) codes. We observe the establishment's municipality and use IBGE's *regiões geográficas imediatas*, which are similar to commuting zones in the U.S., as geographic units.

For most of the analysis, we work with residualized earnings. We construct residualized earnings in two steps. First, we estimate the role of covariates using only within-firm-worker match variation in wages. Our baseline set of covariates is a third degree polynomial in age, interacted with gender and four education categories (no primary education, primary education, high school, and some college or more). Following Card et al. (2018), we normalize age around 40 years old and omit the linear term. Second, we construct residualized earnings by subtracting off the estimated effect of covariates. We use residualized earnings almost everywhere. In the AKM analysis, we show how the conventional AKM regression compares to the results using residualized earnings.

We randomly order paired movers so that sometimes the earnings gap is positive and sometimes negative. To explain why we randomly sort workers, Appendix B shows that if two workers draw earnings from distributions with a different mean (as individual effects imply), then sorting workers to ensure that the gaps are the same sign creates a mechanical upward bias.

3.2 Veneto data

We use cleaning code from Kline, Saggio, and Sølvsten (2020) and so our description follows their computational appendix. We use data from 1985 to 2001. We use workers aged 18 to 64. The earnings concept is daily wages. We use the worker's dominant (highest-earnings) job. We exclude workers with implausibly low wages (below 5 euros or zero days worked), implausibly large wage changes (more than doubling), or those employed in the public sector, having more than ten jobs, or missing gender. The industry codes are five-digit ATECO 91 codes. The employers in the Veneto data are firms. We construct paired movers in a way identical to in the Brazilian data.

4 Results

In this section, we first discuss descriptive statistics. We then present the main results documenting meaningful deviations from additive separability and present extensive robustness. We show that under exogenous mobility the estimate of the variance of match effects is substantial. Finally, we show that the deviations from additive separability are stronger in moves across contexts.

4.1 Descriptive statistics

While canonical estimates of firm effects extrapolate from the sample of all movers (in the connected set) to all workers, this paper extrapolates from paired movers. Table 1 shows summary statistics among all workers, all movers and paired movers. In general, movers are negatively selected relative to all workers in terms of tenure and earnings. In contrast, paired movers are either less negatively or positively selected. Relative to all movers, paired movers come from much larger employers and higher-paying employers. Movers are younger and more male than all workers.

The paired movers come from larger firms and go to larger firms, which suggests that the dominant form of selection into paired mobility is simply chance. To assess the chance hypothesis, we take the origin firms and compute the empirical distribution of the destination firms of movers while not including the paired movers. Using these empirical probabilities, we then simulate mobility using the number of movers we see and compute the number of paired movers. Table 2 shows that this simple logic explains over 40% of the paired movers. This estimate is downward biased because we omit the paired movers. To generate an upward-biased estimate, we include the paired movers which Appendix Table A1 shows can explain about 65% of the paired movers. Appendix Table A2 shows that over half of all paired movers are the only paired movers along the origin-destination-year path, and so it is hard to generate all paired movers by chance.

As another way of comparing paired movers to all movers, Table 3 shows the AKM decomposition in four samples. Columns (1) and (2) show the decomposition in the sample of all workers, where the columns differ by whether we use raw earnings or the residualized earnings, and show minimal differences. Column (3) reports the decomposition among all movers, and column (4) reports results for the paired movers. The baseline finding is that the adjusted-R^2—and thus the fit of the AKM model—

Table 1: Summary statistics

	All workers (1)	All movers (2)	Paired movers (3)
A. Brazil			
Avg. age	34.09	30.90	32.37
Share female	0.40	0.35	0.31
Years of schooling	11.30	11.10	10.90
Avg. Tenure (months)	46.97	23.27	38.13
Median Tenure (months)	24.74	13.52	26.83
Avg. log earnings	2.34	2.19	2.39
# of Worker-Years	555,346,010	79,856,352	3,962,382
Avg. (origin) firm size	14.08	27.21	78.83
Avg. (origin) firm pay	2.41	2.45	2.47
# of (origin) Firm-Years	46,405,377	18,220,954	883,078
B. Veneto			
Avg. age	34.17	30.14	35.39
Share female	0.37	0.37	0.35
Avg. log earnings	10.10	9.64	10.20
# of Worker-years	15,158,383	1,897,250	17,108
Avg. (origin) firm size	8.35	15.79	221.66
Avg. (origin) firm pay	10.10	10.11	10.17
# of (origin) Firm-years	1,815,553	676,890	6,154

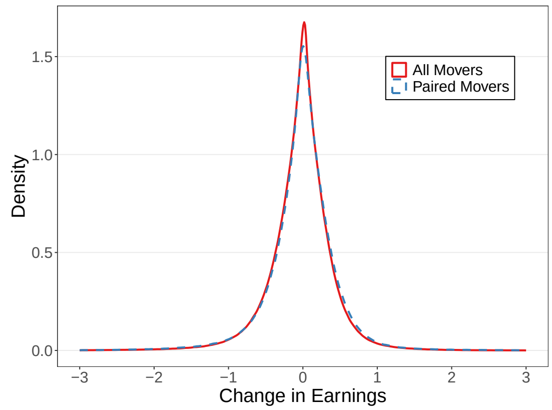
This table shows descriptive statistics of the Brazilian (Panel A) and Veneto (Panel B) data, described in Section 3. Column 1 displays all workers in the data and reports statistics over all employment spells. Columns 2 and 3 display different groups of movers and report statistics of the spell prior the move (origin firm). Column 2 shows all movers in the data and Column 3 shows the sample of paired movers used in estimation. Panel A uses the universe of spells in RAIS from 2007 to 2021, while Panel B uses Veneto Worker History data from 1985 to 2001. In both paired movers samples, we keep only one pair per worker per moving event.

Table 2: Paired movers *vs.* what would be predicted by chance

	N	% of total moves			
		Sector Stayers	Sector Movers	Cluster Stayers	Cluster Movers
	(1)	(2)	(3)	(4)	(5)
A. Brazil					
Paired Movers	1,981,101	45.5	54.5	32.1	67.9
All Movers		33.3	66.7	22.7	77.2
A1. Simulation					
Simulated Movers	823,946	50.3	49.7	35.1	64.9
p5 (200 draws)	823,377	50.3	49.6	35.1	64.8
p95 (200 draws)	824,804	50.4	49.7	35.2	64.9
A2. Random Benchmark					
Paired Movers		3.8	96.2	9.8	90.2
All Movers		6.0	94.0	5.0	95.0
B. Veneto					
Paired Movers	8,554	55.1	44.9	11.2	88.8
All movers		68.0	32.0	12.7	87.3
B1. Simulation					
Simulated Movers	3,749	61.7	38.3	10.4	89.6
p5 (200 draws)	3,689	61.5	37.3	9.8	89.2
p95 (200 draws)	3,761	62.7	38.5	10.8	90.2
B2. Random benchmark					
Paired movers		10	90	6.1	93.9
All movers		6	94	4.4	95.6

This table shows the number of paired movers that would occur by chance given the mobility patterns in the data. We take the sample of movers and drop workers in the paired movers sample. For each origin firm, we compute the distribution of destination firms over a one-year period. We use these distributions to simulate a sample of movers of the same size as in the data, and then create a paired movers sample. Panels A1 (Brazil) and B1 (Veneto) show the number of paired movers in the simulated samples, as well as the 5th and 95th percentiles of the distribution across 200 draws. In Column 1 we show the number of paired moves obtained, and columns 2 to 5 show the share of these moves that are within or across sectors and clusters. The first row shows the actual sample of paired movers in the data, while the second row shows the sample of all movers. In Panels A2 and B2, we compare those with a random benchmark formed by either the sample of paired movers or all movers. The benchmark is formed by taking the empirical probabilities of each origin and destination, but assuming that mobility is random.

Figure 1: Change in Earnings around the Move, Brazil



This figure shows the distribution of the change in log earnings between the destination and origin firms for all movers (Column 2, Table 1) and the final sample of paired movers (Column 3, Table 1).

is similar in the all mover sample and the paired mover sample. Where there are differences is that the variance of firm effects (and the share of the variance of earnings explained) is larger in the all mover than the paired mover sample: in the all mover sample it is about 15% of the variance of earnings, whereas in the paired mover sample it is about 10% of the variance of earnings.

Figure 1 shows that the distribution of earnings changes among the paired movers are broadly similar compared to all movers.

Veneto data: The Veneto data differs in a few ways from the Brazilian data. First, as in the Brazilian data the movers are negatively selected relative to all workers. But the paired movers do not fully undo that selection and are still lower earning than all workers. In addition, the paired movers are more rare than in the Brazilian data.

4.2 Main results

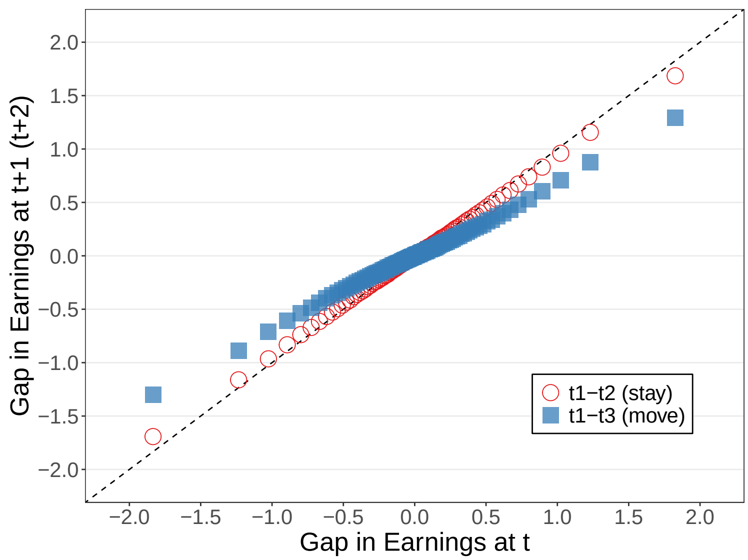
Figure 2 shows the graphical version of our main results. The x-axis shows the gap in earnings between a pair of workers in period t, and the y-axis shows the gap in earnings in a subsequent year. The red dots show the first stage. These dots are approximately linear and lie very close to the 45 degree line, indicating that earnings gaps are quite stable. The blue squares show the reduced-form. These dots are also approximately linear, but the slope is much shallower than between the years when

Table 3: AKM Estimates (KSS-corrected)

	All Workers (1)	All workers* (2)	All movers* (3)	Paired movers* (4)
A. Brazil, 2012–2017				
Var(Y)	0.506	0.501	0.442	0.478
Var(α)	0.283	0.294	0.245	0.307
Var(ψ)	0.072	0.065	0.059	0.046
Cov(ψ, α)	0.050	0.047	0.040	0.033
Var(ε)	0.051	0.048	0.058	0.059
R2	0.929	0.931	0.901	0.910
Adj. R2	0.907	0.910	0.869	0.880
Adj. R2, int.	0.934	0.937	0.924	0.926
Observations	40,950,644	40,950,644	22,599,194	1,059,548
# Firms	911,755	911,755	911,755	49,784
# Workers	8,764,893	8,764,893	4,598,015	211,588
# Movers	4,598,015	4,598,015	4,598,015	211,588
# Moves	14,585,164	14,585,164	14,585,164	644,018
B. Veneto, 1996–2001				
Var(Y)	0.766	0.714	0.697	0.802
Var(α)	0.317	0.305	0.241	0.451
Var(ψ)	0.129	0.104	0.091	0.059
Cov(ψ, α)	0.041	0.039	0.038	0.042
Var(ε)	0.238	0.227	0.289	0.208
R2	0.820	0.816	0.714	0.815
Adj. R2	0.753	0.746	0.618	0.750
Adj. R2, int.	0.797	0.797	0.759	0.812
Observations	5,111,991	5,111,991	1,702,650	10,828
# Firms	61,638	61,638	61,638	524
# Workers	1,332,729	1,332,729	363,956	2,262
# Movers	363,956	363,956	363,956	2,262
# Moves	988,088	988,088	988,088	5,612

This table shows the results of the AKM estimation with the KSS correction in Brazilian (Panel A) and Veneto (Panel B) data (Abowd, Kramarz, and Margolis, 1999; Kline, Saggio, and Sølvsten, 2020). The asterisk (∗) indicates residualized earnings. Columns 1 and 2 show the results for all workers. In Brazil, we use a 25% random sample of workers followed from 2012–2017, while in Veneto we use all work spells from 1996–2001. Column 3 restricts the samples only to workers that move at least once in the period, and Column 4 further restricts the sample to workers in the paired movers sample (and whose associated move in the paired movers sample is in the same period as the AKM estimation). We report the KSS-corrected variance terms, the R^2 and adjusted R^2 of the AKM model, and the adjusted R^2 of a regression of the log earnings on fully interacted worker and firm fixed effects. The last five rows show the number of observations, firms, workers, movers, and moves in the leave-one-out connected set used in estimation.

Figure 2: Changes in Gap in Earnings, Paired



This figure shows the graphical version of the paired movers test. The x-axis shows the gap in earnings in the second-to-last period in the origin firm. We split the gap in earnings into 100 bins. The red circles show the first stage, so the y-axis represents earnings in the last period at the origin firm. The blue squares show the reduced-form, so the y-axis represents earnings in the first period at the destination firm. The dashed line shows the 45-degree line.

the workers do not move. Hence, earnings gaps are not stable when workers move.

Table 4 shows the main results for paired stayers and movers. Paired stayers provide a useful quantitative benchmark in that they allow for the possibility of deviations from the assumptions of additively separability. For example, if there is drift in the individual effects, then the paired movers coefficient would interpret this as evidence for match effects, whereas in the paired stayers setting we would know that persistent match effects are not the explanation. Panel A shows the second stage estimates. For the stayers, the coefficient is 0.96, indicating minimal drift in individual effects between periods. In contrast, for paired movers the coefficient is 0.73, which is meaningfully different from 1 and from the stayers coefficient.

Robustness to residualization: In our baseline results, we residualize for a polynomial in age interacted with gender and education. Appendix Table A3 shows that our results are similar if we instead just residualize for the age polynomial interacted with gender. As an alternative to residualizing earnings, the Table shows similar coefficients when matching on gender and 10-year age bins or gender and 1-year age bins. We focus on the residualized approach because it preserves sample sizes and parallels the treatment of covariates in the AKM literature.

Reweighting by firm size: One of the strongest forms of selection into the paired mover sample is based on firm size. In Appendix Table A4 we reweight the paired movers so that either the origin firm size distribution matches the distribution among all workers, or the destination distribution does so. The point estimates are reasonably close: the baseline estimate is 0.73, reweighting based on origin firm size yields an estimate of 0.75, and reweighting the destination firm size yields 0.67.

Time series properties: In Section 1, we defined match effects to be a permanent feature of the match. Thus, up to measurement error the gap should be stable *within* the match. We already presented some evidence on this aspect of match effects by looking at paired stayers and finding a forecast coefficient of 0.96, which indicates small deviations from permanence.

The correlation between the gaps at various horizons and in various subsamples of paired movers provides further evidence. To help interpret these correlations, we

Table 4: The Paired Mover Coefficient

			Gap in Origin			
	Baseline	Stayers	Within Sector	Across Sector	Cluster Stayers	Cluster Switchers
	(1)	(2)	(3)	(4)	(5)	(6)
A. Brazil, 2nd Stage						
Gap in Origin	0.731	0.961	0.756	0.706	0.767	0.722
	(0.0007)	(0.0003)	(0.0010)	(0.0009)	(0.001)	(0.0010)
p-value, equality test			0.0000		0.0000	
Var(Z)	0.316	0.425	0.335	0.301	0.381	0.314
Γ_Z	0.693	0.954	0.717	0.667	0.729	0.691
Var($\tilde{\mu}_{ij} \mid \mathcal{P}$)	0.029	0.008	0.029	0.030	0.032	0.030
Observations	1,981,101	1,301,604	900,986	1,080,115	435,056	921,572
B. Brazil, 1st Stage						
Gap in Origin (lag)	0.916	0.938	0.917	0.916	0.930	0.919
	(0.0005)	(0.0003)	(0.0007)	(0.0007)	(0.0010)	(0.0007)
F-statistic	3,419,903.6	11,030,472.6	1,725,716.6	1,697,218.0	918,194.8	1,602,926.4
C. Veneto, 2nd Stage						
Gap in origin	0.783	0.942	0.848	0.674	0.707	0.783
	(0.028)	(0.0009)	(0.033)	(0.050)	(0.088)	(0.040)
p-value, equality test			0.0033		0.2627	
Var(Z)	0.394	0.381	0.419	0.363	0.383	0.430
Γ_Z	0.502	0.903	0.535	0.455	0.583	0.503
Var($\tilde{\mu}_{ij} \mid \mathcal{P}$)	0.021	0.010	0.017	0.027	0.033	0.023
Observations	8,554	1,021,944	4,709	3,845	548	4,338
D. Veneto, 1st Stage						
Gap in Origin (lag)	0.427	0.627	0.466	0.374	0.455	0.407
	(0.009)	(0.0005)	(0.012)	(0.013)	(0.032)	(0.012)
F-statistic	2,333.1	1,335,073.4	1,562.4	788.92	197.00	1,114.9

This table shows different versions of the paired movers regression. We estimate a 2SLS regression where the earnings gap in the destination firm is regressed on the earnings gap in the origin firm, instrumented by the lagged earnings gap in the origin firm. Panels A and B (C and D) show the second and first stages of the test in Brazilian (Veneto) data, respectively. Column 1 shows the baseline estimate using the full paired movers sample (Column 3, Table 1). Column 2 shows the estimate using a benchmark sample of paired stayers. Columns 3 and 4 show the estimate using the sample of paired movers split by whether they stay in the same sector or move to a different sector (2-digit CNAE code for Brazil and ATECO 91 code for Veneto). Columns 5 and 6 show the estimate using the sample of paired movers split by whether they stay in the same cluster or move to a different cluster. Clusters are estimated using the entire data following Bonhomme, Lamadon, and Manresa (2019). The clusters are only defined in the connected set, which is why the sample sizes fall. For Columns 3 to 6, we report the p-value of the equality test of the coefficients across columns for the relevant heterogeneity. Finally, we also report the variance of the gap in the origin, the measurement error correction term, and the variance of match effects under exogenous mobility (Equation 12).

convert them to implied AR(1) coefficients at each horizon[1] Panels B and D of Table 5 shows these correlations range from 0.92 to 0.98.

We follow Lachowska et al. (2023, Footnote 9) and pool the autocorrelation coefficients using a regression.[2] For the full sample, we find an estimate of the persistence parameter of around 0.96 (see Appendix Table A5). This estimate compares to 0.976 that Lachowska et al. (2023, pg. 387) find for firm effects. Thus, the match effects are almost as persistent as firm effects.

As another way to show the relative stability of match effects, Panel (a) of Figure 3 shows that our results are not very sensitive to the year in which we measure the earnings gaps. The left-side of the graph looks at paired movers who are at their origin firm and plots stayer coefficients on a rolling basis. The earnings gaps are quite stable at the origin firm. The right-side of the graph instruments for the gap at 0 with the gap at -1 and then looks at outcomes in different years at the destination firm. There is a dramatic drop from the left- to the right-hand side: gaps are much less stable when switching firms than staying within firms. The gaps are also quite stable regardless of which year we use in the destination firm. This evidence is consistent with the match effects being quite persistent.

Reverse regression: For our baseline results, we ask whether earnings gaps at the origin firm predict gaps at the destination firm. Conceptually, we could do the reverse, and ask whether the gaps in the destination firm predict the gaps in the origin firm. Appendix C shows that our results are robust to reverse versions or using alternate periods.

Alternative instruments: The baseline assumption of the instrumental variables approach is that the measurement error is independent across time. Hence, using a single lag is a valid instrument. If we relax this assumption and instead assume the error term follows a moving average process (of various orders), then this assumption justifies using deeper lags as instruments. Panel (b) of Figure 3 shows that the results of deeper lags up to five lags are quite similar to our baseline estimates. We interpret

[1] Let $\rho_{t,t+n}$ be the correlation of the gap between the paired workers at time t and $t+n$. Then under the $AR(1)$ model the persistence coefficient is given by $\frac{\rho_{t,t+2}}{\rho_{t,t+1}}$.

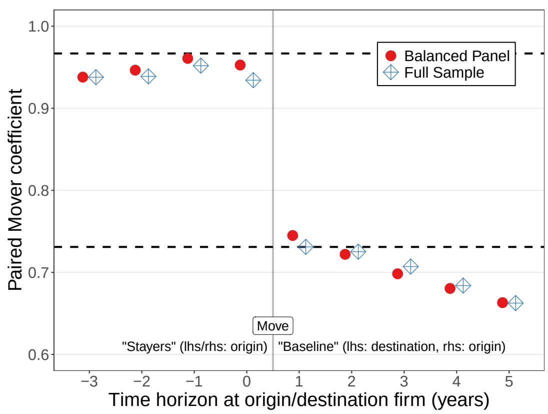
[2] For completeness, we reproduce the logic here. Let $y_t = \delta + \nu y_{t-1} + u_t$ so that the autocorrelation of y_t and y_{t-k} is $\rho(k) = \nu^k$. To estimate ν, fit $\ln(\rho(k)) = \beta k + \epsilon(k)$ where k is the lag order. The antilog of $\hat{\beta}$ is an estimate of ν.

Table 5: Earnings gap autocorrelation and implied AR(1) coefficients

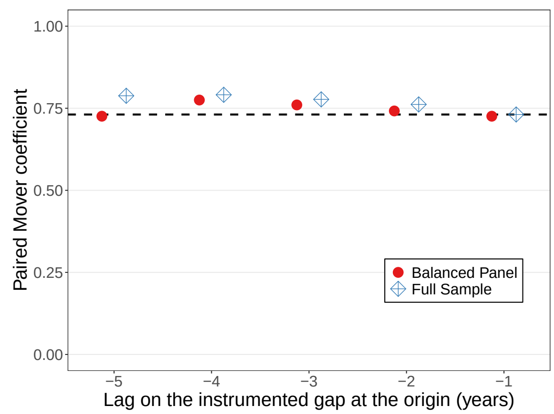
	Balanced Panel					Full Sample				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Origin Firm, autocor.										
Time lag	-1	-2	-3	-4	-5	-1	-2	-3	-4	-5
0	0.843	0.818	0.785	0.753	0.722	0.811	0.797	0.767	0.740	0.716
-1		0.959	0.913	0.875	0.837		0.952	0.896	0.860	0.828
-2			0.941	0.899	0.859			0.930	0.890	0.854
-3				0.937	0.893				0.931	0.888
-4					0.933					0.930
B. Origin Firm, AR(1) coef.										
Time lag	-1	-2	-3	-4	-5	-1	-2	-3	-4	-5
0		0.970	0.960	0.959	0.959		0.983	0.962	0.965	0.968
-1			0.952	0.958	0.957			0.941	0.960	0.963
-2				0.955	0.956				0.957	0.960
-3					0.953					0.954
C. Destination Firm, autocor.										
Time lead		2	3	4	5		2	3	4	5
1		0.922	0.878	0.837	0.774		0.852	0.818	0.789	0.762
2			0.934	0.889	0.820			0.880	0.843	0.812
3				0.933	0.859				0.890	0.853
4					0.897					0.896
D. Destination Firm, AR(1) coef.										
Time lead		2	3	4	5		2	3	4	5
1			0.952	0.953	0.925			0.960	0.965	0.966
2				0.952	0.922				0.958	0.963
3					0.921					0.958

This table shows the autocorrelation and implied AR(1) coefficients of the log earnings gap between paired movers at origin and destination firms for different time lags/leads. Panels A and C show the autocorrelation of the log earnings gap at origin and destination firms, respectively. Hence, the first row of Column 1 in Panel A shows the correlation between the log earnings gap at the last year at the origin firm ($t = 0$) and one year before that ($t = -1$). Panels B and D show the implied AR(1) coefficients at origin and destination firms, respectively, calculated as the ratio of the autocorrelation at time lag/lead k and $k + 1/k - 1$. We show results for both a balanced panel where all paired movers stay at the origin firm from $t = -4$ to $t = 0$ and at the destination firm from $t = 1$ to $t = 5$ (Columns 1 to 5) and the full sample of paired movers observed in a given time lag/lead (Columns 6 to 10).

Figure 3: Temporal Stability



(a) Event-study representation



(b) Deeper lags in the instrument

This figure shows different versions of the paired movers test. In Panel A, the lower dashed horizontal line shows the baseline movers estimate from Column 1 of Table 4. The higher dashed horizontal line shows the baseline stayers estimate from Column 2 of Table 4. The gray vertical line marks the timing of the move between origin and destination firms. To the left of the vertical line, we report the paired movers coefficient for stayers at different horizons. To the right of the vertical line, we report the paired movers coefficient where the first stage is always the last two years at the origin firm and the outcome rolls forward in time. In Panel B, we show the paired movers coefficient using deeper lags of the gap in earnings at the origin as an instrument. The x-axis shows the lag used in the instrument, so the first point corresponds to the baseline specification (lag of 1). The dashed horizontal line shows the baseline estimate from Column 1 of Table 4. In both panels, we report the coefficients from regressions using the full sample (all paired movers observed at that period) and a balanced panel (only paired movers observed at all periods).

this evidence as indicative that there is limited correlation in the errors.

An alternative instrument for the gap would be pre-determined characteristics of the workers such as age, education or gender. Conceptually, however, pre-determined characteristics of a worker affect the person effect and not the match effect. Instrumenting with components of the person effect would test whether the returns to these characteristics were constant across employers, rather than telling us about properties of the match effects. Appendix D presents a more formal discussion.

Veneto data: In the Veneto data, the paired mover coefficient is quite similar: 0.78 versus 0.73. The stayer coefficient is also similar: 0.94 rather than 0.96.

4.3 Match effects under exogenous mobility

Under exogenous mobility, we find the variance of match effects is 0.029.[3] This estimate is about 40% of the variance of firm effects that we found in Table 3 for the full sample, which is a conservative comparison.[4]

This estimator relies on the earnings gap between the paired movers. Panel (a) of Figure 4 shows the CDF of the absolute gap, which is closely related to the variance. While the median is around 0.36, there is a wide support. For the remainder of the figure we divide the paired movers into deciles of the absolute earnings gap.

The paired mover coefficient could driven by the instability of small earnings gaps at the origin firm. Panel (b) shows instead that the coefficient is *declining* in the size of the gap at the origin firm. The coefficient is smallest where the gap is largest.

The earnings gaps might also be less stable when they are larger. By decile, we compute the implied AR(1) coefficient on earnings gaps. Panel (c) shows that besides the first to second deciles, the gaps are *more* stable the larger they are.

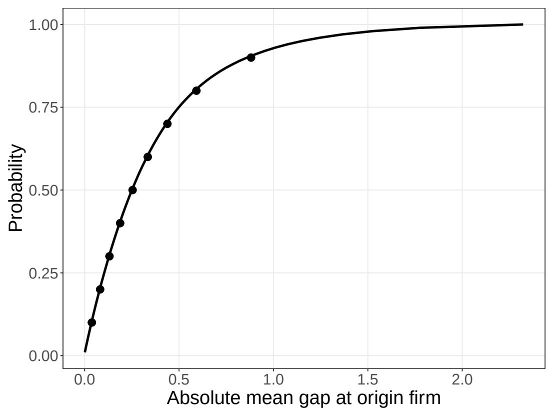
There are corners of the data with a very large role for match effects. Panel (d) combines the estimates from Panel (b) and (c) and computes the variance of match effects under exogenous mobility (see Appendix Table A6 for a tabular version).[5] In

[3] $\text{Var}(\tilde{\mu}_{ij} \mid \mathcal{P}) = (1 - \hat{\beta}_1)\,\text{Var}(Z)\frac{\Gamma_Z}{2} = (1 - 0.731) \times 0.316 \times \frac{0.693}{2} = 0.029$, see equation (12) and Table 4.

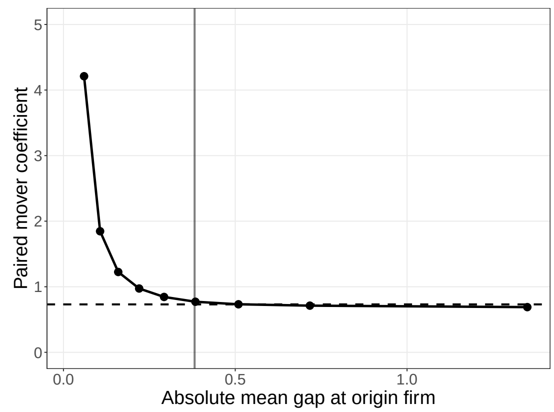
[4] Our estimate of the variance of firm effects is 0.072 in the full sample, 0.065 in the full sample using residualized earnings, and 0.046 in the paired mover sample. Gerard et al. (2021, Table 3) report estimates separately by demographic group and find variances of firm effects ranging from 0.056 to 0.086.

[5] The panel only shows this calculation for the top 6 deciles, because these are the deciles where

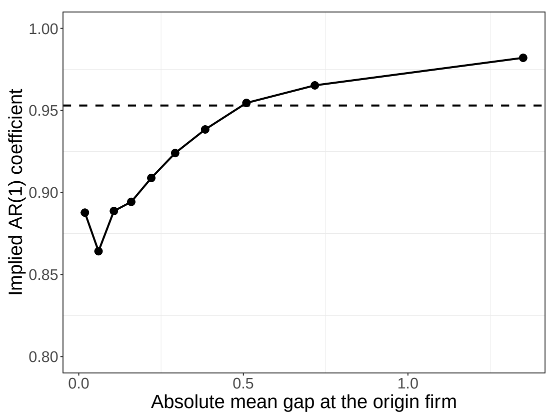
Figure 4: Paired Movers across the Gap Distribution



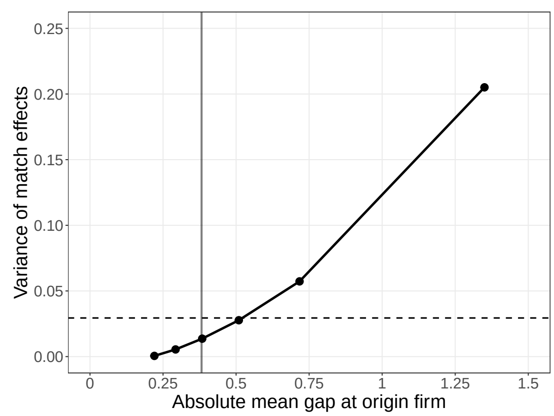
(a) Gap in earnings distribution



(b) Paired movers coefficient by gap decile



(c) Implied AR(1) by gap decile



(d) Variance of match effects

Panel A shows the cumulative distribution of the absolute gap in earnings in the last year at the origin, where the dots mark the cutoff of each decile. Panel B shows the paired movers' coefficient estimated using only movers within a given decile of the gap in earnings at the origin firm. The first two deciles are excluded due to large (imprecise) estimates. The dashed horizontal line shows the baseline estimate from Column 1 of Table 4, and the vertical line shows the average absolute gap. Panel C shows the implied AR(1) coefficient of the log earnings gap at the origin firm. The dashed horizontal line shows the AR(1) coefficient estimated using the full sample (Appendix Table A5). Panel D shows the variance of match effects under the assumption of exogenous mobility according to Equation 12. The first four deciles are not shown since the implied variance is negative. The dashed horizontal line shows the variance of match effects implied by the model (Table 6). In Panel D, summing up the variances across the six deciles and dividing by ten yields a mean implied variance of 0.031. Appendix Table A6 shows the numbers underlying Panels A–D.

the top decile the variance of match effects is over 0.20, which is more than 6 times the overall number. Summing the six estimates and dividing by ten, the implied variance of match effects is 0.031, slightly larger than the pooled estimator.

4.4 Heterogeneity

We consider heterogeneity by whether it is across contexts. We parameterize context first by whether the paired movers stay in the same sector or not, where we define a sector as a two digit CNAE code. Second, we parameterize it by whether or not the pairs stay in the same firm cluster, where we discuss how we construct the firm clusters in Section 6.

For both splits, the paired mover coefficient is lower for the across-context moves. Columns (3) and (4) of Table 4 shows that for sectors, it is 0.76 for stayers versus 0.71 for switchers. Columns (5) and (6) show that it is 0.77 for the cluster stayers and 0.72 for the cluster switchers. For sectors, the finding that earnings gaps are more stable for moves in more similar contexts generalizes. For each origin sector, we rank destination sectors by their frequency. Panel (a) of Figure 5 shows that as we move from more to less frequent destination sectors, the paired mover coefficient falls from 0.74 (for the stayers) to around 0.64 for sectors that are 11th and down in terms of likelihood. Retail trade is a very common destination sector in almost all origin sectors. Separating out retail trade, we find starker patterns.

For these splits, workers are more likely to stay in the context than to switch. Table 2 shows that 55% of moves switch sectors, whereas under random mobility 96% of moves would switch sectors. 68% of moves switch clusters, whereas under random mobility close to 90% of moves would switch clusters.

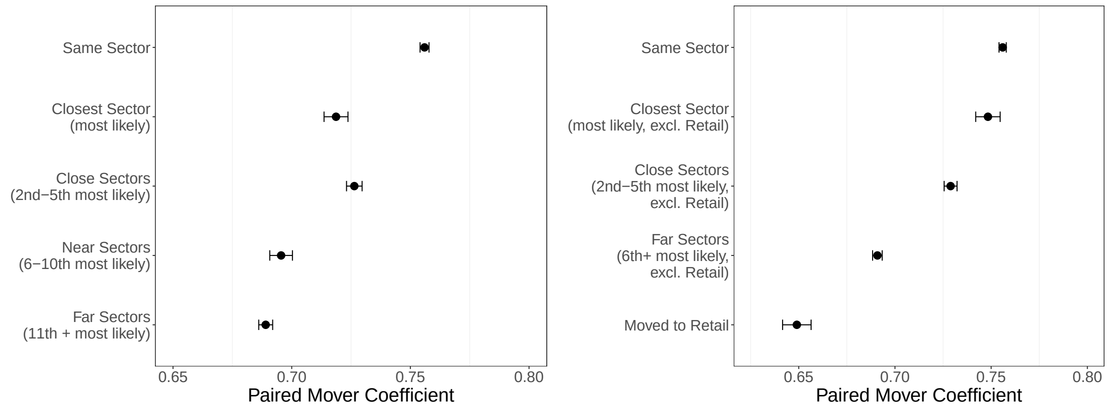
There are also different patterns of selection into who stays in a context versus switches. First, Appendix Table A7 shows that context-stayers tend to be higher-earners. Second, Table 4 shows that the variance of the earnings gap is larger for context-stayers.

These patterns of heterogeneity and selection are consistent with violations of exogenous mobility. If workers take match effects into account, then they should be less likely to switch across contexts that encode match effects. In Section 5, we develop a model of endogenous mobility that is consistent with these patterns.[6]

the paired mover coefficient is less than 1, which is required by the formula.

[6] For some splits and datasets, the variance of match effects under exogenous mobility is smaller

Figure 5: Moves within and across context



(a) By move likelihood (b) By move likelihood, excl. Retail

This figure shows the second stage of the paired movers test for different types of moves in the Brazilian data (Panels A and B). In Panel A, we report the coefficient for movers who stay in the same sector (2-digit CNAE code), as in Column 3 of Table 4, and then moves by decreasing "likelihood", defined as the share of moves from a given sector that go to the destination sector. We first show the coefficient for movers that go to the most likely destination sector, then the second to fifth most likely, and so on. Since the most likely destination sector for half of the sectors is Retail (the largest employer), in Panel B we replicate the exercise by categorizing Retail separately.

Additional heterogeneity: In Appendix Table A8, we present numerous other sample splits. We split moves on the basis of mean pay at the firm where we split firm pay at the median of the distribution, on whether one or both workers in the pair switch occupation, on the basis of worker tenure among the paired movers, and on whether the paired moves are within or across geographies. In Appendix Figure A1 we do so for the time between employment spells. In general, the longer time between spells yields a lower coefficient. Appendix Table A9 splits the sample by whether the higher-earner (in the pair) moved first. The gaps are more stable when the higher earner moved first.

Veneto data: The results on heterogeneity in Veneto are generally more stark than in Brazil. For example, the sector stayers have a coefficient of 0.85 whereas the switchers have a paired mover coefficient of 0.67.

---

for the switchers. The model in Section 5 rationalizes this difference as resulting from endogenous selection into mobility on the basis of match effects.

5 A model of endogenous mobility

Assuming exogenous mobility, we developed an estimator for the variance of match effects. In the last section we developed some evidence that is inconsistent with exogenous mobility. Therefore, in this section, we develop and estimate a partial equilibrium search model to estimate the variance of match effects in the presence of endogenous mobility. We show why the presence of match effects matters. Finally, we develop an extension of the model with a notion of context which allows us to rationalize the patterns of heterogeneity documented in Section 4.4.

5.1 The model

The model is a standard partial equilibrium search model. Workers have a fixed effect $h_i \sim N(0, \sigma_w^2)$. Firms have a firm effect $p_j \sim N(0, \sigma_f^2)$. This distribution is also the offer distribution, denoted by F. The match effects distribution is $\mu_{ij} \sim N(0, \sigma_m^2)$. Denote this distribution by M. A worker's payoff to working at a firm j with firm effect p_j and match effect m_{ij} is given by $h_i + p_j + m_{ij}$.

The value of an employed worker with person effect h_i at a firm with firm effect p_j and with match effect m_{ij} is:

$$W_i(p_j, m_{ij}) = h_i + p_j + m_{ij} + \beta\bigg(\lambda_1(1-\delta)\int_m\int_p \max\{W_i(p_j, m_{ij}), W_i(p', m')\}dF(p')dM(m') \\ + (1-\lambda_1)(1-\delta)W_i(p_j, m_{ij}) + \delta U_i\bigg). \quad (13)$$

An employed worker receives the flow payoff of her match. If her job is not destroyed, then she can receive an outside offer, and she makes a maximizing decision whether or not to accept it. The offer includes both the firm effect and the match effect and the worker considers both in her mobility decision. Thus, this model violates both additive separability and exogenous mobility. If she does not receive an offer (and her job is not destroyed), she remains in her job. With probability δ her job is destroyed.

The value function for an unemployed worker is standard and so is omitted.

5.2 Estimation and fit

We estimate the model by the method of simulated moments. We assume that b is negative infinity so that all offers from unemployment are accepted. We simulate the model forward until a steady state where the distribution of the firm effects, worker effects, and match effects among employed workers is stable.

Estimation proceeds in two steps. In the first step, we estimate the EU rate (δ) and UE rate (λ_0) by matching quarterly flow rates computed in survey data. In the second step, we pick four unknowns, the arrival rate of offers on the job (λ_1), and the variance of person, firm and match effects, to match four empirical moments: the EE rate, the paired mover regression coefficient, the variance of firm effects (estimated using Kline, Saggio, and Sølvsten (2020)), and the measurement error-corrected variance of the earnings gap between paired movers that we documented in Table 4.[7] We search the parameter space using Metropolis-Hastings.

The first part of Table 6 shows that we are able to closely replicate the four moments that we match. Appendix E discusses how model parameters map into the paired mover coefficient.

5.3 Results

The fourth part of Table 6 shows that the variance of match effects in the steady state distribution are quite similar to those in the offer distribution. Hence, our main estimate is that the variance of the match effects is about 50% of the variance of the firm effects, or 20% larger than the estimate under exogenous mobility.

Panel (a) of Figure 6a shows that earnings changes on accepted offers do a poor job of predicting earnings changes on rejected offers. At the firm-level, the x-axis computes the mean change in earnings for workers who move from A to B on EE moves (and thus all earnings changes are positive). For those same firm pairs, the y-axis plots the change in earnings for the rejected offers. We then bin the firms. The notable feature of the Figure is first that the intercept is negative, consistent with the fact that offers that would lead to earnings cuts are rejected in this model. Second, the slope is not 1 (even when we split the sample and instrument). Thus,

[7]This model has no way of generating a correlation between person effects and firm effects. The way to generate such a correlation would be to allow the arrival rate of offers to be increasing in the worker type, as in Lentz (2010).

Table 6: Model results

	(1)	(2)	(3)	(4)	(5)
	EE rate	β_1	Var(Z)$\times\Gamma_Z$	Var(ψ_j)	Var(α_i) (untargeted)
Empirical Moments (Data)	0.038	0.731	0.219	0.065	0.294
Simulated Moments (Model)	0.038	0.734	0.219	0.065	0.076
	λ_0 (UE rate)	δ (EU rate)			
Calibrated Parameters	0.072	0.076			
	λ_1	$Var(m_{ij})$	$Var(h_i)$	$Var(p_j)$	
Estimated Parameters	0.114	0.035	0.078	0.072	
Steady-state (employed)		m_{ij}	h_i	p_j	
Mean		0.060	0.000	0.099	
Variance		0.035	0.078	0.065	

This table shows the model calibration. We fix λ_0 and δ to quarterly UE and EU rates in Brazil, measured using labor force surveys (PNAD). We estimate the four remaining parameters of the model, λ_1, $Var(\mu_{ij})$, $Var(h_i)$, and $Var(p_j)$, to match: the EE rate, the paired movers coefficient, the variance of the gap, and the variance of firm effects. The first two rows show the empirical moments in the data and the simulated moments in the model. The next two rows show the calibrated and estimated parameters of the model. The last two rows show the steady-state moments of employed workers in the model.

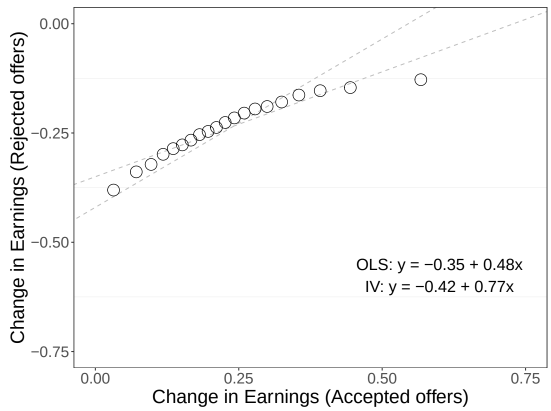
the presence of match effects means that the magnitude of earnings changes among accepted offers is a poor guide to the earnings changes among other potential moves (the rejected offers).

5.4 Extended model with a notion of context

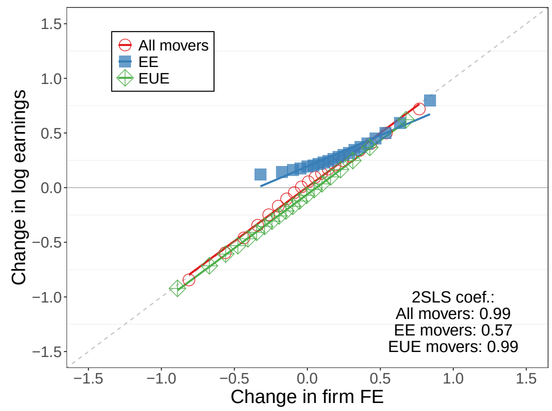
Our results above show that the paired mover coefficient looks quite different for moves within and across contexts, and that moves across context are relatively rare. In addition, there are two patterns of selection: the within-context moves tend to happen among higher-earnings workers and the variance of the gaps are larger for the moves within-context. Combined, this evidence suggests violations of exogenous mobility where workers are taking into account match effects in mobility.

To formalize this logic, consider an extended version of the model that is the same as above except that the match component has two dimensions: a sector component and a firm-specific component. If the worker stays in the same sector, then they retain their sector component. If the worker switches sectors, then they get a new draw of the sector component. (In this sense, the model is identical to Neal (1999)).

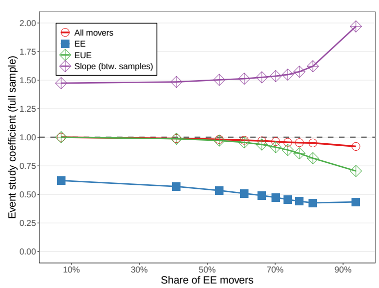
Figure 6: Model Simulated Data



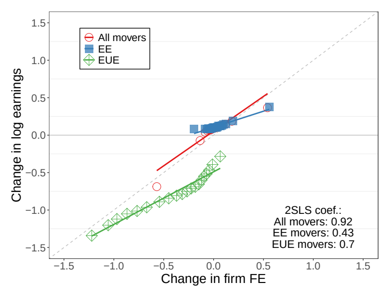
(a) Accepted *vs* Rejected Offers



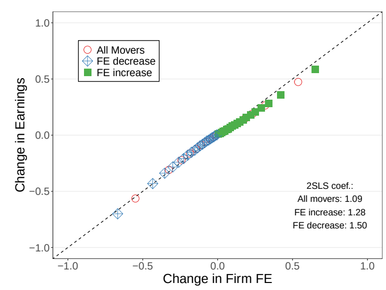
(b) Event Study, $\lambda_1 = 0.11$



(c) Event study by share EE



(d) Event Study, $\lambda_1 = 0.99$



(e) Event Study, data

This figure shows results using data simulated from the model. In Panel A, we compare the change in log earnings for accepted and rejected offers. We match workers who move from firm A to B with workers at A who reject offers at B. The regression is run on the individual-level data, and the IV coefficient comes from splitting the data randomly. Panels B and D show the event study test in different samples. We plot the change in log earnings between destination and origin firms against the change in firm effects. Panel B uses the baseline parameter $\lambda_1 = 0.12$, while Panel D sets $\lambda_1 = 0.99$. Iin Panel C, we vary the parameter λ_1 from 0 to 0.99 to show how changing the share of EE moves affects results. We also plot the between-sample slope. Panel E shows the event study estimated in Brazilian data (Column 1, Appendix Table A12), split by whether the move has an increase in firm effects or not. In all panels we split the sample randomly to generate two estimates of the firm effects to be able to instrument for the change in the firm effects.

Because a move across sectors entails new draws of both dimensions of match quality whereas a move within-sector only entails a draw of one dimension, the model has a force to generate a lower paired mover coefficient on sector-switching.

In addition, rather than being normally distributed, we make the sector component be Pareto distributed. For most distributions, endogenous mobility will imply that the sector stayers (but firm switchers) will be higher-earnings than the sector movers. The thick right tail means that sector stayers will also have a larger variance of the earnings gap, because the high-earners will be further apart than the low earners.

To quantify this intuition, we use 20 sectors. Appendix Table A10 shows that when we target the difference between the paired mover coefficient for within- and between-sector moves we qualitatively replicate the relative rarity of between- versus within sector moves in that workers are twice as likely to accept within sector as between sector offers. Similarly, the within-sector movers are higher paid and have a larger variance of the gap. Finally, fewer moves are cross-sector than a random mobility benchmark would predict.

6 Revisiting specification tests in the literature

So far we have developed evidence indicating a large role for match effects both as a share of the variance of earnings, as well as in driving mobility, i.e., endogenous mobility. This evidence is inconsistent with the results of three exercises in the literature that are supportive of both exogenous mobility and additive separability. In this section, we use a combination of the estimated model and the paired movers to revisit this evidence.

6.1 Event study test

The test that launched the modern Abowd, Kramarz, and Margolis (1999) literature is the seminal event study analysis in Card, Heining, and Kline (2013). The event study test asks whether changes in the firm effects (or mean co-worker earnings) predicts changes in earnings within worker. The event study test is a concrete feature of the data that is consistent with exogenous mobility. Therefore, if the model that we developed in the last section could not pass the event study test, then it would not be a compelling way of interpreting the paired mover evidence under endogenous mobility. So it is important to know if the model can pass the event study test and, if so, why.

We construct the event study test in the following way. We randomly split the sample of workers and estimate two sets of firm effects. We then instrument the measurement of the change in firm effects estimated in the first sample with the change in firm effects estimated in the second sample.

Event study test in the model: Intuitively, EE moves give more scope to select on match effects than EUE moves and so should generate a lower event study coefficient. Consistent with this intuition, Panel (b) of Figure 6 shows that on EE moves the coefficient is 0.57 and on EUE moves it is 0.99 (Appendix Table A11 shows regression versions). Thus, here the test works exactly how researchers expect it to and would detect that there are deviations from exogenous mobility on the EE moves.

What is surprising, however, is that in the full sample we find an event study coefficient of 0.99. We might expect that the overall coefficient would be between the coefficient on the EE split and the EUE split. Instead, the coefficient is (very slightly) above the coefficient on the EUE split. Even though there is a significant portion of the data that features endogenous mobility (41% of moves are EE), the model generates data that "passes" the event study test.

To explain this result, recall that OLS combines regression coefficients in two subsamples as a function of three things: relative sample sizes, the within-group variances, and the between-group differences in sample means. The relative sample size and variance logic implies that the pooled coefficient should lie in between the coefficients in the two samples. The difference in sample means can generate unintuitive behavior; for example, Simpson's paradox occurs when the pooled coefficient is a different sign than the coefficient in each subsample.

Panel (c) of Figure 6 shows that the event study test fails to find convincing evidence against exogenous mobility for a wide range of the shares of moves that are EE. Increasing the arrival rate of offers on the job, λ_1, increases the share of EE moves. As the share of EE moves rises both the EE and EUE event study coefficients fall. In both cases, a higher share of EE moves means that workers are more selected on match effects at their origin firm. Even as the coefficients in the subsamples decrease, the pooled event study coefficient remains above 0.90, exceeding the coefficient in either subsample. The reason is that the between-sample slope is above one and so pushes up the overall coefficient. Panel (d) looks at the simulated data for the highest share of EE moves where it is visually easy to see the role of the between-sample comparison.

Event study test in the data: Panel (e) of Figure 6 and Appendix Table A12 show that the forecast coefficient in the event study test is 1.09. The standard errors are tight enough that we can reject 1, but we do not interpret this deviation as economically large.[8] In the Veneto data, the event study coefficient is 1.18.[9]

Panel (e) of Figure 6 shows that the between-sample logic is present in the data. We split the moves based on whether there are increases or decreases in the firm fixed effects. In these two subsamples the event study coefficient *exceeds* the coefficient in the pooled regression: it is 1.28 on the increases and 1.50 on the decreases. Thus, the event study test looks less "good" in these subsamples and pooling makes the event study test look better. In Appendix Figure A2 we repeat this sample split in model-generated data. We do so in an endogenous mobility version of the model (what we introduced above) and an exogenous mobility version where workers accept all offers. In the exogenous mobility version of the model, in both subsamples and in the pooled regression we find an event study coefficient of 1. In the endogenous mobility version of the model the coefficient is 0.86 on the increases and 1.12 on the decreases. Quantitatively, the gap between the coefficients on the increases and decreases is very similar to what we find in the data.

Discussion: Sample means are likely to be different on endogenous moves than exogenous moves. In the context of the model, workers are more likely to see increases in firm effects and earnings on endogenous than exogenous moves. In other applied economics contexts, it is likely that endogenous moves also feature more increases in unit effects and in the outcome of interest. This pattern generates the between-sample slope. This slope is also likely to be above one: insofar as there are some

[8] In Appendix Figure A1, we proxy for the EE/EUE split by we splitting the data by deciles of the time between employment spells. We find non-monotonic patterns that are loosely consistent with the model: the event study coefficient is the furthest from one at low and high levels of the gap. The increasing event study coefficient and decreasing match paired movers coefficient could be rationalized in an extended version of the multi-sector model where the decision to search outside the home sector is endogenous and its likelihood is increasing in unemployment duration.

[9] The Brazilian and Veneto data differ in the event study coefficient in the paired mover sample: in the Brazilian data it is 1.07 while in Veneto it is 0.66. The reweighting exercise in columns (3) through (5) shows that this difference can be explained by the set of firms we are considering.

match effects, a given increase in the firm (or unit) fixed effects will likely translate more than one for one into individual outcomes.

Thus, even if there are large corners of the data where workers move on the basis of match effects, the test might not fail if there are also non-trivial corners where workers do not move on the basis of match effects. Unless the researcher ex-ante knows the corner of the data where exogenous mobility fails, the event study test is unlikely to detect this failure.

This between-sample reason why the event study test can fail to reject is new to the literature. Bonhomme, Lamadon, and Manresa (2019) emphasize an explanation similar to the distinction between exogenous mobility and additive separability discussed in Section 1: even if there are match effects, if mobility is exogenous then symmetry will hold. Relatedly, Borovičková and Shimer (2024) emphasize that if the wage equation is additively separable, then the data will pass the event study test.

6.2 Adjusted-R^2 in AKM and the interacted model

Card, Heining, and Kline (2013) show that the adjusted-R^2 only increases slightly between an additively separable model and one with match fixed effects, which is interpreted as evidence against the importance of match effects. Versions of this exercise are standard in the literature (e.g., Card, Heining, and Kline (2013, pg. 995) and Bonhomme, Lamadon, and Manresa (2019, pg. 718-9). Table 3 shows that going from an additively separable model to a model with arbitrary firm and worker interactions increases the adjusted-R^2 from 0.91 to 0.94, which implies a variance of match effects of 0.013 (in the paired mover sample it is 0.020).[10] This variance of match effects is smaller than what the paired mover estimator finds.

We use the model to assess this R^2-evidence. Holding fixed the parameters of the model, we simulate data under two assumptions on mobility and two panel lengths. The assumptions on mobility are exogenous mobility—where workers accept all offers—and endogenous mobility, which is the baseline model. The panel lengths are 100 periods and 7 periods, the number of periods in Card, Heining, and Kline (2013).

We report three ways of estimating the variance of match effects. We report two

[10] $(0.937 - 0.910) \times 0.501 \times 0.937 = 0.013$, where the last 0.937 is a simple way of adjusting the variance of earnings for error. This calculation uses the full sample with residualized earnings. The calculation using only the paired movers is $(0.926 - 0.880) \times 0.478 \times 0.926 = 0.020$.

versions of the adjusted-R^2 calculation. The first compares the adjusted-R^2 from the interacted and uninteracted model and multiplies this by the variance of earnings to get an estimate of the variance of match effects (in the data, we also adjust for noise by multiplying by the adjusted-R^2 from the interacted model). The second is similar to the first but uses the KSS variance components to construct an implied R^2 of the uninteracted model, and then assigns an R^2 of 1 to the interacted model because there is no error term. The third is the paired mover estimator.

Table 7 shows that under exogenous mobility and long panels the KSS-version of the adjusted R^2 calculation works. Specifically, in column (1) the KSS version matches the structural variance of match effects. Similarly, the paired mover estimator exactly recovers the structural variance of match effects. In contrast, the standard adjusted-R^2 estimator is biased down.

With endogenous mobility (but retaining the long panels), the KSS approach is approximately unbiased, while the adjusted-R^2 approach is further downward biased, as is the paired mover estimator.

With short panels, the adjusted-R^2 logic is quite biased. With exogenous mobility it is biased down by 67% and with endogenous mobility it is biased down by an order of magnitude: the implied variance of match effects is only 11% of the truth. The KSS-based estimator is also quite biased, while the paired mover estimator has identical bias (or its absence) in short panels as in long panels.

Hence, in the presence of realistic data generating processes the adjusted-R^2 estimator is a downward biased estimator of the variance of match effects, which means that there is more scope for match effects than previously thought.

6.3 Bonhomme, Lamadon, and Manresa (2019)

Bonhomme, Lamadon, and Manresa (2019) (BLM) develops methods designed to detect deviations from additive separability. Do the paired movers just rediscover features of the data already documented by BLM? In this section, we show that the BLM approach does not reliably detect the deviations that paired movers detect.

The BLM model and estimation procedure: We consider the static model in BLM. In the static model, firms are assigned to classes, and workers are assigned to types. The static model in BLM parallels the AKM model in its key assumptions.

Table 7: Match effects estimators under different mobility assumptions

	Long Panel ($T = 100$)		Short Panel ($T = 7$)	
	Exogenous (1)	Endogenous (2)	Exogenous (3)	Endogenous (4)
A. Structural parameters				
Var(Y)	0.185	0.180	0.185	0.180
Var(h_i)	0.078	0.078	0.078	0.078
Var(p_j)	0.072	0.069	0.072	0.069
Var(μ_{ij})	0.035	0.035	0.035	0.035
B. Variance components estimates				
Var(α_i) - AKM	0.084	0.088	0.105	0.109
Var(ψ_j) - AKM	0.072	0.068	0.072	0.062
Var(α_i) - KSS	0.078	0.077	0.095	0.105
Var(ψ_j) - KSS	0.072	0.068	0.072	0.061
C. Intermediate statistics				
R2	0.843	0.866	0.958	0.983
Adj. R2	0.840	0.863	0.941	0.976
Adj. R2, interactions	1.000	1.000	1.000	1.000
Paired mover coefficient (β)	0.689	0.738	0.689	0.731
Gap Variance (Var(Z))	0.226	0.220	0.226	0.220
D. Implied variance of match effects under exogenous mobility				
Var(m_{ij}) - KSS	0.035	0.036	0.018	0.014
Var(m_{ij}) - Adj. R2	0.030	0.025	0.011	0.004
Var(m_{ij}) - Equation 12	0.035	0.029	0.035	0.030

This table shows the performance of several match effects estimators with different panel lengths and mobility assumptions. We simulate data from the model in Table 6 under two scenarios: exogenous mobility, where workers accept any outside offer, and endogenous mobility. We simulate two panel lengths: a long panel with 100 periods (Column 1 and 2) and a short panel with 7 periods (Column 3 and 4). Panel A shows structural parameters: the true variances of wages, worker effects, firm effects, and match effects. Panel B shows the variance of worker and firm effects estimates. Panel C shows intermediate statistics: the (adjusted) R^2 of the two-way fixed effects regression, the adjusted R^2 of a fixed effects regression with worker-firm interactions, the paired mover coefficient, and the absolute mean gap between paired movers in the simulated data. Finally, Panel D shows the variance of match effects implied by the KSS estimates, $Var(Y) \times \left(1 - \frac{Var(\alpha_i)+Var(\psi_j)}{Var(Y)}\right)$ (the model generates a covariance term of zero and has no measurement error), the adjusted R^2 estimates $(1 - \text{Adj. } R^2) \times Var(Y)$, and the formula in Equation 12 (which assumes exogenous mobility). All results are averaged over 100 simulations.

Where the static model is richer than the AKM model is that it allows for interactions in the wage equation.

Estimation proceeds as in BLM. The key step is that we give the k-means clustering algorithm 30 percentiles of the log-earnings distribution among stayers. We use 10 firm clusters and 6 worker types. The static model implies worker-type specific transition probabilities across firm classes.

We compute the steady state distribution implied by these transition probabilities. From steady state, we simulate a set of moves. To construct wages, we we draw from the error term. For the origin firm, we draw two error terms in order to be able to exactly replicate what we do in the data and instrument. In this simulated dataset, we then run the paired movers regression and construct other statistics about the paired movers.

We estimate the BLM model in two datasets. First, we use all movers. This parallels the standard implementation of BLM. Second, we use only the paired movers. This enable an apples-to-apples comparison in that the BLM estimator is looking at the exact same data.

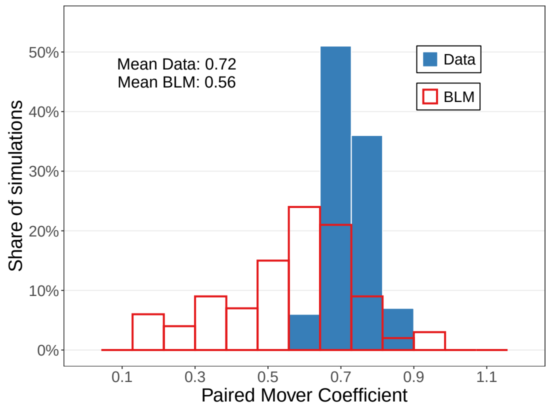
For each of these exercises, we draw 100 random subsamples of the data in order to show the distribution of BLM-implied estimates of the paired mover coefficient (and mean absolute earnings gaps) compare to what we find in the data. For all movers, we use subsamples of 500,000. For paired movers, we use subsamples of 50,000.

The paired movers test in data based on the BLM model: Figure 7 shows that the data simulated from BLM estimates does not reflect the underlying patterns in the data and the direction of the bias cannot be signed. Panel (a) shows that the BLM-implied estimates of the paired mover coefficient are biased down relative to the data. Moreover, they are much more dispersed. In contrast, Panel (b) shows that using only the paired movers the BLM-implied estimates of the paired mover coefficients are biased up.

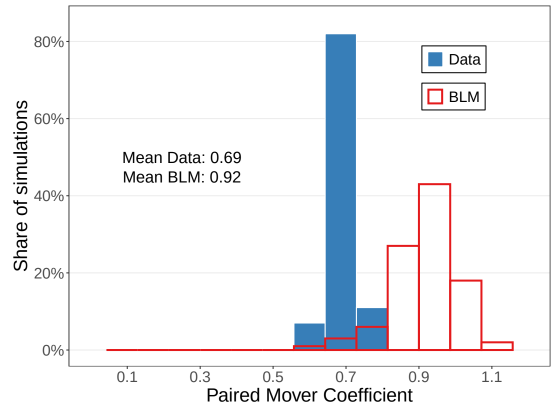
We interpret this evidence as indicative that the deviations from additive separability that the paired movers document are different from what the literature has previously found.

Monte Carlo evidence using the model: The evidence so far suggests that BLM estimation is not able to reliably recover properties of the paired movers, which is

Figure 7: Paired Movers and BLM



(a) All Movers



(b) Paired Movers

This figure shows the distribution of the paired movers coefficient in the data and in data simulated from the estimated BLM model across 100 iterations. In Panel A, in each run we sample 500,000 workers from the entire dataset of workers (Column 1 of Table 3). We then compute the paired movers coefficient using this sample, estimate BLM on the same sample, and simulate data from the estimated BLM model to obtain the paired movers coefficient in the simulated data. In Panel B, we repeat the same exercise but we sample 50,000 workers from the sample of paired movers (Column 4 of Table 3).

the minimal feature of the data that speaks directly to the presence of match effects. To understand why this failure occurs, we simulate data from the economic model in Section 5. We work with a discrete heterogeneity version of the model where the number of firm types and worker types matches that assumed in the BLM estimation procedure. As above, we consider exogenous and endogenous mobility versions.

Appendix Figure A5 shows that BLM accurately recovers the paired mover coefficient under exogenous but not endogenous mobility. The figure further shows that under either type of mobility the BLM estimator successfully recovers the wage equation. Where BLM struggles is that under endogenous mobility it misses on the mobility patterns: specifically, it finds that the destination firms are almost uniformly distributed among firm types, while in reality the higher firm types are more likely destinations.[11] Because the paired mover coefficient and the gap variance are functions of the mobility patterns, missing on these leads to inaccuracies in the estimator.

Thus, BLM does not accurately recover the data underlying paired movers.

[11]This finding is consistent with the Appendix of BLM, which reports Monte Carlo evidence, but only about the wage equation and not mobility patterns.

6.4 Quantitative comparison to Borovičková and Shimer (2024)

Borovičková and Shimer (2024) calibrate their model to match Kline, Saggio, and Sølvsten (2020)'s result on Veneto data. They attribute the fact that their model can pass the event study test to the approximate additive separability in their model. In Appendix Table A13 we report results of the paired mover and event study tests using data simulated from their model. The paired mover coefficient is 0.62. Thus, there is evidence against additive separability in the model-generated data. The event study coefficient is around 0.96. We run the event study test splitting the data by EE and EUE moves. Both coefficients are below the coefficient in the pooled sample (0.95 in the EUE sample and 0.35 in the EE sample). Therefore, the between-sample logic explains why the model-generated data passes the event study test in their model.

7 Discussion

The idea of this paper is that paired movers—that is, workers who both move from firm A to firm B—are the simplest feature of the data that allows us to see match effects. We study these paired movers in administrative data and find meaningful deviations from the additively separable wage equation that is popular in empirical work. Under exogenous mobility, we find a variance of match effects that is about 40% of the variance of firm effects. Strikingly, there are larger deviations for statistically rarer moves. This link between match effects and mobility is consistent with deviations from exogenous mobility.

We estimate a simple job search model to interpret the paired mover evidence under endogenous mobility. The variance of match effects is about a fifth larger than under exogenous mobility. Because of match effects, the earnings changes we see in the data are a poor guide to earnings changes on the moves that we do not see.

Finally, we revisit three pieces of evidence taken as supportive of the AKM specification. We first show that it is possible to pass the influential event study test because of the between-sample slope, even in the presence of exogenous mobility. We then show that the adjusted-R^2 logic produces a severely downward biased estimate of the variance of match effects given endogenous mobility and short panels. Finally, we show that the Bonhomme, Lamadon, and Manresa (2019) approach does not reliably detect deviations from additive separability using our paired mover test.

This paper suggests two exciting avenues for future work. First, the results in this paper suggest that—since they are large—trying to understand the structure and determinants of match effects is an important task. Second, while this paper has developed results in the context of an additively separable wage equation, there are numerous other applied contexts where researchers have used movers designs (e.g., Finkelstein, Gentzkow, and Williams (2016), Chetty and Hendren (2018), and Lachowska, Sorkin, and Woodbury (2025)). Studying the paired mover sample could be informative as to the presence of match effects in these contexts as well.

References

Abowd, John M, Francis Kramarz, and David N Margolis. 1999. "High wage workers and high wage firms." *Econometrica* 67 (2):251–333.

Beauregard, Pierre-Loup, Thomas Lemieux, Derek Messacar, and Raffaele Saggio. 2025. "Why Do Union Jobs Pay More? New Evidence from Matched Employer-Employee Data." Tech. Rep. 33740, National Bureau of Economic Research.

Benedetto, Gary, John Haltiwanger, Julia Lane, and Kevin McKinney. 2007. "Using worker flows to measure firm dynamics." *Journal of Business & Economic Statistics* 25 (3):299–313.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "A distributional framework for matched employer employee data." *Econometrica* 87 (3):699–739.

Borovičková, Katarína and Robert Shimer. 2024. "Assortative Matching and Wages: The Role of Selection." Tech. rep., National Bureau of Economic Research.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and labor market inequality: Evidence and some theory." *Journal of Labor Economics* 36 (S1):S13–S70.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women." *Quarterly Journal of Economics* 131 (2):633–686.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace heterogeneity and the rise of West German wage inequality." *Quarterly Journal of Economics* 128 (3):967–1015.

Chetty, Raj and Nathaniel Hendren. 2018. "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects." *Quarterly Journal of Economics* 133 (3):1107–1162.

Eeckhout, Jan and Philipp Kircher. 2011. "Identifying sorting—in theory." *The Review of Economic Studies* 78 (3):872–906.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of geographic variation in health care: Evidence from patient migration." *Quarterly Journal of Economics* 131 (4):1681–1726.

Gerard, François, Lorenzo Lagos, Edson Severnini, and David Card. 2021. "Assortative matching or exclusionary hiring? The impact of employment and pay policies on racial wage differences in Brazil." *American Economic Review* 111 (10):3418–3457.

Gibbons, Robert, Lawrence F Katz, Thomas Lemieux, and Daniel Parent. 2005. "Comparative advantage, learning, and sectoral wage determination." *Journal of Labor Economics* 23 (4):681–724.

Hagedorn, Marcus, Tzuo-Hann Law, and Iourii Manovskii. 2017. "Identifying equilibrium models of labor market sorting." *Econometrica* 85 (1):29–65.

Kantenga, Kory and Tzuo-Hann Law. 2016. "Sorting and Wage Inequality." Tech. rep., University of Pennsylvania. URL `https://red-files-public.s3.amazonaws.com/meetpapers/2016/paper_660.pdf`.

Kline, Patrick. 2024. "Firm wage effects." *Handbook of Labor Economics* 5:115–181.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2020. "Leave-out estimation of variance components." *Econometrica* 88 (5):1859–1898.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A Woodbury. 2023. "Do firm effects drift? Evidence from Washington administrative data." *Journal of Econometrics* 233 (2):375–395.

Lachowska, Marta, Alexandre Mas, and Stephen A Woodbury. 2020. "Sources of displaced workers' long-term earnings losses." *American Economic Review* 110 (10):3231–3266.

Lachowska, Marta, Isaac Sorkin, and Stephen A Woodbury. 2025. "Employers and unemployment insurance take-up." *American Economic Review* 115 (8):2529–2573.

Lentz, Rasmus. 2010. "Sorting by search intensity." *Journal of Economic Theory* 145 (4):1436–1452.

Neal, Derek. 1999. "The complexity of job mobility among young men." *Journal of Labor Economics* 17 (2):237–261.

Sorkin, Isaac. 2018. "Ranking firms using revealed preference." *Quarterly Journal of Economics* 133 (3):1331–1393.

Woodcock, Simon D. 2015. "Match effects." *Research in Economics* 69 (1):100–121.

Supplementary Appendix to Paired Movers

Isaac Sorkin and Lucas Warwar

A Match effects under exogenous mobility

Here we derive the estimator for the variance of match effects under exogenous mobility. Define $\mathcal{P}$ to be the set of paired movers.

Step 1: solve for $\text{Var}(\alpha_i - \alpha_{i\prime})$**:** Define the mean worker effect in the origin firm j among the paired movers to be: $\bar{\alpha}_j \equiv \mathbb{E}[\alpha_i \mid j(i) = j, \exists i' \text{ s.t } (i, i' \in \mathcal{P})]$. Using this notation, define the within-firm demeaned individual effect as $\tilde{\alpha}_{ij} \equiv \alpha_i - \bar{\alpha}_j$. Hence,

$$\mathbb{E}[\tilde{\alpha}_{ij} \mid j(i) = 1, \exists i' \text{ s.t } (i, i' \in \mathcal{P})] = 0. \tag{A1}$$

Now:

$$\begin{aligned}\text{Var}(\alpha_i - \alpha_{i'} \mid (i, i') \in \mathcal{P}) &= \text{Var}(\tilde{\alpha}_{ij} - \tilde{\alpha}_{i'j} \mid (i, i') \in \mathcal{P}) &\text{(A2)}\\ &= \text{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P}) + \text{Var}(\tilde{\alpha}_{i'j} \mid \mathcal{P}) - 2Cov(\tilde{\alpha}_{ij}, \tilde{\alpha}_{i'j} \mid \mathcal{P}). &\text{(A3)}\end{aligned}$$

By construction (because we match workers within a firm randomly) $Cov(\tilde{\alpha}_{ij}, \tilde{\alpha}_{i'j} \mid \mathcal{P}) = 0$. Hence,

$$\text{Var}(\alpha_i - \alpha_{i'} \mid (i, i') \in \mathcal{P}) = 2\,\text{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P}). \tag{A4}$$

Step 2: write the regression coefficient in the demeaned notation: Under exogenous mobility,

$$\text{plim}\,\hat{\beta}_1 = \frac{\text{Var}(\alpha_i - \alpha_{i'} \mid (i, i') \in \mathcal{P})}{\text{Var}(\alpha_i - \alpha_{i'} \mid (i, i') \in \mathcal{P}) + \text{Var}(\mu_{ij} - \mu_{i'j} \mid (i, i') \in \mathcal{P})}. \tag{A5}$$

From above:

$$\begin{aligned}\text{Var}(\alpha_i - \alpha_{i'} \mid (i, i') \in \mathcal{P}) &= 2\,\text{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P}) &\text{(A6)}\\ \text{Var}(\mu_{ij} - \mu_{i'j} \mid (i, i') \in \mathcal{P}) &= 2\,\text{Var}(\tilde{\mu}_{ij} \mid \mathcal{P}). &\text{(A7)}\end{aligned}$$

Hence:

$$\operatorname{plim}\hat{\beta}_1 = \frac{Var(\tilde{\alpha}_{ij} \mid \mathcal{P})}{\operatorname{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P}) + \operatorname{Var}(\tilde{\mu}_{ij} \mid \mathcal{P})}. \tag{A8}$$

Step 3: write the variance of the gap in the demeaned notation: Define

$$Z \equiv (\alpha_i - \alpha_{i'}) + (\mu_{ij} - \mu_{i'j}) + (\tilde{\epsilon}_{ij} - \tilde{\epsilon}_{i'j}). \tag{A9}$$

Under exogenous mobility these three terms are not correlated. Then:

$$\operatorname{Var}(Z) = \operatorname{Var}(\alpha_i - \alpha_{i'} \mid (i,i') \in \mathcal{P}) + \operatorname{Var}(\mu_{ij} - \mu_{i'j} \mid (i,i') \in \mathcal{P}) + \operatorname{Var}(\tilde{\epsilon}_{ij} - \tilde{\epsilon}_{i'j} \mid (i,i') \in \mathcal{P}) \tag{A10}$$

$$= 2\operatorname{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P}) + 2\operatorname{Var}(\tilde{\mu}_{ij} \mid \mathcal{P}) + 2\operatorname{Var}(\tilde{\epsilon}_{ij} \mid \mathcal{P}). \tag{A11}$$

Let Γ_Z be the signal to noise ratio in $Var(Z)$ such that

$$\operatorname{Var}(Z)\frac{\Gamma_Z}{2} = \operatorname{Var}(\tilde{\alpha}_{ij} \mid \mathcal{P}) + \operatorname{Var}(\tilde{\mu}_{ij} \mid \mathcal{P}). \tag{A12}$$

Step 4: combine moments to generate an estimator: Now:

$$1 - \hat{\beta}_1 = \frac{Var(\tilde{\mu}_{ij} \mid \mathcal{P})}{\operatorname{Var}(Z)\frac{\Gamma_Z}{2}} \tag{A13}$$

$$Var(\tilde{\mu}_{ij} \mid \mathcal{P}) = (1 - \hat{\beta}_1)\operatorname{Var}(Z)\frac{\Gamma_Z}{2}. \tag{A14}$$

Step 5: an estimator for the signal to noise ratio: The simplest estimator for the signal to noise ratiom Γ_Z relies on the strong assumption that the noise in Z is independent over time and the variance is stationary. Then:

$$\Gamma_Z = \frac{\operatorname{Cov}(Z_{t-1}, Z_t)}{\operatorname{Var}(Z_t)}. \tag{A15}$$

B Pairing workers randomly *vs.* sorting

In the baseline paired movers regression, we pair workers and compute their earnings gap randomly, i.e., without sorting which worker is the highest earner in the origin firm. One alternative implementation is to sort paired workers so that the gap in earnings at the origin firm is always positive. Appendix Figure A3 presents Monte Carlo evidence that sorting workers creates a mechanical upward bias in the paired movers coefficient when the income process of the two workers have different means, which would be the case in the empirically realistic setting where there are differences in individual effects.

In particular, we simulate the income process of two workers by drawing two correlated random variables, one for the origin firm and one for the destination firm. We then add an independent mean-zero shock to each and create a lagged variable for the origin firm using an AR(1) process. We either leave the order of workers random, or sort them so that the gap in earnings at the origin firm is positive and compute the paired movers coefficient using the same IV regression as in the baseline. We repeat this exercise 1,000 times.

The figure shows that when the two workers have the same mean, sorting does not create any bias. However, when there is a difference in the means, sorting creates a mechanical upward bias in the paired movers coefficient. Therefore, we prefer the baseline implementation of the paired movers test, which does not sort workers.

C Forward, Backward, and Donut Paired Movers

In the baseline implementation, we regress the gap in earnings at the destination firm at $t = 1$ (first year at the destination) on the gap in earnings at the origin firm at $t = 0$ (last year at the origin), instrumenting with the lagged gap in earnings at the origin firm at $t = -1$. However, we can use any three periods in which we observe two workers in two firms. These three periods do not need to be consecutive, and this regression can be implemented in either direction (regressing the gap at the destination on the gap at the origin, or vice-versa).

Panel A of Appendix Table A14 shows results from the baseline, or forward, version of the regression, where the outcome is the gap in the destination firm. Column 1 replicates the baseline results from Column 1 of Table 4 – where the outcome is the gap in earnings at the destination firm at $t = 1$ and the endogenous variable is the gap in earnings at the origin firm at $t = 0$, instrumented with the lagged gap in earnings at the origin firm at $t = -1$.

Column 2 shows results when we keep the outcome the same but change the right hand side variable. Instead of the last year at the origin firm ($t = 0$), we use the gap in earnings at the origin firm at $t = -1$ (second to last period at the origin firm) instrumented with the lagged gap in earnings at the origin firm at $t = -2$ (third to last period at the origin firm). Column 3 shows results when we exclude the first year at the destination ($t = 1$) from the sample, using as the outcome the gap in earnings at the destination firm in the second year, $t = 2$, and as the regressor the gap in earnings at the origin firm at $t = 0$, instrumented with the lagged gap in earnings at the origin firm at $t = -1$.

Finally, Column 4 shows results from a donut version of the test, combining Columns 2 and 3 so that we omit the last year at the origin firm and the first year at the destination firm. Columns 5 to 8 replicate Columns 1 to 4 within a fixed sample of movers observed in all periods. The results are very similar across all specifications, with coefficients ranging from 0.73 to 0.75.

Panel B of Appendix Table A14 shows results from the backward version of the paired movers test, where the outcome is the earnings gap in the origin firm, and the independent variable is the earnings at the destination. Across specifications, the coefficient is larger than the baseline when the right hand side variable is the gap in earnings in the first year at the destination—around 0.85—and closer to the baseline when the right hand side variable is the gap in earnings in the second year at the destination—around 0.70. This pattern is consistent with earnings in the first year at the destination being noisier, which can also be seen in Table 5. Overall, the results suggest that our findings are not driven by the specific way we implement the paired movers regression.

D Alternative instruments for the gap

In the main text, we instrument for the gap using lag(s) of the earnings gap. Recall that the gap between i and i' in period t is given by:

$$Z_{ii't} = \underbrace{\alpha_i - \alpha_{i'}}_{\text{person}} + \underbrace{\mu_{ij} - \mu_{i'j}}_{\text{match}} + \underbrace{\tilde{\epsilon}_{ijt} - \tilde{\epsilon}_{i'jt}}_{\text{error}}, \tag{A16}$$

where the underbraces indicate that the gap has three components: the difference in the person effects, the difference in the match effects, and the difference in the error term. Conceptually, in finding instruments we are hoping to isolate the person and match effect component. Under the assumption that the error term is independent across observations, lags of the gap do so.

One alternative would be to instrument for the gap with pre-determined characteristics of a person, such as age, education, or gender. Here we briefly explain why these alternative instruments estimate different parameters than the ones we are interested in. The basic idea is that person characteristics such as education are instruments for the person effect component of the gap, and not also the match effects component. Thus, instrumenting for the gap using education is a test for whether the observational returns to education are the same across employers, rather than about the stability of the earnings gap.

To see this point, decompose the person effect into two parts: $\alpha_i = \alpha_i^s + \alpha_i^r$, where the first superscript is s for schooling and the second superscript is r for random. Suppose that the returns to schooling differ across employers so that the wage equation is:

$$y_{it} = \alpha_i^r + \gamma_{\mathbf{j}(i,t)}\alpha_i^s + \psi_{\mathbf{j}(i,t)} + \beta X_{it} + \epsilon_{it}, \tag{A17}$$

where the change relative to equation (2) is that now the person effect is decomposed into two, and the schooling part has returns that vary across employers (the $\gamma_{\mathbf{j}(i,t)}$ term).

Suppose we observe i and i' at two different employers, j and j' in periods t and t'. Then

$$Z_{ii't} = \alpha_i^r - \alpha_{i'}^r + \gamma_{\mathbf{j}(i,t)}(\alpha_i^s - \alpha_{i'}^s) + \mu_{ij} - \mu_{i'j} + \tilde{\epsilon}_{ijt} - \tilde{\epsilon}_{i'jt} \tag{A18}$$

$$Z_{ii't'} = \alpha_i^r - \alpha_{i'}^r + \gamma_{\mathbf{j}(i,t')}(\alpha_i^s - \alpha_{i'}^s) + \mu_{ij'} - \mu_{i'j'} + \tilde{\epsilon}_{ij't'} - \tilde{\epsilon}_{i'j't'}. \tag{A19}$$

Now suppose that we instrument for the gap using the difference in schooling, $\alpha_i^s - \alpha_{i'}^s$. Suppose that this component is independent of all other terms in the gap. The Wald estimator is then:

$$\frac{Cov(Z_{ii't}, \alpha_i^s - \alpha_{i'}^s)}{Cov(Z_{ii't'}, \alpha_i^s - \alpha_{i'}^s)} = \frac{\mathbb{E}[\gamma_{\mathbf{j}(i,t)}]}{\mathbb{E}[\gamma_{\mathbf{j}(i,t')}]}, \tag{A20}$$

That is, this instrument is just testing whether the returns to a component of the person effect is constant across employers, and is not a function of the match effects.

E Heuristic identification

Appendix Figure A4 shows heuristic identification plots, which explains how model parameters map into the paired mover coefficient. Under exogenous mobility, we showed that the paired mover coefficient is directly proportional to the variance of match effects. In panel (a) of the Figure we hold fixed our estimated parameters, vary the variance of the match effects and plot the resulting paired movers coefficient. The paired movers coefficient is decreasing in the variance of the match effects. Panel (b) shows that all else equal the paired mover coefficient is increasing in the variance of worker effects. Holding fixed the variance of match effects, the larger is the variance of person effects the more stable are earnings gaps across firms.

Because the model does not generate any sorting on the basis of person effects, all of the covariance terms between person effects and match effects are zero. So the only term that matters is $Cov(m_{iA} - m_{i'A}, m_{iB} - m_{i'B})$, the covariance in match effects gaps between workers i and i' across firms. This covariance depends on how long the job ladder is.

The "length" of the job ladder depends on how many job-to-job transitions a worker makes before re-entering unemployment. It is increasing in λ_1 (the arrival rate of offers on the job) and decreasing in δ (the job destruction rate). Panels (c) and (d) show that these two parameters move the implied paired mover coefficient in opposite directions. As δ increases and workers have fewer jobs (all else equal) the paired mover coefficient declines. The reason is that with fewer jobs the match effects distribution is unselected and so the covariance term of the match effects across firms is smaller. Similarly, as λ_1 increases and workers have more jobs the paired mover coefficient increases because the covariance of match effects across jobs increases. In contrast, the job finding rate when unemployed (λ_0) has no impact on the length of the job ladder and so varying λ_0 has no impact on the estimated paired mover coefficient.

Table A1: Bias in Mobility Simulations

	Observations (1)	% of Initial Dataset (2)
A. Brazil		
Paired Movers dataset	1,981,101	100
Baseline simulation	825,704	41.7
Including paired movers	1,264,999	63.9
B. Veneto		
Paired Movers dataset	8,554	100
Baseline simulation	3,694	43.2
Including paired movers	6,129	71.7

This table shows a quantification exercise of the downward bias in the simulations presented in Table 2. Since we do not use paired movers to construct the empirical distribution of between-firm mobility, the simulated data do not have the same number of paired movers as the data. This table shows the number of paired movers in the estimation sample, of a baseline simulation from Table 2 and a simulation that includes paired movers in the construction of the empirical distribution of between-firm mobility. Panel A shows results for Brazil, while Panel B shows results for Veneto.

Table A2: Distribution of # of Workers per Origin-Destination Pair

	Min (1)	p5 (2)	p10 (3)	p25 (4)	p50 (5)	p75 (6)	p90 (7)	p95 (8)	Max (9)
All movers	1	1	2	2	2	3	6	9	1,507
Within paired movers sample	2	2	2	2	2	4	6	9	1,507

This table shows the distribution of the number of eligible workers per origin-destination pair. Eligible workers are those that are observed for at least two years at the origin firm. The table presents summary statistics of the distribution of the number of eligible workers per origin-destination pair, weighted by the number of workers across all eligible movers (first row) and among paired movers in each pair (second row). A destination-origin pair is defined by origin and destination firms and a year.

Table A3: Paired Movers Test, Robustness to Residualization Approach

	Residualization		Matching by age and gender	
	Baseline (CHK) (1)	Without education (2)	1-year age bin (3)	10-year age bin (4)
A. Brazil, Second Stage				
Gap in Origin	0.731	0.723	0.742	0.700
	(0.0007)	(0.0007)	(0.001)	(0.0009)
Observations	1,981,101	1,981,101	549,834	1,408,531
B. Brazil, 1st Stage				
Gap in Origin (lag)	0.916	0.913	0.920	0.915
	(0.0005)	(0.0005)	(0.001)	(0.0007)
F-statistic	3,419,903.6	3,248,141.8	822,427.0	1,925,253.4

Column 1 shows the baseline results where we residualize log earnings on a cubic in (normalized) age and a full set of sex and education groups interactions, omitting the linear term in age. Column 2 shows the results when we do not include education. Columns 3 and 4 show results when we match paired movers on age and gender. In Column 3 we use one-year age bins, while in Column 4 we use ten-year age bins.

Table A4: Paired Movers Test, Reweighting for the Firm Size Distribution

		Reweighted ... to match overall firm size distribution	
	Baseline (1)	Origin (2)	Destination (3)
A. Second Stage			
Gap in Origin	0.731	0.747	0.671
	(0.0007)	(0.0008)	(0.0008)
Observations	1,981,101	1,981,101	1,981,101
B. Brazil, 1st Stage			
Gap in Origin (lag)	0.916	0.894	0.874
	(0.0005)	(0.0005)	(0.0005)
F-statistic	3,419,903.6	3,231,317.7	3,167,361.9

This table shows results from a robustness exercise in which we reweight the sample of paired movers to match the overall firm size distribution in the data. We first divide the firm size distribution in the full data (Column 1, Table 1) into 100 bins with the same number of firms. We then reweight the sample of paired movers so that the distribution of firm sizes at origin (Column 2) and destination firms (Column 3) matches the overall distribution in the data, within the shared support of firm sizes in the paired movers sample. In practice, the first 27 bins of the firm size distribution are not in the support of the paired movers sample, and more than 95% of the firms in the paired movers sample are in the top 10% of the firm size distribution (larger than 15 employees).

Table A5: Earnings Gap AR(1) Coefficients

	Full sample (1)	Destination (2)	Origin (3)
A. Balanced Panel			
k	-0.048	-0.061	-0.042
	(0.002)	(0.004)	(0.002)
AR(1) Coefficient	0.953	0.941	0.959
B. Full Sample			
k	-0.036	-0.039	-0.036
	(0.004)	(0.001)	(0.006)
AR(1) Coefficient	0.964	0.962	0.965
Observations	25	10	15

This table shows coefficients from regression $\ln(\rho_k) = \beta k + \varepsilon_k$, where ρ_k is the autocorrelation of the log earnings gap between paired movers at time lag/lead k. Hence, $\exp(\beta)$ is the AR(1) coefficient of the log earnings gap. The autocorrelations are the same as those shown in Table 5. Panel A shows the results for the balanced panel, while Panel B shows the results for the full sample. Column 1 shows the results using all available autocorrelations, while Columns 2 and 3 show the results using only the autocorrelations at destination and origin firms. Regressions are weighted by the number of observations used to calculate each autocorrelation. Standard errors are in parentheses.

Table A6: $\mathrm{Var}(\tilde{\mu}_{ij})$ by decile of the absolute earnings gap

Decile	Mean Gap (1)	β (2)	$\mathrm{Var}(Z)$ (3)	$\mathrm{Cov}(Z_t, Z_{t-1})$ (4)	Γ_Z (5)	$\mathrm{Var}(\tilde{\mu}_{ij} \mid \mathcal{P})$ (6)
5	0.220	0.973	0.049	0.040	0.809	0.001
6	0.293	0.844	0.086	0.070	0.811	0.005
7	0.384	0.771	0.148	0.119	0.805	0.014
8	0.509	0.732	0.261	0.207	0.792	0.028
9	0.717	0.712	0.521	0.398	0.763	0.057
10	1.350	0.689	2.057	1.320	0.642	0.205

This Table shows the underlying numbers for Figure 4. Column 1 shows the mean earnings gap at origin firms for each decile of the gap distribution (deciles 5 to 10). Column 2 shows the paired movers test coefficient β estimated separately for each decile. Column 3 shows the variance of the earnings gap. Column 4 shows the covariance of the earnings gap across periods at the origin firm. Column 5 shows the implied signal to noise ratio. Finally, Column 5 shows the variance of match effects implied by the estimates in Columns 1-5 using equation (12).

Table A7: Summary statistics: paired movers

	All pairs (1)	Sector stayers (2)	Sector movers (3)	Cluster stayers (4)	Cluster movers (5)	Low-Low (6)	Low-High (7)	High-High (8)	High-Low (9)
A. Brazil									
Avg. age	32.37	33.02	31.84	32.44	31.97	32.21	31.25	32.72	33.01
Share female	0.31	0.34	0.29	0.32	0.31	0.38	0.31	0.25	0.27
Years of schooling	10.90	10.89	10.91	11.35	10.88	10.16	10.81	11.83	10.48
Avg. Tenure (months)	38.13	39.49	37.00	41.10	38.16	33.59	34.75	43.37	39.62
Median Tenure (months)	26.90	28.03	26.03	29.23	27.03	24.50	25.50	30.37	27.03
Avg. log earnings	2.39	2.41	2.37	2.60	2.37	1.99	2.18	2.86	2.37
Avg. log gap origin	0.35	0.36	0.35	0.40	0.35	0.27	0.31	0.46	0.36
Var. log gap origin	0.18	0.19	0.17	0.21	0.17	0.11	0.14	0.24	0.17
p90 - p10 log gap origin	0.85	0.88	0.83	0.96	0.84	0.65	0.74	1.05	0.85
% EE move	0.14	0.14	0.14	0.16	0.14	0.12	0.17	0.17	0.10
% different gender	0.21	0.23	0.19	0.22	0.21	0.23	0.20	0.19	0.19
% Higher earner first	0.45	0.46	0.44	0.47	0.46	0.42	0.45	0.48	0.45
Mean time btw moves	120.93	132.20	111.52	140.23	124.15	105.89	114.18	140.50	115.41
# of Worker-years	3,962,202	1,801,972	2,160,230	870,112	1,843,144	1,447,000	534,214	1,446,752	534,236
Avg. origin firm size	78.83	98.83	107.69	178.38	115.14	68.62	116.91	144.38	174.18
Avg. origin firm pay	2.47	2.48	2.49	2.56	2.49	2.09	2.14	2.75	2.61
Avg. destination firm size	77.07	95.05	106.17	172.11	111.70	67.76	181.05	138.09	111.15
Avg. destination firm pay	2.49	2.49	2.52	2.57	2.52	2.10	2.70	2.79	2.16
# of Firm-Years	1,579,732	917,036	939,545	406,417	795,772	776,634	313,316	505,651	315,713
B. Veneto									
Avg. age	35.39	36.26	34.33	34.85	34.89	36.08	33.95	35.41	35.30
Share female	0.35	0.33	0.37	0.36	0.35	0.58	0.34	0.16	0.28
Avg. log earnings	10.20	10.26	10.14	10.22	10.22	9.76	10.05	10.63	10.40
Avg. log gap origin	0.41	0.42	0.39	0.41	0.44	0.44	0.39	0.39	0.40
Var. log gap origin	0.23	0.24	0.21	0.22	0.24	0.31	0.21	0.16	0.20
p90 - p10 log gap origin	0.92	0.93	0.89	0.84	0.96	1.02	0.88	0.85	0.89
# of Worker-years	17,108	9,418	7,690	1,096	8,676	5,868	2,670	5,900	2,670
Avg. origin firm size	221.66	253.99	235.63	287.25	287.69	166.01	162.39	319.21	307.30
Avg. origin firm pay	10.17	10.19	10.15	10.11	10.16	9.53	9.65	10.52	10.37
Avg. destination firm size	238.80	266.18	283.15	275.31	298.16	196.43	396.18	328.88	217.93
Avg. destination firm pay	10.25	10.24	10.26	10.16	10.22	9.64	10.50	10.59	9.86
# of Firm-years	11,854	7,031	6,049	1,029	6,204	4,438	2,265	4,310	2,285

This table shows descriptive statistics of the paired movers sample. Panel A displays the Brazilian data, while Panel B displays the Veneto data. Column 1 displays all paired movers and uses the same samples as Column 3 in Table 1. Columns 2 and 3 display movers split by whether they stay in the same sector or move to a different sector (2-digit CNAE code for Brazil and 2-digit ATECO 91 code for Veneto). Columns 4 and 5 display movers split by whether they stay in the same cluster or move to a different cluster. Clusters are estimated using the entire data (Column 1, Table 1) following Bonhomme, Lamadon, and Manresa (2019). Clusters are only estimated in the connected set of firms. Columns 6 to 9 display movers across different pay bins. We define low/high pay as the bottom/top 50% of the firm pay distribution among firms in the sample. Firm pay is computed yearly across all workers, not only paired movers.

Table A7: Summary statistics: paired movers (cont.)

	No ocup. change (10)	1 ocup. change (11)	2 ocup. change (12)	Both Low Tenure (13)	Both High Tenure (14)	One Low Tenure (15)	CZ stayers (16)	CZ switchers (17)
C. Brazil, cont.								
Avg. age	33.98	32.59	30.51	31.41	34.22	31.92	32.00	33.56
Share female	0.31	0.30	0.31	0.27	0.33	0.34	0.34	0.21
Years of schooling	10.54	11.09	11.13	10.41	11.34	11.11	11.04	10.44
Avg. Tenure (months)	36.84	39.47	38.49	16.88	67.98	37.87	39.72	33.13
Median Tenure (months)	26.43	27.63	26.90	16.70	54.03	30.02	28.23	23.20
Avg. log earnings	2.37	2.47	2.35	2.26	2.57	2.38	2.36	2.48
Avg. log gap origin	0.30	0.45	0.34	0.32	0.36	0.38	0.35	0.35
Var. log gap origin	0.15	0.21	0.17	0.16	0.18	0.19	0.18	0.17
p90 - p10 log gap origin	0.74	1.00	0.83	0.79	0.88	0.91	0.85	0.87
EE move	0.13	0.13	0.16	0.15	0.11	0.16	0.15	0.12
% different gender	0.18	0.25	0.21	0.18	0.22	0.23	0.23	0.15
% Higher earner first	0.43	0.48	0.45	0.44	0.45	0.46	0.46	0.43
Mean time btw moves	122.79	127.17	114.22	103.60	136.23	128.86	128.36	97.43
# of Worker-years	1,494,252	1,063,542	1,404,408	1,574,964	1,132,260	1,254,978	3,009,938	952,264
Avg. origin firm size	111.26	125.47	123.88	108.68	130.33	120.18	82.74	175.39
Avg. origin firm pay	2.48	2.50	2.50	2.45	2.53	2.49	2.47	2.52
Avg. destination firm size	108.98	121.96	120.96	111.38	122.09	118.94	80.83	172.17
Avg. destination firm pay	2.51	2.53	2.53	2.49	2.55	2.52	2.49	2.55
# of Firm-Years	763,078	637,657	706,074	779,300	649,875	697,137	1,382,737	394,409

This table shows descriptive statistics of the Brazilian paired movers sample. Columns 10 to 12 display movers split by whether they change occupation or not (3-digit CBO codes). Column 10 shows moves in which both workers in the pair remain in the same occupation, Column 11 shows moves in which one worker changes occupation and Column 12 shows moves in which both workers change occupation. Columns 13 to 15 display movers split by whether they have low or high tenure at the origin firm. We define low tenure as tenure below the median tenure of all paired movers (30 months). Column 13 shows pairs where both workers have low tenure, Column 14 shows pairs where one worker has low tenure and Column 15 shows pairs where both workers have high tenure. Finally, Columns 16 and 17 display movers split by whether they stay in the same commuting zone (CZ) or not. There are 510 CZs, defined as the "immediate regions" designated by the Brazilian Institute of Geography and Statistics (IBGE), which have a similar logic to the US commuting zones.

Table A8: The Paired Mover Coefficient in Brazilian Data: Additional Splits

	Firm Pay				Change Ocup.			High Tenure			CZ	
	LL (1)	LH (2)	HH (3)	HL (4)	None (5)	One (6)	Both (7)	None (8)	One (9)	Both (10)	Stayers (11)	Switchers (12)
A. 2nd Stage												
Gap in Origin	0.740	0.728	0.747	0.651	0.781	0.748	0.663	0.741	0.719	0.734	0.734	0.723
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0008)	(0.001)
p-value, equality test	0.0000		0.0000			0.0000			0.0000		0.0000	
$\text{Var}(Z)$	0.209	0.254	0.447	0.316	0.263	0.416	0.299	0.284	0.347	0.328	0.318	0.311
Γ_Z	0.663	0.682	0.717	0.650	0.672	0.742	0.655	0.627	0.715	0.741	0.695	0.678
$\text{Var}(\tilde{\mu}_{ij} \mid \mathcal{P})$	0.018	0.024	0.041	0.036	0.019	0.039	0.033	0.023	0.035	0.032	0.029	0.029
Observations	723,500	267,107	723,376	267,118	747,126	531,771	702,204	787,482	627,489	566,130	1,504,969	476,132
B. 1st Stage												
Gap in Origin (lag)	0.901	0.903	0.927	0.912	0.898	0.937	0.910	0.896	0.919	0.934	0.919	0.907
	(0.0009)	(0.001)	(0.0008)	(0.001)	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0008)	(0.0008)	(0.0006)	(0.001)
F-statistic	1,075,134.2	428,423.4	1,430,678.1	388,663.6	1,136,649.3	1,214,484.9	1,036,385.7	1,009,861.8	1,200,476.1	1,273,019.9	2,662,795.2	759,954.3

Columns 1 to 4 show the estimate using the sample of paired movers split by whether origin and destination firms have low or high pay (defined as below or above the median firm pay in the sample, respectively). Columns 5 to 7 show the estimate using the sample of paired movers split by whether none, one or both workers change occupation (3-digit CBO code). Columns 8 to 10 show the estimate using the sample of paired movers split by whether both workers have low tenure (less than 30 months), one has low tenure and the other has high tenure, or both have high tenure. Columns 11 and 12 show the estimate using the sample of paired movers split by whether they stay in the same commuting zone (CZ) or move to a different CZ. The CZs are defined as IBGE's immediate geographic regions (Regiões Geográficas Imediatas).

Table A9: Paired Mover Coefficient by Earnings of the First Mover

	Gap in Destination			
	All paired movers		Gap higher than 0.2 log points	
Higher earner first	No (1)	Yes (2)	No (3)	Yes (4)
A. Second Stage				
Gap in Origin	0.708	0.750	0.684	0.734
	(0.001)	(0.001)	(0.001)	(0.002)
$\text{Var}(Z)$	0.259	0.387	0.460	0.620
Γ_Z	0.696	0.687	0.694	0.686
$\text{Var}(\tilde{\mu}_{ij} \mid \mathcal{P})$	0.026	0.033	0.050	0.057
Observations	1,090,494	890,607	601,575	549,306
B. First Stage				
Gap in Origin (lag)	0.893	0.936	0.955	0.983
	(0.0007)	(0.0007)	(0.0009)	(0.0009)
F-statistic	1,789,919.4	1,607,552.1	1,182,693.7	1,134,672.5

Columns 1 and 2 show the results for all paired movers, while Columns 3 and 4 show the results for paired movers where the gap in earnings is higher than 0.2 log points. Columns 1 and 3 show the results for pairs where the higher earner in the origin firm is the second mover, while Columns 2 and 4 show the results for pairs where the higher earner in the origin firm is the first mover. We also report components of the exogenous mobility estimator for the variance of match following, following Equation 12.

Table A10: Multi-sector Model results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	EE rate	β_{Stayer}	β_{Switcher}	Var(Z_{Stayer})$\times\Gamma_{Z_{\text{Stayer}}}$	Var(Z_{Switcher})$\times\Gamma_{Z_{\text{Switcher}}}$	Var(ψ_j)	Var(α_i) (untargeted)
Empirical Moments (Data)	0.038	0.756	0.706	0.240	0.201	0.065	0.294
Simulated Moments (Model)	0.040	0.827	0.513	0.243	0.215	0.064	0.049
	λ_0 (UE rate)	δ (EU rate)					
Fixed Parameters	0.072	0.076					
	λ_1	$Var(m_{ij})$	m_{is}	$Var(h_i)$	$Var(p_j)$		
Calibrated Parameters	0.090	0.020	0.115	0.068	0.050		
Mobility (EE)	% Offers Accepted	% Offers Cross-Sector	% Within-Sector Offers Accepted	% Cross-Sector Offers Accepted	% Moves Cross-Sector		
	0.251	0.969	0.508	0.243	0.937		

This table shows the results of the multi-sector model estimation. We estimate the model similarly to the baseline model (Table 6), but now allowing for sector-specific match effects m_{is}, in addition to firm-specific match effects m_{ij}. We simulate 200 firms across 20 sectors, and sector-specific match effects follow a generalized Pareto distribution with shape parameter $k = 0.25$ and scale parameter σ. We use the same moments as in Table 6, but now we split the paired movers' coefficients and gap variances by whether the move is within the same sector or across sectors. We also report EE mobility patterns: the percentage of offers that are accepted, the percentage of offers that are from a different sector, the percentage of accepted offers that are within the same sector, the percentage of accepted offers that are from a different sector, and the percentage of moves that are across sectors.

Table A11: Event Study Test in Model Simulated Data

	All movers (1)	EE moves (2)	EUE moves (3)
A. Brazil			
Change in $\hat{\psi}$ (IV)	0.990	0.570	0.987
	(0.0007)	(0.001)	(0.001)
Observations	982,046	393,711	588,335

This table shows the event study test in data simulated from the estimated model (Table 6). We estimate firm effects in two randomly split samples of the data and run a 2SLS regression as in Table A12. Column 1 shows the results for all movers, while Columns 2 and 3 show the results for EE and EUE moves, respectively. A graphical version of this table is shown in Figure 6b.

Table A12: Event Study Test

	Change in $\hat{\psi}_1$ / Change in Log Earnings				
	All movers	Paired movers	Weight (1) to match ... in Paired Movers:		
			Origin Firms	Destination Firms	Origin-Destination Pairs
	(1)	(2)	(3)	(4)	(5)
A. Brazil, 2nd Stage					
Change in $\hat{\psi}_1$	1.09	1.07	0.994	0.995	1.02
	(0.0010)	(0.003)	(0.0010)	(0.0010)	(0.002)
Observations	4,457,623	427,166	3,825,306	3,879,875	1,066,222
B. Brazil, 1st Stage					
Change in $\hat{\psi}_2$	0.746	0.812	0.880	0.882	0.889
	(0.0003)	(0.0009)	(0.0002)	(0.0002)	(0.0004)
F-statistic	5,535,025.1	831,373.4	13,059,664.2	13,253,291.6	3,928,520.8
C. Veneto, 2nd Stage					
Change in $\hat{\psi}_1$	1.18	0.663	0.976	0.967	0.720
	(0.007)	(0.045)	(0.008)	(0.008)	(0.027)
Observations	281,075	3,829	110,305	130,334	11,294
D. Veneto, 1st Stage					
Change in $\hat{\psi}_2$	0.500	0.740	0.755	0.717	0.771
	(0.002)	(0.011)	(0.002)	(0.002)	(0.006)
F-statistic	95,639.4	4,279.4	143,340.4	127,550.6	16,263.4

This table shows the results of the standard event study test using AKM estimates. We use a sample of movers to estimate a 2SLS regression where the change in the log earnings between destination and origin firms is regressed on the change in the estimated firm effect across origin-destination firms, where we use the firm effects estimated in one sample to instrument the firm effects in the second sample. We estimate two sets of firm fixed effects by randomly splitting the full data (Column 1, Table 1) into two samples. Column 1 shows the results for all movers used in the AKM estimation. Column 2 subsets the sample for movers in the final paired movers sample (Column 3, Table 1). Columns 3 to 5 show the results the sample in Column 1 when we weight the observations to match the distribution of origin firms, destination firms, and origin-destination pairs in the paired movers sample. Panel A shows the second stage of the test in Brazilian data, while Panel C shows the second stage in Veneto data. The first stage is shown in Panels B and D.

Table A13: Simulated Borovičková and Shimer (2024) data

	Paired Mover	Event Study		
		All movers	EE moves	EUE moves
	(1)	(2)	(3)	(4)
Gap in origin	0.622			
	(0.002)			
Change in $\hat{\psi}$ (IV)		0.958	0.347	0.950
		(0.002)	(0.002)	(0.003)
Observations	180,560	2,516,273	1,338,002	1,178,271
Gap Variance	0.017			
Var($\tilde{\mu}_{ij}$) – Exog. mobility	0.003			

This table shows the results using data simulated from the model of Borovičková and Shimer (2024), with the same parameters as estimated in the original paper. We report the paired movers and event study tests (split by EE and EUE moves), as well as the absolute mean gap and the variance of match effects implied by the paired movers test under exogenous mobility (Equation 12), without a correction for measurement error ($\Gamma_Z = 1$).

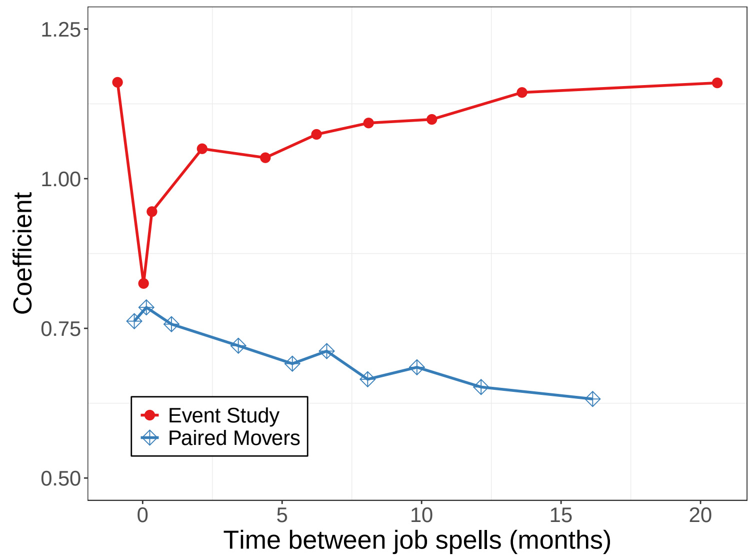
Table A14: Forward and Backward Paired Mover Test

A. "Forward" Dependent Var.:	Gap in Destination, $t = 1$		Gap in Destination, $t = 2$		Gap in Destination, $t = 1$		Gap in Destination, $t = 2$	
	Baseline (1)	Excl. $t = 0$ (2)	Excl. $t = 1$ (3)	"Donut" (4)	Baseline (5)	Excl. $t = 0$ (6)	Excl. $t = 1$ (7)	"Donut" (8)
A1. 2^{nd} Stage								
Gap in Origin, $t = 0$	0.731		0.728		0.746		0.733	
	(0.0007)		(0.0009)		(0.0009)		(0.001)	
Gap in Origin, $t = -1$		0.731		0.728		0.739		0.728
		(0.0008)		(0.001)		(0.0009)		(0.001)
Observations	1,981,101	1,334,489	1,365,736	924,117	924,117	924,117	924,117	924,117
A2. 1^{st} Stage								
Gap in Origin, $t = -1$	0.916		0.921		0.926		0.926	
	(0.0005)		(0.0006)		(0.0007)		(0.0007)	
Gap in Origin, $t = -2$		0.750		0.748		0.748		0.748
		(0.0004)		(0.0005)		(0.0005)		(0.0005)
F-statistic	3,419,903.6	3,089,371.7	2,366,829.6	2,148,765.0	1,832,889.2	2,148,765.0	1,832,889.2	2,148,765.0

B. "Backward" Dependent Var.:	Gap in Origin, $t = 0$		Gap in Origin, $t = -1$		Gap in Origin, $t = 0$		Gap in Origin, $t = -1$	
	Baseline (1)	Excl. $t = 0$ (2)	Excl. $t = 1$ (3)	"Donut" (4)	Baseline (5)	Excl. $t = 0$ (6)	Excl. $t = 1$ (7)	"Donut" (8)
B1. 2^{nd} Stage								
Gap in Destination, $t = 1$	0.849		0.841		0.853		0.848	
	(0.001)		(0.0009)		(0.001)		(0.001)	
Gap in Destination, $t = 2$		0.710		0.704		0.710		0.704
		(0.001)		(0.001)		(0.001)		(0.001)
Observations	1,365,736	987,776	1,365,736	987,776	987,776	987,776	987,776	987,776
B2. 1^{st} Stage								
Gap in Destination, $t = 2$	0.636		0.636		0.666		0.666	
	(0.0005)		(0.0005)		(0.0006)		(0.0006)	
Gap in Destination, $t = 3$		0.696		0.696		0.696		0.696
		(0.0006)		(0.0006)		(0.0006)		(0.0006)
F-statistic	1,471,896.4	1,176,657.8	1,471,896.4	1,176,657.8	1,189,980.3	1,176,657.8	1,189,980.3	1,176,657.8

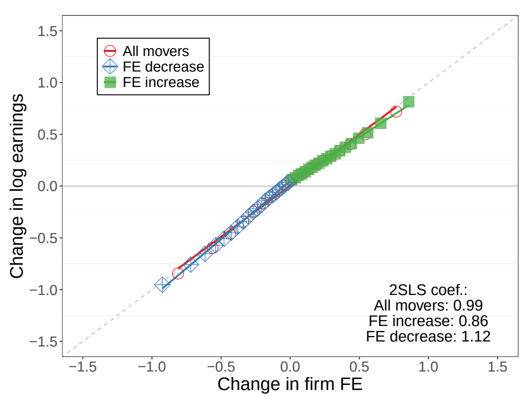
This table shows results from the forward (baseline) and backward versions of the paired movers test. Panel A shows results from the forward version, where the outcome is the earnings gap in the destination firm and the right hand side is the earnings gap at the origin firm. Panel B shows results from the backward version, where the outcome is the earnings gap at the origin firm and the right hand side is the earnings gap at the destination firm. In both panels, the first stage results are shown in the lower half and the second stage results are shown in the upper half. Columns 1 and 5 show baseline results using the gap in earnings at $t = 0$ (Panel A) or $t = 1$ (Panel B) as regressor. Columns 2 and 6 exclude the contemporaneous period from the regression, i.e., they use only the lagged gap in earnings at the origin (Panel A) or destination (Panel B) as regressor. Columns 3 and 7 exclude one period after/before the move from the regression, i.e., they use only the gap in earnings at $t = 2$ (Panel A) or $t = -1$ (Panel B) as dependent variable. Finally, Columns 4 and 8 implement a "donut" version of the paired movers test, where both the contemporaneous period and one period after/before the move are excluded from the regression.

Figure A1: Paired Movers and Event Study tests by time between employment spells

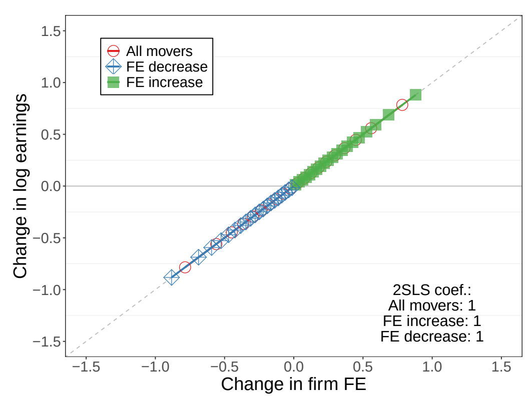


This figure shows the paired mover and event study test results by the time between employment spells, defined as the difference between the hiring date at the destination firm and the separation date at the origin firm. In each sample, we compute deciles of the distribution of time between employment spells and run the paired movers and event study tests separately for each decile. For the paired movers test, we only keep pairs of movers where both workers in the pair are in the same decile of the distribution of time between employment spells.

Figure A2: Event Study Test in the Model: Endogenous vs Exogenous Mobility



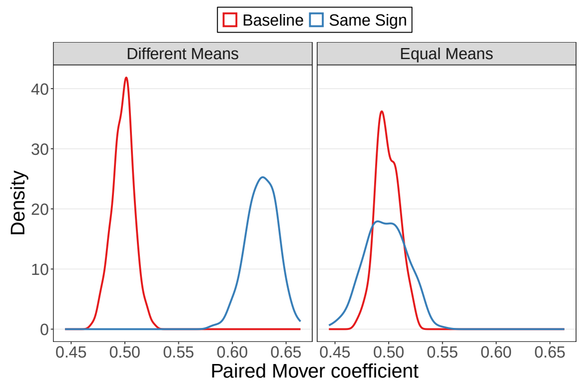
(a) Endogenous Mobility



(b) Exogenous Mobility

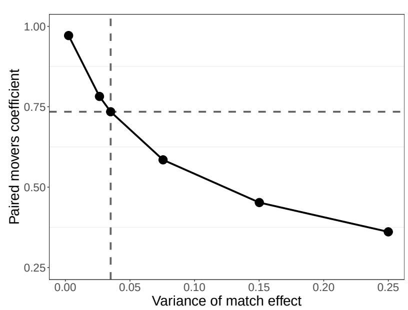
This figure shows the event study test in model simulated data under endogenous (Panel A) and exogenous (Panel B) mobility. In both panels, we estimate firm effects in two randomly split samples of the data and run 2SLS as in Appendix Table A12. We plot the reduced form scatterplot and the 2SLS estimates for all movers, moves with firm FE increases, and moves with firm FE decreases.

Figure A3: Bias in Same Sign Paired Movers Test: Monte Carlo Simulations

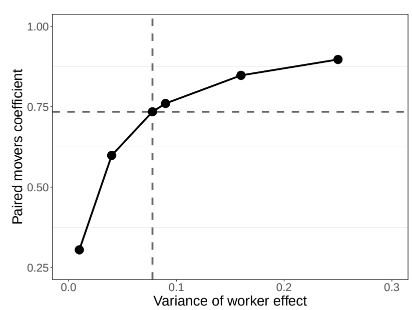


This figure shows the results of 1,000 simulations to assess bias in the same sign paired movers regression. We simulate the income process of two workers by drawing two correlated random variables, one for the origin firm and one for the destination firm and add an independent mean-zero shock to each. We then create a lagged variable for the origin firm using an AR(1) process and either leave the order of workers random, or sort them so that the gap in earnings at the origin firm is positive. We compute the paired movers coefficient using the same IV regression as in the baseline.

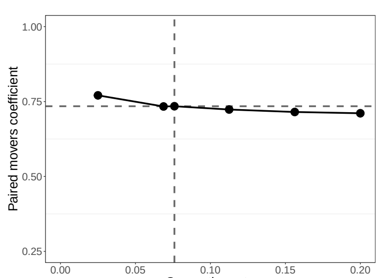
Figure A4: Paired Movers coefficient and model parameters



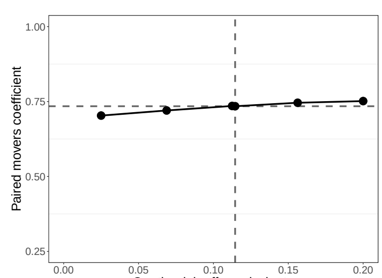
(a) Match Effects ($Var(\mu_{ij})$)



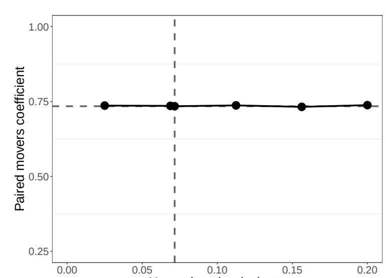
(b) Worker Effects ($Var(h_i)$)



(c) Separation Rate (δ)



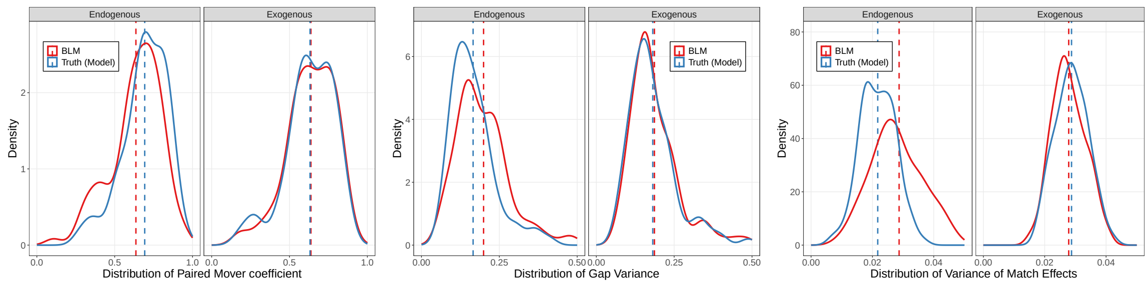
(d) Emp. arrival rate (λ_1)



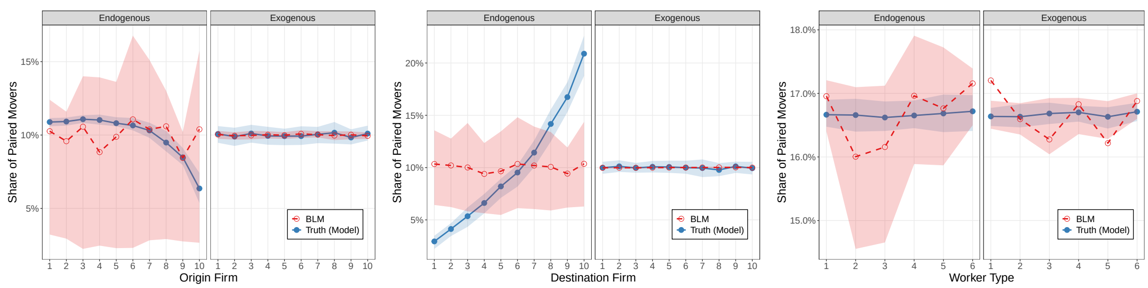
(e) Unemp. arrival rate (λ_0)

This figure shows heuristic identification plots. In each panel we start with the estimated parameter (shown in the vertical dashed line). We then vary that estimated parameter (holding all other model parameters constant) and show how changing that parameter affects the paired movers coefficient.

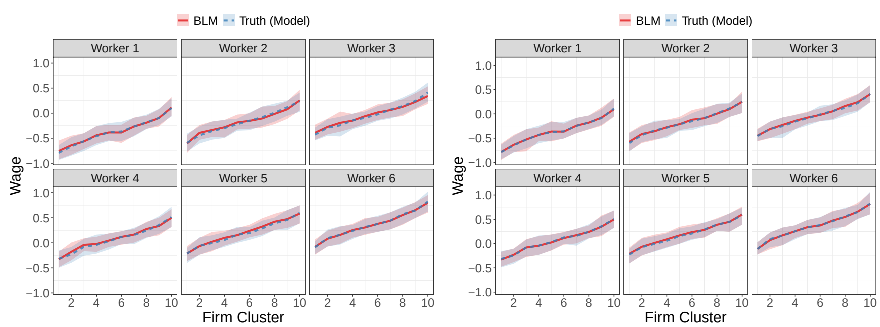
Figure A5: Monte Carlo Simulations: BLM with discrete heterogeneity



(a) Paired Movers Coefficient (b) Gap Variance (c) Variance of Match Effects



(d) Origin Firm (e) Destination Firm (f) Workers



(g) Wages: endogenous mobility (h) Wages: exogenous mobility

This figure shows the distribution of the paired movers test statistics across BLM estimations. We draw 100 simulations each of the endogenous and exogenous versions of the model with discrete heterogeneity (6 worker types and 10 firm types) and estimate the BLM model in each simulation. In Panels A–C, we plot the distribution of the paired movers coefficient (Panel A), the gap variance (Panel B), and the variance of match effects implied by the paired movers test under exogenous mobility (Panel C) in the true data (fed into BLM) and from data simulated from the estimated BLM model. In each panel, the dashed vertical line shows the average value in the true and estimated data across simulations. In Panel D–F, we plot the average distribution of origin firms (Panel D), destination firms (Panel E), and workers (Panel F) across simulations. Panels G and H show the average wage for each worker type and firm type under endogenous and exogenous mobility, respectively, in the true data (simulated from our model) and in the estimated BLM model. In each panel, bands represent p25 and p75 across simulations.