Paired movers: A new test of (log) additive separability

Isaac Sorkin Lucas Warwar *

September 2025

Abstract

Consider a pair of workers who both move from firm A to firm B. After accounting for measurement error, a model of wages with additively separable firm and worker effects implies that the gap in pay between these two workers at firm A is an unbiased forecast of the gap at firm B because the only persistent component of the gap is in the individual effect (the forecast coefficient is 1). We find a forecast coefficient of around 0.70 in Brazil and 0.78 in Italy, which statistically rejects additive separability, thus suggesting the presence of match effects in the wage equation. The forecast coefficient varies in ways consistent with the presence of match effects among similar firms. For example, the coefficient is larger for sector stayers than sector switchers. We estimate a simple job ladder model with endogenous mobility to match the forecast coefficient. Our estimates imply a large role for match effects. Despite this prominent role for match effects and endogenous mobility, our model-generated data passes a specification check which is typically interpreted as evidence for exogenous mobility. We show that the prominent role for match effects implies that attempts to extrapolate earnings patterns from accepted offers (moves we see in the data) to rejected offers (moves we do not see in the data) can be quite misleading.

*Sorkin: Stanford University and NBER, sorkin@stanford.edu. Warwar: Stanford University, warwar@stanford.edu. Thanks to Jediphi Cabal, Michael Dinerstein, Paul Goldsmith-Pinkham and Hamish Low for helpful conversations. 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.

A common research design is the movers design. A key assumption in this design is (log) additive separability. In the context of wages and firms (Abowd, Kramarz, and Margolis (1999) and Card, Heining, and Kline (2013)), this assumption means that we can write the (log) wage as the sum of a person effect, a firm effect, and an independent error term. This equation has been enormously influential and sparked a large body of work on firms in wage-setting (e.g., Card, Cardoso, and Kline (2016), Sorkin (2018), and Gerard et al. (2021)).

This design is also very controversial because it is inconsistent 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) for discussion). Nonetheless, rigorous attempts to explore the validity of this equation have tended to find that it is a reasonable approximation to the data (e.g. Card, Heining, and Kline (2013) and Bonhomme, Lamadon, and Manresa (2019)). Specifically, Card, Heining, and Kline (2013) introduce what we call the event study test and show that for movers across firms the change in the firm effects does a good job of predicting changes in earnings.

In this paper, we revisit this evidence by developing a new test of additive separability. The basis of the test is to consider a pair of workers ("paired movers") who both move from firm A to firm B. Additive separability implies that on average the gap between the earnings of the workers should be determined only by the difference in their person effects and thus that this difference should be stable across firms. We formulate this test as a forecast bias test and document meaningful deviations from additive separability. This deviations are larger for moves across larger contexts, where it is plausible that workers are changing their match quality: sector switchers, larger changes in firm pay, or across firm clusters determined by a k-means algorithm.

We begin by explaining how additive separability, the assumption that this paper is testing, is stronger than exogenous mobility mobility, the assumption needed to interpret a two-way fixed effects estimator as capturing average firm effects. 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 it is clear that additive separability is not satisfied because there are match effects in the wage equation. Thus, finding evidence against additive separability holding does not necessarily mean that exogenous mobility is not satisfied. Nonetheless, below we develop auxilliary evidence that points against exogenous mobility.

We then introduce the paired movers test, which relies on the idea that earnings gaps between workers should be stable across firms. When additive separability holds and the errors are independent, the regression of the gap between the paired movers at firm A on the gap at firm B is just a regression of the difference in worker effects on itself. Even when additive separability holds, however, the OLS estimator is downward biased because of measurement error. Therefore, in our implementation we only use workers where we have two observations of earnings at the origin firm so that we can instrument and correct for measurement error.

We use administrative data from two countries: Italy and Brazil. For the Italian data, we use the Veneto Worker History data and exactly follow the data construction in Kline, Saggio, and

Sølvsten (2020). For the Brazilian data, we use the RAIS data and follow the data construction—except for geographic restrictions—in Gerard et al. (2021). Because the Brazilian dataset is much larger (and less noisy), we focus on the results from Brazil. In the body of the paper, we briefly discuss where results in Veneto differ from those in Brazil.

In terms of descriptive statistics, we focus on how selected the paired movers are. Paired movers make up about 5% of all movers. Paired movers are older, earn more, and have higher tenure than all movers. The most important difference between paired movers and all movers is that paired movers start from—and end up at—much larger firms. We show that given mobility patterns, the number of paired movers that we have in the data is close to what we would expect by chance. So in this sense while the paired movers are a narrow subset of the data, it is roughly as narrow as we would expect based on the distribution of mobility that we observe. Therefore, we think it is plausible to extrapolate from the paired mover sample.

One possible concern is that if the paired movers are very similar then we would find large deviations from additive separability according to the forecast bias test even if these deviations are economically small. We show that the mean earnings gap among the paired movers is about 80% of a randomly selected pair of workers within a firm. So this concern does not seem large. In robustness, we reweight the paired mover sample to match this random benchmark and find stronger evidence of deviations from additive separability. More generally, the paired mover coefficient is decreasing in the size of the gap at the origin firm.

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 bias coefficient is 0.97. Among all firm switchers, the coefficient is 0.70. Under a strong "no sorting" benchmark, the paired mover coefficient is proportional to the relative variance of person and match effects. Under this benchmark, this coefficient implies that the variance of match effects is larger than the variance of firm effects.

We then develop subgroup analyses which are suggestive of violations of exogenous mobility. We find lower coefficients for moves across "contexts" than within contexts, which is what we expect if there were context specific match effects. For within-sector moves the coefficient is 0.73, and for across sector moves it is 0.67. In the Veneto data, these differences are starker. One important feature of the data is that workers are disproportionately likely to stay in the same "context," rather than switching contexts. Such a pattern is consistent with workers taking match effects into account when deciding whether to move, which would violate exogenous mobility. This pattern occurs in models of match effects where the importance of match effects implies that much mobility is not realized (e.g., Borovičková and Shimer (2024)).

For comparison to the literature, we implement the event study test in our data. The forecast coefficient is 1.07, which is consistent with the general finding in the literature that changes in firm effects do a good job of predicting changes in earnings. Interestingly, when we restrict to the paired mover sample, we find a very similar coefficient of 1.05, which is further evidence that the paired mover sample is not unusual relative to the broader sample. In the Veneto data, the sample

matters a lot more for the results of the event study test.

To develop a point estimate of the role of match effects, we estimate a simple random job search model where workers are heterogeneous in productivity and move on the basis of both a firm effect and a match effect. We match an "AKM moment" of the variance of the firm effects, the earnings gap between paired movers (which depends on person effects and match effects), the paired mover regression coefficients, as well as EE, EU and UE rates.

Quantitatively, the model finds that the variance of match effects in the steady state distribution is just over half the variance of firm effects and is thus smaller than that implied by our "no sorting" benchmark. The main reason is that workers are sorted across firms and so the within firm wage dispersion is much less than what is implied by the no-sorting benchmark. Calibrating to the earnings gap between paired movers captures this sorting.

This simple model of mobility based on match quality and a firm effect—which thus violates exogenous mobility—generates an event study forecast bias coefficient of 0.99. This finding is consistent with the idea that the event study test is low-powered against alternative hypotheses that include an important role for endogenous mobility.

The model also shows that 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 the slope departs from 1. Thus, in the context of the model accepted offers are a poor guide to average treatment effects in the population.

To show that we are detecting deviations from additive separability that are new to the literature, 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. We do not find that the BLM estimator is able to reliably detect the deviations from additive separability that are detected by the paired mover test. 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 an unpublished paper, Kantenga and Law (2016). Section 6 of that paper 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. Relative to Kantenga and Law (2016), this paper develops a more transparent approach to looking at this implication. By formulating it as a forecast bias test, this paper also connects to other features of the data that researchers studying the AKM model have explored.

Woodcock (2015) is another paper that develops an estimate of match effects in the context of the AKM model. The paper pursues multiple estimation approaches. The easiest to explain

is the orthogonal match effects model where the match effect is the mean residual of the AKM model within firm-worker match. The paper finds that the variance of match effects is about 2/3 the variance of firm effects, and so our estimates are quantitatively similar. 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 these same edges but looks at a 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) develop a U-AKM model, where the worker's fixed effect is allowed 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. In this sense, these exercises are reminiscent of the event study test. A typical finding is that the 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 also related to Borovičková and Shimer (2024). The main thrust of that paper is to develop a model with endogenous mobility both because there are match effects, and because workers and firms are "picky" about whom they match with. The paper shows that the model implies a wage equation that is additively separable in levels (and approximately so in logs). Most provocatively, 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. In our model we replicate the finding that it is possible to pass the event study test in an economic model that violates exogenous mobility. 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 (since prior work had supported this functional form restriction), our goal is to develop a direct test of this functional form restriction.

1 Additive separability versus exogenous mobility

In this section, we discuss the difference between exogenous mobility and additive separability, which is important to interpreting our results.

1.1 Informal discussion

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 employers 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). We will see, however, that the conventional assumption deviates from additive separability.

For expositional purposes, it is helpful to depart from the conventional equation and explicitly write the wage equation to include match effects, which now reflects a deviation from additive separability:

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

where $m_{ij} \equiv \mathbb{E}[\epsilon_{it}|i,j]$ so that m_{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 $\tilde{\epsilon}_{it} \equiv \epsilon_{it} - m_{ij}$. This definition of a match effect as a permanent component of the match parallels how person and firm effects are defined, and parallels the definition in equation (3) in Woodcock (2015). Below, we characterize the time series properties of the match effect. 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 m_{ij} as match effects in what follows, though the m_{ij} are a statistical—and not an economic—object.

Conventionally, researchers impose an exogenous mobility assumption. Informally, exogenous mobility says that workers do not, on average, select a firm based on the error term.

To see the distinction between exogenous mobility and additive separability, suppose that we see each worker at all firms for an extended period of time. Then we can estimate the average treatment effect of each firm on earnings in the whole population (relative to some omitted firm). For example, suppose that there are N people, J firms and T time periods. Posit asymptotics where N and T grow large. Assume that $\frac{T}{J}$ (the number of periods that each worker spends at each firm) is an integer. Suppose we index firms by numbers and omit the 0^{th} firm. Then the plug-in

estimator for the firm j effect is:

$$\hat{\psi}_j = \frac{J}{NT}\sum_i\sum_t y_{it|\mathbf{j}(i,t)=j} - \frac{J}{NT}\sum_i\sum_t y_{it|\mathbf{j}(i,t)=0}. \tag{3}$$

Substituting in for the definition of y, we have:

$$\hat{\psi}_j = \frac{J}{NT}\sum_i\sum_t \left[\alpha_i + \psi_{\mathbf{j}(i,t)=j} + \beta X_{it} + m_{i\mathbf{j}(i,t)=j} + \tilde{\epsilon}_{it}\right] - \frac{J}{NT}\sum_i\sum_t \left[\alpha_i + \psi_{\mathbf{j}(i,t)=0} + \beta X_{it} + m_{i\mathbf{j}(i,t)=0} + \tilde{\epsilon}_{it}\right]. \tag{4}$$

We can see that the α_i cancel, the assumption that the m_{ij} and $\tilde{\epsilon}_{it}$ are mean zero for each $i-j$ combination means that these cancel out. So then $\psi_j = \frac{J}{NT}\sum_i\sum_t \psi_{\mathbf{j}(i,t)=j} - \frac{J}{NT}\sum_i\sum_t \left[\psi_{\mathbf{j}(i,t)=0}\right]$.

While ψ_j captures the average effect of firm j on earnings, this effect averages over arbitrary heterogeneity. For worker i, the average change in earnings of going from firm j' to j is

$$\psi_j + m_{ij} - \psi_{j'} - m_{ij'}.$$

Exogenous mobility allows the treatment effect of a firm to be heterogeneous and thus is consistent with the presence of match effects. The key implication of exogenous mobility is that matches are not formed on the basis of the match effects.

For the interpretation of the results in this paper, the relevant takeaway is that 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. More deeply, specification tests for exogenous mobility do not directly test additive separability. Below, however, we use our machinery to develop some auxilliary evidence about the plausibility of exogenous mobility.

1.2 (More) Formal discussion

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, pg. 14)):

Assumption 1. *Assume that:*

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

and

$$\mathbb{E}[m_{ij}|\mathbf{j}(i,s), X_{is} = x] = 0, \tag{6}$$

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* $t \in \{1,....,T\} \equiv [T]$*), 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.

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

Assumption 2. *Assume that:*

$$\mathbb{E}[\tilde{\epsilon}_{it}|\mathbf{j}(i,s), X_{is} = x] = 0, \tag{7}$$

and

$$m_{ij} = 0 \ \forall \ i, j, \tag{8}$$

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* $t \in \{1, ..., T\} \equiv [T]$*), all possible firm assignments* $j \in [J]$ *(where* $j \in \{1, ..., J\} \equiv [J]$*.*

Exogenous mobility says that on average the match effects in realized matches are zero. Additive separability says that the permanent component of the match effects are all zero.

The paired mover test that we develop in this paper is a test of additive separability. It is logically possible to reject additive separability but for exogenous mobility to hold. In this case, the firm effects are average treatment effects. There are many questions for which it is interesting to know the average treatment effect of a firm.

2 A test of (log) additive separability

In this section, we introduce the paired movers test. We compare the paired movers test to the event study test. We discuss important implementation issues, and how to interpret deviations of the forecast bias coefficients in both tests from the benchmark value of 1.

2.1 The paired movers test

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:

$$y_{iA} - y_{i'A} = (\alpha_i - \alpha_{i'}) + (m_{iA} - m_{i'A}) + (\tilde{\epsilon}_{iA} - \tilde{\epsilon}_{i'A}). \tag{9}$$

By differencing across workers within firm, the firm effect drops out. Additive separability implies that $\mathbb{E}[\epsilon_{iA} - \epsilon_{i'A}|\mathbf{j}(i,s) = A, \mathbf{j}(i',s) = A] = 0$ and $m_{iA} = m_{i'A} = 0$. So that $\mathbb{E}[y_{iA} - y_{i'A}] = (\alpha_i - \alpha_{i'})$. Now consider the difference between worker i and i' at firm B:

$$y_{iB} - y_{i'B} = (\alpha_i - \alpha_{i'}) + (m_{iB} - m_{i'B}) + (\tilde{\epsilon}_{iB} - \tilde{\epsilon}_{i'B}). \tag{10}$$

To construct the test, we consider whether the earnings *gaps* at firm A predict earnings *gaps* at firm B:

$$(y_{iB} - y_{i'B}) = \beta_0 + \beta_1(y_{iA} - y_{i'A}) + \varepsilon_{i,i',A,B}. \tag{11}$$

We can substitute in for the definitions of the earnings gaps to write:

$$(\alpha_i - \alpha_{i'} + (m_{iB} - m_{i'B}) + \tilde{\epsilon}_{iB} - \tilde{\epsilon}_{i'B}) = \beta_0 + \beta_1((\alpha_i - \alpha_{i'}) + (m_{iA} - m_{i'A}) + \tilde{\epsilon}_{iA} - \tilde{\epsilon}_{i'A}) + \varepsilon_{i,i',A,B}. \tag{12}$$

Under the assumption of additive separability, 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 exogenous mobility, which implies the presence of match effects. Below, we solve for the regression coefficient in terms of variances and covariances and discuss their economic interpretation.

2.2 Comparison to the 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). This analysis asks whether the change in the firm effects on average predict the change in earnings when workers switch firms. This test is thus testing whether firm effects represent average treatment effects, which is what the exogenous mobility assumption implies.

To write that test in the same notation, consider worker i at firm A and B:

$$y_{iA} = \alpha_i + \psi_A + m_{iA} + \tilde{\epsilon}_{iA} \tag{13}$$

and

$$y_{iB} = \alpha_i + \psi_B + m_{iB} + \tilde{\epsilon}_{iB}. \tag{14}$$

The basis of the event study test is to ask whether changes in the firm effects (or mean co-worker earnings) predicts changes in earnings within worker. The regression version is:

$$y_{iA} - y_{iB} = \beta_0 + \beta_1(\psi_A - \psi_B) + \varepsilon. \tag{15}$$

Plugging in for the definition of y_{iA} and y_{iB} we have:

$$(\psi_A - \psi_B) + (m_{iA} - m_{iB}) + (\tilde{\epsilon}_{iA} - \tilde{\epsilon}_{iB}) = \beta_0 + \beta_1(\psi_A - \psi_B) + \varepsilon. \tag{16}$$

Under the assumption of exogenous mobility and abstracting from difficulties in estimating the ψ, the probability limit of β_1 is 1 because this regression is effectively a regression of $\psi_A - \psi_B$ on itself (on average, the m and $\tilde{\epsilon}$ terms are zero). Deviations from 1 are indicative of deviations from

exogenous mobility. In implementation below, we address concerns about measurement error by splitting the sample on the basis of workers and computing two measures of the ψ_A and the ψ_B so that we can instrument.

2.3 Implementation issues

There are two key implementation issues: first, noise in estimates, and, second, the role of covariates.

The first concern is that the earnings gap at firm A might be measured with noise. In this case, the coefficient on β_1 would be biased down and we would tend to overreject. The solution is that we restrict to cases where we have two observations of earnings gaps at firm A and we instrument one measure of earnings gaps with the second measure.

In section 1, we defined the match effect as the permanent component of the match (in the wage equation). Given a finite dataset, one cannot hope to actually identify this parameter. By instrumenting, we implicitly define the match effect as the component of the match effect that is persistent between two periods. In the empirical section, we empirically characterize the covariance structure of the gap over time and find that it is well-approximated by an AR(1) structure with a persistence coefficient above 0.93.

The second concern is how to control for covariates. Conceptually, if one worker is young and on the steep part of her lifecycle earnings profile and the other worker is old and on the flat part of her lifecycle profile, then we would not find stable earnings gaps across firms. This instability would just be about the lifecycle earnings profile, and not about match effects. To address this issue, we residualize earnings for flexible functions of age, education and gender (we also show robustness to matching on age and gender).

2.4 Interpretation of paired movers test

To interpret deviations from 1 in the forecast coefficient, we compute the probability limit of the forecast coefficient. The asymptotic thought experiment is that the number of paired movers grows. Then the probability limit is given by:

$$\begin{aligned}\text{plim } \hat{\beta}_1 = &\frac{Var(\alpha_i - \alpha_{i'}) + Cov(m_{iA} - m_{i'A}, m_{iB} - m_{i'B})}{Var(\alpha_i - \alpha_{i'}) + 2Cov(\alpha_i - \alpha_{i'}, m_{iA} - m_{i'A}) + Var(m_{iA} - m_{i'A})} \\ &+ \frac{Cov(\alpha_i - \alpha_{i'}, m_{iB} - m_{i'B}) + Cov(\alpha_i - \alpha_{i'}, m_{iA} - m_{i'A})}{Var(\alpha_i - \alpha_{i'}) + 2Cov(\alpha_i - \alpha_{i'}, m_{iA} - m_{i'A}) + Var(m_{iA} - m_{i'A})}. \end{aligned} \tag{17}$$

A useful special case is when all of the covariance terms are zero. This case corresponds to a "no sorting" benchmark and is useful to understand how we can map the paired mover coefficient into an estimate of the role of match effects. In this case, the probability limit reduces to:

$$\text{plim } \hat{\beta}_1 = \frac{Var(\alpha_i - \alpha_{i'})}{Var(\alpha_i - \alpha_{i'}) + Var(m_{iA} - m_{i'A})}. \tag{18}$$

In this special case, the probability limit reflects the relative dispersion of person effects between

the paired movers to the dispersion in the match effects among the movers. In this special case, we can invert the paired mover coefficient to tell us about the role of match effects relative to person effects. It also illustrates a potential pitfall of the paired movers test: to the extent that the paired movers are very similar, i.e., $Var(\alpha_i - \alpha_{i'}) \approx 0$, then this test does not tell us very much about the role of match effects. For example, there is an extreme case where the paired movers always have the same person effects. In this case, even if match effects are economically insignificant, we might still find a large role for match effects. Below, we discuss how the variance of the earnings gaps—which includes the person effects—in the paired movers sample compares to other benchmarks.

The covariance terms correspond to different forms of selection. For example, for the $Cov(\alpha_i - \alpha_{i'}, m_{iB} - m_{i'B})$ term, natural forms of selection would imply that these terms would tend to be negative: the firm will only hire a worker if she is good enough, which is the sum of the person and match effect. So conditional on two people ending up at the same firm, these are negatively related. The $Cov(m_{iA} - m_{i'A}, m_{iB} - m_{i'B})$ term is about the persistence of match effects. This covariance would be negative if workers i and i' are very different and so the firms where i needs a good match effect to match are those where i' needs a bad match effect to match and vice versa. In the economic model below, we map the structure of the model to these statistical objects.

2.5 Interpretation of event study test

As for the paired movers test, we now compute the probability limit of the forecast coefficient. Using the notation in equation (16) the probability limit is given by:

$$\text{plim } \hat{\beta}_1 = \frac{Var(\psi_A - \psi_B) + Cov(m_{iA} - m_{iB}, \psi_A - \psi_B)}{Var(\psi_A - \psi_B)}. \tag{19}$$

If mobility is exogenous, then this coefficient equals 1. The forecast coefficient is below 1 when the differences in match quality are negatively correlated with the differences in firm effects, and positive when the differences in match quality are positively correlated with the differences in firm effects. This natural way of generating a coefficient above 1 differs from the paired movers test.

Similar to the paired movers test, the interpretation of the deviation rests on how "large" the relevant moves are that go into the event study test. If most workers make relatively small moves in firm effects space, then the coefficient could depart meaningfully from 1 without corresponding to a large economic deviation.

3 Data

3.1 Brazilian data

We use *Relação Anual de Informações Sociais* (RAIS) data from Brazil, which is a matched employer-employee data set that contains information on all formal job contracts. Each observation in the data is a job spell, the notion of employer is a establishment within a firm, and the

frequency is annual. We start from the universe of spells from 2006 to 2022 and obtain a single observation per worker per year by keeping the job with the highest earnings (average monthly wage $\times$ months worked) in each year. Each year, we drop workers on temporary contracts, public servants, and those with zero earnings. Finally, we focus on workers aged 18 to 64. We refer to this sample as the sample of "all workers."

We define a mover as a worker who changes employers between two consecutive years and (i) separated from the previous employer and (ii) such separation was either voluntary (employee-initiated) or involuntary (employer-initiated). The first condition excludes workers who change their "dominant" job but remain employed at the original firm. The second condition excludes workers who transfer within establishments of the same firm. We often refer to this sample as the sample of "all movers."

To construct the set of paired movers, we make a few restrictions. 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. We also explore the time series properties of the gaps and so restrict to samples where the paired workers are both present for more years in both the origin and destination firms. Second, we want to ensure that the paired movers are actually moving firms, and so 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)). By construction, we also need to observe the worker at least once at the destination firm, and she must move with at least one co-worker.

A move is defined by origin, destination, and year. 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. However, workers can appear more than once in the sample since they can move between different firms in different years. We also construct a sample of "paired stayers," where A and B is the same employer. Since the resulting sample is too large, 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 – there is, we drop the stayers. The third sample includes only workers in the second sample who are in the paired movers sample (and her move in the paired movers sample is between 2012 and 2017).

The main variables we report from the data are gender, age, years of education, tenure, occupation, earnings, and firm-level variables such as sector, size, average pay, and location.

Our baseline notion of earnings is the average log hourly wage through the year. Job contracts in Brazil are defined by monthly wages and weekly working hours. We compute earnings as average monthly wage $\times$ (weekly hours $\times$ 30/7) (Gerard et al., 2021). We compute firm size as the number of workers employed on December 31st of each year. Average pay is the average log hourly wage

of these workers.

Age is age 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 as the difference between the date of hire and the date of separation (or Dec. 31st). Occupation is classified using 3-digit *Classificação Brasileira de Ocupações* (CBO) codes. Sectors are defined using 2-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:

$$y_{it} = X_{it}\beta + m_{ij(i,t)} + \epsilon_{it},$$

where m_{ij} is a dummy for the interaction of worker i and firm $j(i,t)$. Our baseline set of covariates is a fourth degree polynomial in age, interacted with gender and four education categories (no primary education, primary education, high school, and some college or more). Second, we construct residualized earnings by subtracting off the estimated effect of covariates:

$$\tilde{y}_{it} = y_{it} - X_{it}\hat{\beta}.$$

We use residualized earnings throughout the analysis except that in the AKM analysis we show how the "conventional" AKM regression compares to the results using residualized earnings. In the Appendix, we also present an alternative approach where we match workers on bins of age and gender.

3.2 Veneto data

The Veneto Worker History data in its full version runs from 1976 to 2001.

In terms of cleaning the data, we use cleaning code used in Kline, Saggio, and Sølvsten (2020) and so our description follows their computational appendix. The wage record consists of a start date, end date, number of days worked in a given year, and the total wage compensation paid in a given year. We can thus compute a notion of a daily wage in each year.

We focus on workers aged 18 to 64. We focus on the worker's "dominant" job, which is the job from which the worker made the most money in a given year. We also exclude workers with implausibly low wages (below 5 euros or zero days worked), implausibly large wage changes (more than 1), or those employed in the public sector, having more than ten jobs or missing gender.

We use the data from 1985 to 2001 because the view is that data on days worked is unreliable before this date.

Based on the description of the Veneto data in Card, Devicienti, and Maida (2014) the industry codes are 5-digit ATECO 91 codes. The employers in the Veneto data are described as firms.

To construct the set of paired movers, we make analogous restrictions to what we did in the

Brazilian data. In many cases, there are more than two workers who leave a firm in a particular year. In this case, we form all possible pairs.

4 Results

In this section, we first discuss descriptive statistics on our sample. The key question is understanding who the paired movers are and how they are selected. We then present the main results documenting meaningful deviations from additive separability and present extensive robustness. Finally, we show that the deviations from additive separability are stronger in moves across contexts.

4.1 Descriptive statistics

Table 1 shows summary statistics in three samples. The first sample is all workers. The second sample is all workers who switch firms. The third sample is the set of paired movers. In general, movers are negatively selected relative to all workers in terms of tenure and earnings. In contrast, paired movers are positively selected. Relative to all movers, paired movers come from much larger employers and higher-paying employers. Movers are a 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 over a one year window—not including the paired movers. Using these empirical probabilities, we then simulate mobility year by year using the number of movers we see each year and compute the number of paired movers, where a paired move has to occur in the same year. As can be expected from the fact that paired movers have disproportionately large origin and destination firms, Table 2 shows that there are, if anything, fewer paired movers than we would expected based on the mobility patterns of other workers. We interpret this evidence as reassuring that the paired movers are not too unusual and so it makes some sense to extrapolate from them to other movers.

If there are very small earnings gaps between the paired movers in their origin firms, then being unable to predict the earnings gap at the destination firm would not be strong evidence of match effects because the base would be so small. In Table 3 we report statistics about the distribution of the gap in earnings at the origin firm among the paired movers. The mean of the gap is around 36 log points. For comparison, Table 3 also reports the distribution of earnings gaps between pairs of workers constructed in two different ways. First, we take a random set of workers who are "similar" to the paired movers: we look at random pairs of workers who are observed at the same firm for two years and then we see again for at least a year (either in the origin firm, or at a different firm). These workers must be in the same gender and age bin, so could be paired workers. Second, we take random pairs of stayers and compute the gap. The stayers mechanically are present for three years. The stayers have more dispersion than the paired movers (45 log points, versus 36 log points for the paired movers). Because of these differences, below we consider the effect of reweighting

the paired movers to match the random benchmark.

Finally, 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. Second, using mobility patterns generates fewer paired movers than we see, whereas in Brazil it generates more. Third, the mean earnings gaps are surprisingly similar.

4.2 Main results

Figure 2 shows the graphical version of our main IV 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 the next period. The red dots show the first-stage of the gap in earnings in two consecutive periods in the paired mover sample at the origin firm. The dots are approximately linear and lie very close to the 45 degree line, indicating that earnings gaps are quite stable. The blue squares show earnings gaps from the year before the move to the year after the move. These dots are also approximately linear, but the slope is much shallower than between the years when the workers do not move. This indicates that earnings gaps are not stable across contexts.

Table 4 shows the main IV results in a table. We first show results for paired stayers. Paired stayers provide a quantitative benchmark that is perhaps more empirically realistic than forecast unbiasedness in that it allows for the possibility of deviations from the assumptions of the benchmark additively separable specification that are different from match effects. For example, if there is drift in the individual effects, then the paired movers test 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 0.97, indicating minimal "drift" in match effects between periods. In contrast, for paired movers the coefficient is 0.70, which is meaningfully different from 1 and from the stayers coefficient. Consistent with the first stage dots lying close to the 45 degree line, Panel B shows that the first stage coefficient is around 0.93 for the baseline paired movers and 0.94 for the stayers.

What does the magnitude of 0.70 mean? As we showed in Section 2, the probability limit of the regression coefficient is a complicated function of variances and covariances. We discussed a special case where there is no sorting between worker effects and match effects. In this case, the probability limit tells us about the ratio of the variance of the (difference) in individual effects to that of match effects. In this special case, $\hat{\beta} = 0.7$ implies that $Var(m_{iA} - m_{i'A}) = \frac{0.3}{0.7} Var(\alpha_i - \alpha_{i'})$. If we further assume that there is no covariance between the worker effects or the match effects, then this implies that the variance of match effects is about 40% of the variance of individual effects. In our data, this benchmark implies that variance of match effects is about 50% *larger* than the variance of firm

effects. We return to this benchmark below in the context of the model.

Robustness to residualization: In our benchmark results, we residualize for a fourth degree polynomial in age interacted with gender and education. Appendix Table A4 shows that our results are similar if we instead just residualize for the age polynomial interacted with gender. Rather than residualizing earnings, we can also match workers on age and gender. Appendix Figure A1 shows that the paired mover coefficient is quite similar whether we match on 10 year age bins or 1 year age bins, with the coefficient ranging from 0.70 to 0.73. We prefer to focus on the residualized approach because it preserves sample sizes.

Time series properties: In Section 1, we defined match effects to be a permanent feature of the match. The implication of this assumption is that 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 0.97, which indicates small deviations from permanence.

Table 7 presents further evidence that the gaps are quite persistent by zooming in on the sample of paired movers and computing the correlation between the gaps at various horizons and in various samples (balanced panel and full sample) at both the origin and destination firms. To help interpret these correlations, we convert them to implied AR(1) coefficients at each horizon[1] The bottom panel of Table 7 shows these coefficients range from 0.93 to 0.97 (in the balanced panel).

We follow Lachowska et al. (2023, Footnote 9) and pool these coefficients using a regression.[2] For the balanced panel, we find an estimate of the persistence parameter of around 0.93. 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 the firm effects.

As an alternative way of showing the relative stability of the match effects, Figure 4 shows that our results are not very sensitive to which year we measure the earnings gaps. It provides an event study representation of our main results. The left-side of the graph looks at paired movers who are at their origin firm and plots "stayer" coefficients. The dot at -3 is workers who are at their origin firm in -5, -4, and -3 and we instrument the gap at -4 with the earnings in -5 and see how related they are to -3. We then repeat the exercise in -2, -1 and 0. The figure shows this regression coefficient for both a balanced panel and for the "full sample" of paired movers. The basic message of the figure is that the earnings gaps are quite stable over time at the origin firm. The right-side of the graph instruments for gap at 0 with the gap at -1 and then looks at outcomes in different years at the destination firm. The first thing to note is the dramatic drop from the left- to the right-hand side: gaps are much less stable when switching firms than staying within firms.

[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.

The second thing to note is that the gaps are quite stable regardless of which year we use in the destination firm. This evidence is consistent with the match effects being quite persistent.

For our primary results, we ask whether earnings gaps at the origin firm predict gaps at the destination firm. Conceptually, we could also do the test in reverse, and ask whether the gaps in the destination firm predict the gaps in the origin firm. In Appendix Table A6 we construct a sample where we observe workers for two years at both the origin and destination employers. This way we can consider both exercises. We find that the paired mover coefficients are a bit higher for the "reverse" exercise: the benchmark coefficient is 0.82 rather than 0.70. In ongoing work, we have shown that this result is because the earnings gaps in the first year at a new firm appear to have a larger transitory component than the gaps in other years (something that can be seen in Table 7). When we do this reverse exercise doing other years (albeit for a subset of matches), we get more stable coefficients.

How coefficient varies with the size of the gap and robustness to reweighting: One hypothesis is that the earnings gaps are unstable for small gaps and stable for large gaps, and so the deviations are less important than they might appear. As discussed above, the gaps in earnings between the paired movers at their origin firms are smaller than various random benchmarks. Panel (a) of Figure 3 shows that there is a substantial dispersion in the gaps at the origin firms; for example, 75th percentile of the distribution of the gap at the origin firm has a gap of 0.5, so about 25% of the sample has gaps larger than the mean of the random within firm benchmark.

Panel (b) shows that when we estimate the paired mover coefficient separately by percentiles of the distribution of gap at the origin firm we find that the paired mover coefficient is declining in the size of the gap. (For very small gaps at the origin firms, we have very small first stages and so estimate very large paired mover coefficients; to make the graph legible, we pool the first 25 percentiles). The basic pattern is that rather than the coefficient *increasing* in the size of the gap (so that large gaps are more stable) we find that the coefficient is *decreasing* in the size of the gap and remarkably stable across the support of the gap. Notably, for gaps from the median of the distribution to the 98th percent, the paired mover coefficient is visually indistinguishable from our baseline estimate, and the 99th percentile coefficient is slightly smaller. Thus, paired movers with larger gaps see less stable earnings gaps than those with smaller gaps.

This finding is important when it comes to translating the paired mover coefficient into an estimate of the role of match effects because the size of the match effects will be proportional to the variance of the difference in the person effects among the paired movers.

To summarize this observations that the forecast coefficient is smaller for paired movers with larger gaps at the origin firm, Appendix Table A5 shows that when we reweight the paired movers to match the random benchmark we find smaller forecast coefficients.

Veneto data: The Veneto data generates a paired mover coefficient that is quite similar to that in Brazil: 0.78 versus 0.76. The stayer coefficient is slightly lower: 0.93 rather than 0.97. Notably, and indicating that the data are noisier, the first stage coefficients are much lower. The effect of

reweighting is similar in the Veneto data in that it just lowers the coefficient slightly. In contrast to the Brazilian data, the paired mover coefficient is more similar for the "forward" and "backward" regression: 0.72 vs. 0.75.

4.3 Heterogeneity

Table A1 shows additional summary statistics on the paired movers, and all of the subsamples of paired movers that we discuss below.

We first split by whether the paired movers are moving across sector or within sector. We define sector as a two digit CNAE code. Appendix Table A1 shows how the sector stayers compare to the sector movers. Sector stayers are older, higher paid, higher tenure and have more schooling than sector movers. In terms of what would be predicted by chance Table 2 shows that the mobility patterns of other workers would predict that about 47% of paired movers would be within sector, whereas in the data 46% of them are. Thus, the notable feature of the data is that mobility patterns are relatively concentrated within sector, rather than that the paired movers are unusual.

Columns (3) and (4) of Table 4 show that the forecast bias coefficient is higher for the stayers than the movers: 0.73 versus 0.67. The finding that earnings gaps are more stable for "small" sector moves can be generalized. For each origin sector, we rank destination sectors by the frequency with which workers arrive. Panel (a) of Figure 5 shows that as we move from the most frequent sector move to less frequent moves, 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. Inspection of the data reveals that retail trade is a very common destination sector in almost all origin sectors. Separating out retail trade, we get a coefficient of 0.73 for the closest sector, falling to 0.63 for the least likely sectors, and 0.59 for retail trade.

Columns (5) to (8) of Table 4 shows heterogeneity where we split moves on the basis of mean pay at the firm and we split firm pay at the median of the distribution. Appendix Table A1 shows important differences by whether workers are staying or switching between different parts of the pay distribution. Those originating in high-paying firms have higher tenure and earn more than those originating in low paying firms. Columns (6) through (9) of Table 2 shows that again the distribution of paired movers in terms of staying at the same level of firm pay or changing is similar to what would be predicted by chance. And, again, the notable feature of the data is the low share of "switchers." Turning to the regression results, Columns (5) to (8) of Table 4 shows that earnings are gaps are more stable for workers staying in the same "context" than switching. This gap is especially large for workers starting in high-paying firms, where those who move to another high-paying firm have a paired mover coefficient of 0.72 whereas those who switch to low-paying firm have a coefficient of 0.62.

Cluster stayers versus switchers also show differences in the paired mover coefficient. We discuss how we construct the clusters of firms in Section 7). Columns (9) and (10) of Table 4 shows that the paired mover coefficient is 0.74 for the stayers and 0.68 for the switchers.

Another way of measuring context is by occupation. In columns (11) through (13) we consider

moves where neither worker changes occupation, one worker changes occupation or both change occupation. Here we find the starkest differences: the paired mover coefficient is 0.77 for moves where both workers keep their occupations but only 0.60 where both workers switch their occupations.

Columns (14) through (16) show, interestingly, that the earnings gaps are most stable for high-tenure paired movers than for low-tenure paired movers. Perhaps surprisingly, columns (17) and (18) show that there are minimal differences for paired movers within or across geography.

Figure 7 summarizes these results in a figure.

Across these splits, workers are more likely to stay in the context than to switch. Table 2 shows that 53% of moves switch sectors, whereas random mobility would say that 96% of moves should switch sectors. 68% of moves switch clusters, whereas random mobility would say that close to 90% of moves should switch clusters. Similar patterns apply to the earnings stayers.

One model that qualitatively generate some of the patterns we have documented is a model inspired by Neal (1999). Suppose the payoff to working at a firm in a sector is the firm effect plus a sector match effect. If the worker changes firms within a sector, then she keeps the sector match effect, whereas if she switches to a firm in a different sector then she redraws the sector effect. This model generates the pattern that for moves within sectors changes in firm effects do a good job of predicting changes in earnings, while for moves across sectors changes in firm effects are less good predictors. Similarly, it generates the pattern that cross-sector moves are unlikely relative to within sector moves precisely because match effects are important. We return to a model in this spirit in Section 6.

Another aspect of mobility that plausibly selects for match quality is the nature of the move: was it a voluntary or involuntary move, and, if so, how long did it take the worker to make the move? The notion of voluntary refers to the reason for separation in the RAIS data. Figure 6 shows that there are meaningful differences between voluntary and involuntary movers both in terms of the average earnings changes (involuntary movers see smaller earnings gains or cuts rather than gains as compared to the voluntary movers) and also in terms of the non-employment duration between the jobs. Table 5 splits paired movers both by voluntary and involuntary, and, within each category, by the time between jobs. Note that both workers in the pair satisfy the same criteria, so this table does not show all moves. (Appendix Table A3 shows the comprehensive splits by voluntary and involuntary, as well as mixed status). There are not sharp patterns about whether the paired mover coefficient is lower or higher for the involuntary or voluntary movers. Where there is a sharp pattern is that the coefficient declines in non-employment duration. For the involuntary moves, it starts at 0.80 for moves within a month, and declines to 0.62 for moves with a gap of 6 or more months. For voluntary moves the slope is shallower, and goes from 0.74 to 0.69. We develop a model-based interpretation of this pattern in Section 6.

Veneto data: The results on heterogeneity in Veneto are generally more stark than in Brazil. For example, the sector stayers have a coefficient of 0.91 whereas the switchers have a paired mover coefficient of 0.58.

5 Event study test

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.

Table 6 shows that the forecast coefficient in the event study test is 1.07. The standard errors are tight enough that we can reject 1, but we do not interpret this deviation as economically large. In Column (2) we restrict to the paired mover sample and find a coefficient of 1.05. The small difference between the event study test with all movers and paired movers is further reassurance that the paired mover sample is not strongly selected relative to all movers. In columns (3) through (5) we reweight the sample to match the paired mover sample in terms of origin firms, destination firms, and their interaction. These changes slightly reduce the forecast coefficient to be closer to 1.

Appendix Table A8 parallels Table 5 in doing the event study test splitting moves by voluntary and involuntary and, within these categories, by nonemployment duration. As with the paired mover test, there is a declining coefficient in nonemployment duration within reason: for voluntary moves it is from 1.13 to 1.02 and for involuntary moves it is 1.09 to 0.92.

Veneto data: The patterns in the Veneto data are a bit different than in the Brazilian data. The main coefficient is 1.14, which is very similar to what we find in Brazil. What differs in the Veneto data is that when we move to the paired mover sample the forecast coefficient dramatically declines: to 0.73. The reweighting exercise in columns (3) through (5) shows that this difference can be explained by the set of firms we are considering.

6 Model-based interpretation of the results

In this section we develop and estimate a simple partial equilibrium search model with endogenous mobility to translate the forecast bias coefficients into economic magnitudes.

6.1 The model

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

$$\alpha_i + \psi_j + m_{ij}.$$

As we formalize in the next paragraph, the worker sees the m_{ij} before making her mobility decision, so this model violates both additive separability and exogenous mobility.

The value of unemployment for a worker is:

$$U = b + \beta \left(\lambda_0 \int_m \int_\psi \max\{W(\psi', m'), U\} dF(\psi') dM(m') + (1 - \lambda_0) U \right). \tag{20}$$

This equation says that a worker receives the flow value of unemployment of b. Then the worker receives an offer with probability λ_0, and accepts it if it increases her value. If she does not receive an offer, then she remains unemployed. The notable feature is that the offer includes both the firm effect and the match effect.

The value of an employed worker with person effect α_i at a firm with firm effect ψ and with match effect m is:

$$W(\alpha_i, \psi_j, m_{ij}) = \alpha_i + \psi_j + m_{ij} + \beta \Bigg(\lambda_1 (1 - \delta) \int_m \int_\psi \max\{W(\alpha_i, \psi_j, m_{ij}), W(\alpha_i, \psi', m')\} dF(\psi') dM(m') + (1 - \lambda_1)(1 - \delta) W(\alpha_i, \psi_j, m_{ij}) + \delta U \Bigg). \tag{21}$$

This equation says that an employed worker receives the flow payoff of her match. If her job is not destroyed, she can receive an outside offer, and she makes a maximizing decision whether or not to accept it. 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.

6.2 Estimation

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

We estimate parameter in two steps. We estimate the EU rate (δ) and UE rate (λ_0) by matching quarterly flow rates computed in survey data. We are then left with four unknowns: λ_1, and the variance of person, firm and match effects. We match four empirical moments: the EE rate, the paired mover regression coefficient, the variance of firm effects (estimated using AKM and the Kline, Saggio, and Sølvsten (2020) correction), and the mean absolute earnings gap between paired movers that we documented in Table 3. The reason to target the earnings gap between paired movers is that the paired movers coefficient fundamentally depends on this gap (see equation (17)). Targeting the gap is also a way of capturing the extent to which workers are sorted across firms.[3] We search the parameter space using Metropolis-Hastings.

[3]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).

6.3 Model fit and validation

The first part of Table 9 shows that we are able to closely replicate the four moments that we match. The paired mover coefficient is 0.69 in the data and 0.70 in the model, the variance of firm effects is 0.07 in both, and the mean absolute gap is 0.36 in the data and 0.38 in the model. Finally, the EE rate is 0.04 in both.

Besides matching the paired mover coefficient, Column (4) of Table 10 shows that the model also generates an event study coefficient of 0.99, which is close to what we see in the data. This finding is reminiscent of Borovičková and Shimer (2024) and emphasizes that the event study test is a low-powered test. The reason is that in this model workers move on the basis of the match component. The estimates of the match component are large. So the model structure is very far from satisfying exogenous mobility. And yet the test that has been emphasized in the literature as supportive of the exogenous mobility assumption fails to detect this deviation.

The model can also qualitatively match the patterns in Table 5. That table showed that the paired mover coefficient was lower for moves with a longer time of nonemployment between the employment spells. We can interpret this distinction as mapping onto EE and EUE moves. It is not obvious that the model would be able to match this pattern. On the one hand, in the model EE moves are all voluntary and so workers select on match effects more strongly than for EUE moves. So we might think that we would find a lower coefficient on EE moves in the model than for EUE moves. On the other hand, to the extent that the selection on match effects is similar across pairs of workers then we might expect a lower coefficient on EUE moves. Table 10 shows that the model generates a paired mover coefficient of 0.79 for the EE moves and 0.65 on the EUE moves. While this distinction is quantitatively larger than what is in the data, it is qualitatively consistent. In contrast, the model implies that the event study test coefficient should be lower on EE than EUE moves, which is not seen in the data.

Adjusted R^2 in AKM and the interacted model: Besides the event study test, one important piece of evidence in Card, Heining, and Kline (2013) against the importance of match effects is the small change in adjusted R^2 in going from the additively separable model to one with match fixed effects. For example, Appendix Table A9 shows that in our data going from an additively separable model to a model with arbitrary firm and worker interactions increases the adjusted R^2 from 0.90 to 0.93.

This evidence is less compelling than many have interpreted it to be. Specifically, the change in adjusted R^2 depends both on the presence of match effects and mobility patterns. To take an extreme example, suppose that there are N firms and $N-1$ workers. Worker 1 moves from firm 1 to 2, worker 2 moves from firm 2 to 3 and etc. so that worker $N-1$ moves from firm $N-1$ to N. In this case, we can separately identify firm effects for $N-1$ firms (up to an omitted firm). The R^2 of the AKM model and the unrestricted model would be the same, however, because there is only one mover to identify each relative firm effect. Put differently, with only one mover to identify the firm effect, it is not possible to tell apart firm and match effects. Thus, generically, whether it is

possible to tell apart firm and match effects depends on the mobility patterns in the data.

There is some evidence consistent with this case in the data. In Appendix Table A9, the change in adjusted R^2 using all workers is 0.029. This model includes both movers and non-movers. For non-movers, there cannot be a difference between the AKM and interacted model because there is nothing in the data that identifies match effects. If we instead restrict to only paired movers, the change in adjusted R^2 is 0.047, which is over twice the increase.

We are currently working on exploring denser parts of the mobility graph.

6.4 Heuristic identification

Figure 9 shows heuristic identification plots, which explains how certain model parameters map into the paired mover coefficient. To structure the discussion it is useful to reproduce equation (17):

$$\text{plim } \hat{\beta}_1 = \frac{Var(\alpha_i - \alpha_{i'}) + Cov(m_{iA} - m_{i'A}, m_{iB} - m_{i'B})}{Var(\alpha_i - \alpha_{i'}) + 2Cov(\alpha_i - \alpha_{i'}, m_{iA} - m_{i'A}) + Var(m_{iA} - m_{i'A})} + \frac{Cov(\alpha_i - \alpha_{i'}, m_{iB} - m_{i'B}) + Cov(\alpha_i - \alpha_{i'}, m_{iA} - m_{i'A})}{Var(\alpha_i - \alpha_{i'}) + 2Cov(\alpha_i - \alpha_{i'}, m_{iA} - m_{i'A}) + Var(m_{iA} - m_{i'A})}.$$

In the special case of no sorting, this expression reduces to $\frac{Var(\alpha)}{Var(\alpha)+Var(m)}$. Thus, the paired mover coefficient is directly proportional to the variance of match effects. Panel (a) of the Figure illustrates this point. Holding fixed our estimated parameters, we vary the variance of the match effects and plot the resulting paired movers coefficient. The Figure captures the fact that the paired movers coefficient is decreasing in the variance of the match effects.

Outside of this special case, what terms are relevant? 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 one term that can matter is $Cov(m_{iA} - m_{i'A}, m_{iB} - m_{i'B})$. This term captures the extent to which gaps in match effects between workers i and $i\prime$ are correlated across firms. This correlation 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 (b) 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. Similarly, as λ_1 decreases and workers have fewer jobs the paired mover coefficient also declines. 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.

6.5 Results

For the bottom line of the paper, the main take-away from in Table 9 is that the variance of the match effects in the offer distribution is just over half of the variance of firm effects.

The fourth part of Table 9 shows the mean and variances of the structural match effects, person effects and firm effects in the steady state distribution. The steady state distribution differs from the offer distribution because workers reject offers. Despite these rejections, the variance of the firm effects and 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 55% of the variance of the firm effects. This result quantifies the role of match effects implied by the paired mover coefficient.

To understand where this estimate of the role of match effects comes from, recall that above we showed that in a special case of no sorting, the paired mover coefficient implies that the variance of match effects would be about 40% of the variance of individual effects. Rather than calibrating to the variance of AKM person effects in the population, we have calibrated to the within firm gaps so the dispersion of person effects in our estimated model reflects the within firm dispersion. Hence, the person effects in the model reflect the no-sorting benchmark (thus, we miss on the variance of person effects). In Table 9 this implies a variance of match effects of $0.42 \times 0.074 = 0.031$. This estimate is *smaller* than what the model finds, which is 0.04. Appendix Table A11 displays the model-based components of equation (17) and shows that the term that contributes to the gap is the covariance in the gaps of match effects across firms.

These estimates are calibrated to explain the paired mover coefficient and the average gap in earnings among paired movers. As we saw in Figure 3, the paired mover coefficient is stable or decreasing in the size of the gap. So if we calibrated our model to these corners of the data, then we would find a much larger role for match effects.

Why do we care about the presence of match effects? One answer to this question is that predictions of changes in earnings based on accepted offers provides a poor guide to the changes in earnings for other workers, specifically, workers who rejected the offers. Figure 8 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). The y-axis considers for those same firm pairs 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, the presence of match effects means that the magnitude of earnings changes among accepted offers are a poor guide to the earnings changes among other potential moves (the rejected offers).

Veneto data: We also estimate a version of the model to match the Veneto results. Our estimates are more preliminary, but also suggest that the variance of match effects is much larger than the variance of firm effects. In this case, about twice as large.

6.6 Extended model with a notion of content

Our results above suggest that the paired mover test looks quite different for "larger" and "smaller" moves, and that the larger moves are relatively rarer. Combined, this evidence heuristically points to violations of exogenous mobility in that workers are taking into account match effects in mobility. Here, we develop an extended version of the model which includes a notion of sector. The goal is to estimate variance of match effects at the firm and sector level, and to show that this can rationalize both the difference in the paired mover coefficient, and the mobility patterns.

The model is 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)). We use 20 sectors. Appendix Table A12 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, fewer moves are cross-sector than a random mobility benchmark would predict. But quantitatively we are quite far off.

6.7 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. In Table 11 we report results of the paired mover and event study tests using data simulated from their model. The paired mover coefficient is 0.64, whereas it is around 0.78 in the Veneto data. The event study coefficient is 0.96.[4] Thus, while this paper develops a model that delivers an additively separable wage equation (in levels), there are meaningful deviations from additive separability in model generated data.

7 Quantitative comparison to Bonhomme, Lamadon, and Manresa (2019)

Bonhomme, Lamadon, and Manresa (2019) (BLM) develops methods designed to detect deviations from additive separability. In this section, we show that the BLM approach does not reliably detect the deviations that our paired mover test detects.

7.1 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 some of its key assumptions: first, that mobility can depend on the type of the worker and the class of the firm,

[4]Interestingly, this model generates a small increase in R^2 from adding interactions. We believe that this is because the mobility network generated by this model is much less dense than that in our model.

but not directly on earnings. Put differently, mobility cannot depend on the error term. Second, earnings after a job move do not depend on the worker's previous earnings so that only the current firm class is relevant. Where the static model is richer than the AKM model is that it allows for interactions in the wage equation. Using notation parallel to equation (2) the wage equation is:

$$y_{ijt} = \alpha_i + \psi_{j(i,t)} + g(\alpha_i, \psi_{j(i,t)}) + \epsilon_{ijt}, \tag{22}$$

where the i is now overloaded and refers to the type of worker i, the j is overloaded and refers to the class of firm j, $g(\cdot,\cdot)$ is arbitrary interactions of worker type and firm class, and the ϵ is independent of worker type and firm class.

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 a set of worker-type specific transition probabilities (a transition matrix) across firm classes. We compute the steady state distribution implied by this transition matrix. From this steady state distribution, we then simulate a set of moves. By firm class, we then pair workers up. To construct wages, we use the fact that the estimation also gives us an additive error term and we draw from this error term to construct wages. 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.

Based on the estimates, we simulate a dataset. In this simulated dataset, we then run the paired movers regression.

We estimate the BLM model in two different ways. 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 to the paired mover test, in that the BLM estimator is looking at the exact same data as the paired movers test.

7.2 The paired movers test in simulated data

Brazilian data: Column (1) of Table 12 shows that we find a paired mover coefficient of 0.69 when we use all movers in the BLM estimator. This number is virtually identical to what we find in the data. This finding is encouraging for BLM in that it is suggests that the deviations from additive separability that the paired mover test detects are detected by BLM.

Column (2) of Table 12 shows that we find a paired mover coefficient of 0.92 when we use only paired movers in the BLM estimator. This number is quite far from what we find in the data. This finding is discouraging for BLM in that it suggests it is not picking up the deviations from additive separability that are in the data.

Veneto data: Panel B of Table 12 repeats the same two exercises in the Veneto data. The pattern is the opposite from the Brazilian data, where the paired mover coefficient is "close" to 1, 0.92, when the BLM estimator uses all movers, while it is 0.50 (and thus far below what is in the data) when we use only paired movers in the BLM estimator.

Discussion: Combined, these four estimates suggest that the BLM estimator does not reliably detect the deviations from additive separability that the paired mover test detects. We interpret this evidence as indicative that the deviations from additive separability that the paired mover test documents are different from what the literature has previously found.

8 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 the presence of match effects. These match effects represent a deviation from the wage equation that is additively separable in firms and workers that is popular in empirical work. We formulate this paired mover test as a forecast bias test.

We study these paired movers in administrative data from two countries and find meaningful deviations from additive separability in the sense that the forecast bias coefficient is well below 1. Strikingly, there are larger deviations for statistically rarer moves, which are those across "contexts": sector, firm-pay level, cluster or occupations. This link between match effects and mobility is consistent with deviations from exogenous mobility.

To translate the forecast bias test into economically interpretable estimates, we estimate a simple model of on the job search and treat the forecast coefficient as a moment to be matched. We find that the variance of match effects is over half the variance of firm effects. Despite the fact that the model violates exogenous mobility, data simulated from the model passes the event study test, which is typically taken as evidence in favor of exogenous mobility. An implication of the important role for match effects is that extrapolating from the earnings changes we see in the data to the earnings changes that other workers would experience if they made similar moves is misleading. Specifically, the earnings changes on accepted offers are a different sign than those on rejected offers, and the differences in earnings changes across accepted offers are a poor guide to those on rejected offers.

We show that the deviations from additive separability we document are distinct from those captured by the Bonhomme, Lamadon, and Manresa (2019) approach. Specifically, we estimate their model on our data. In data simulated from their model, we do not reliably detect deviations from additive separability using our paired mover test.

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, 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, Francesco Devicienti, and Agata Maida. 2014. "Rent-sharing, holdup, and wages: Evidence from matched panel data." *Review of Economic Studies* 81 (1):84–111.

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 D Saggio, and Stephen A Woodbury. 2023. "Do firm effects drift? Evidence from Washington administrative data." *Journal of Econometrics* 233 (2).

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.

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
Tenure (months)	46.97	23.27	38.13
Avg. log earnings	2.34	2.19	2.39
# of Worker-Years	555,346,010	79,856,352	3,962,382
Avg. firm size	14.08	27.21	78.82
Avg. firm pay	2.41	2.45	2.47
# of Firm-Years	46,405,377	18,220,954	883,078
B. Veneto			
Avg. age	34.17	30.14	33.91
Share female	0.37	0.37	0.38
Avg. log earnings	10.10	9.64	9.99
# of Worker-years	15,158,383	1,897,250	10,508
Avg. (origin) firm size	8.35	15.79	263.82
Avg. (origin) firm pay	10.10	10.11	10.15
# of (origin) Firm-years	1,815,553	676,890	3,350

This table shows descriptive statistics of different subsamples of the Brazilian (Panel A) and Veneto (Panel B) data, described in Section 3. Column 1 displays all workers in the data regardless of mobility status and reports statistics over all employment spells. Columns 2 and 3 display different groups of workers who change firms and report statistics of the spell prior the move in the origin firm. Column 2 shows all movers in the data and Column 3 shows the sample of paired movers that we use in estimation, which matches on age and gender. 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. As described in the text, in the Brazilian data we keep only one pair per worker per moving event in the paired movers sample, while in the Veneto data we form all possible pairs due to the small number of movers.

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

		% of total moves							
	N	Sector Stayers	Sector Movers	Cluster Stayers	Cluster Movers	Low - Low	Low - High	High - High	High - Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Brazil									
Paired Movers	1,981,191	45.5	54.5	31.7	68.3	36.5	13.5	36.5	13.5
All Movers		33.3	66.7	22.7	77.2	33.1	16.8	33.7	16.4
Eligible Movers		48.5	51.5	32.5	67.5	38.4	11.6	38.6	11.4
A. Year-specific distributions									
Simulated Movers	1,402,411	47.5	52.5	32.1	67.9	35.5	12.2	36.0	11.7
p5 (200 draws)	1,402,132	47.5	52.5	32.1	67.9	35.5	12.2	36.0	11.7
p95 (200 draws)	1,403,071	47.5	52.5	32.1	67.9	35.5	12.2	36.0	11.7
B. Single distribution									
Simulated Movers	1,233,808	48.6	51.4	32.7	67.2	31.7	10.8	32.2	10.4
p5 (200 draws)	1,232,589	48.5	51.3	32.7	67.2	31.7	10.8	32.2	10.4
p95 (200 draws)	1,234,106	48.7	51.5	32.8	67.3	31.7	10.8	32.3	10.5
C. Random Benchmark									
Paired Movers		3.8	96.2	9.8	90.2	25.0	25.0	25.0	25.0
All Movers		6.0	94.0	5.0	95.0	25.0	25.0	25.0	25.0
Eligible Movers		4.0	96.0	9.8	90.2	25.0	25.0	25.0	25.0
B. Veneto									
Paired Movers	8167	60.7	39.3	27.1	70.4	36.5	13.3	36.7	13.5
All movers									
D. Year-specific distributions									
Simulated Movers	6804	57.7	42.3	26.5	71.8	37.1	12.7	37.4	12.7
p5 (200 draws)	6627	56.9	41.2	25.6	70.8	36.7	12.5	36.7	12.5
p95 (200 draws)	7025	58.8	43.1	27.3	72.6	37.5	13.3	37.6	13.3
E. Single distribution									
Simulated Movers	4525	59.0	41.0	23.2	74.2	33.9	16.0	34.1	16.0
p5 (200 draws)	4406	57.3	40.2	22.7	72.8	32.9	15.6	33.0	15.6
p95 (200 draws)	4577	59.8	42.7	24.7	74.8	34.4	17.1	34.4	17.1
F. Random benchmark									
Paired movers		10.1	89.9	12.3	87.7		25.2		23.4
All movers		7.5	92.5	9.6	90.4		25.2		24.7

This table shows sample sizes of paired movers that would be obtained by chance as implied by the mobility patterns in the data. We take the sample of eligible movers (as defined in Section 3) and drop workers in the final paired movers sample (Column 3 in Table 1). We then compute, for each origin firm, the empirical distribution of the 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 as we do with the real data. In Panel A, we do these steps year by year, while in Panel B we use a single distribution for each origin firm, pooling years. The table shows the number of paired movers in the simulated samples, as well as the 5th and 95th percentiles of the distribution of the number of paired movers across 200 draws. In Column 1 we show the number of paired moves obtained, and columns 2 to 9 show the share of these moves that are within or across sectors, clusters, and pay levels. The first row shows the actual sample of paired movers in the data, while the second row shows the sample of all movers. In Panel C, we compare those with a random benchmark formed by either the sample of paired movers or all movers. The benchmark is formed as follows. For example, for sectors, we compute the likelihood in the data of a move being from each sector and the likelihood of a move being to each sector. The probability of a move being within sector is the product of these two probabilities summed over sectors.

Table 3: Gap distribution

	Mean (1)	Variance (2)	p10 (3)	p90 (4)
A. Brazil				
Paired movers	0.364	0.167	0.032	0.865
Benchmark	0.470	0.227	0.047	1.095
Paired stayers	0.450	0.210	0.043	1.043
B. Veneto				
Paired movers	0.335	0.198	0.019	0.788
Benchmark	0.419	0.242	0.036	0.983
Paired stayers	0.337	0.150	0.027	0.806

This table shows moments of the distribution of the (absolute) gap in earnings between paired movers at their origin firms in Brazil (Panel A) and Veneto (Panel B). The first row shows the distribution of the gap in the sample of paired movers. The second row shows the distribution of the gap in the sample of workers randomly paired within the same firm and year, conditioning on workers having at least two years of tenure at the origin firm and one more observation after that (Benchmark 2). The third row shows the distribution of the gap in the sample of paired stayers.

Table 4: The Paired Mover Test

					Gap in Origin					
	Baseline	Stayers	Within Sector	Across Sector	Low - Low	Low - High	High - High	High - Low	Cluster Stayers	Cluster Switchers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Brazil, 2nd Stage										
Gap in Origin	0.699	0.967	0.733	0.666	0.709	0.681	0.718	0.619	0.740	0.683
	(0.0007)	(0.0003)	(0.0010)	(0.0010)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Observations	1,981,191	1,302,157	900,977	1,080,214	723,590	267,048	723,497	267,056	418,334	901,398
p-value, equality test			0.0000		0.0000		0.0000		0.0000	
B. Brazil, 1st Stage										
Gap in Origin (lag)	0.925	0.941	0.925	0.924	0.920	0.915	0.929	0.922	0.934	0.931
	(0.0005)	(0.0003)	(0.0007)	(0.0007)	(0.0010)	(0.001)	(0.0008)	(0.002)	(0.001)	(0.0008)
F-statistic	3,133,829.8	10,611,459.4	1,590,750.9	1,545,680.3	931,451.9	378,268.1	1,345,544.8	361,093.4	793,000.8	1,440,341.2
C. Veneto, 2nd Stage										
Gap in origin	0.776	0.930	0.912	0.576	0.895	0.504	0.812	0.613	1.03	0.698
	(0.035)	(0.0004)	(0.044)	(0.057)	(0.060)	(0.084)	(0.049)	(0.129)	(0.074)	(0.040)
p-value, equality test			0.0000		0.0000		0.1329		0.0000	
Observations	8,167	8,577,101	4,960	3,207	2,866	1,062	3,133	1,106	2,211	5,751
D. Veneto, 1st Stage										
Gap in Origin (lag)	0.343	0.572	0.357	0.325	0.309	0.356	0.399	0.340	0.347	0.347
	(0.009)	(0.0002)	(0.011)	(0.015)	(0.013)	(0.030)	(0.015)	(0.026)	(0.017)	(0.011)
F-statistic	1,445.2	7,348,531.4	1,033.5	460.02	519.97	143.78	692.81	169.41	436.84	996.07

This table shows different versions of the paired movers test. We estimate a 2SLS regression where the log earnings gap in the destination firm is regressed on the log earnings gap in the origin firm, instrumented by the lagged earnings gap in the origin firm. Panels A and B show the second and first stages of the test in Brazilian Data, respectively. Panels C and D show the second and first stages in Veneto data. 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 6 to 9 show the estimate using the sample of paired movers across different pay bins. We define low/high pay as the bottom/top 50% of the firm pay distribution among firms in the estimation sample. Columns 9 and 10 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). For Columns 3 to 10, we report the p-value of the equality test of the coefficients across columns for the relevant heterogeneity.

Table 4: The Paired Mover Test in Brazilian Data (cont.)

	None change ocup (11)	One change ocup (12)	Both change ocup (13)	Both low tenure (14)	One low tenure (15)	Both high tenure (16)	CZ Stayers (17)	CZ Switchers (18)
A. 2nd Stage								
Gap in Origin	0.768	0.723	0.602	0.706	0.680	0.716	0.700	0.698
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0008)	(0.001)
Observations	746,586	532,681	701,924	787,939	626,757	566,495	1,505,131	476,060
B. 1st Stage								
Gap in Origin (lag)	0.907	0.943	0.919	0.906	0.927	0.940	0.928	0.914
	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0009)	(0.0009)	(0.0006)	(0.001)
F-statistic	1,026,008.2	1,140,317.1	930,876.4	867,467.7	1,125,726.9	1,206,425.1	2,433,827.8	702,763.0

This table shows different versions of the paired movers test in Brazilian data. We estimate a 2SLS regression where the log earnings gap in the destination firm is regressed on the log earnings gap in the origin firm, instrumented by the lagged earnings gap in the origin firm. Panel A shows the second stage of the test, while Panel B shows the first stage. Columns 11 to 13 show the estimate using the sample of paired movers split by whether none, one or both workers change occupation (3-digit CBO code). Columns 14 to 16 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 17 and 18 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). For Columns 11 to 18, we report the p-value of the equality test of the coefficients across columns for the relevant heterogeneity.

Table 5: Paired Mover Test by Reason of Separation and Time to Re-employment

	Gap in Destination							
	Involuntary				Voluntary			
	Full sample	Within 1 month	Btw. 2–6 months	After 6 months	Full sample	Within 1 month	Btw. 2–6 months	After 6 months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Second Stage								
Gap in Origin	0.671	0.801	0.699	0.620	0.736	0.743	0.692	0.690
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.009)	(0.009)
Observations	835,395	131,395	149,509	554,491	145,425	113,446	13,329	18,650
B. First Stage								
Gap in Origin (lag)	0.919	0.913	0.911	0.923	0.910	0.914	0.870	0.894
	(0.0009)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.006)	(0.007)
F-statistic	1,116,888.0	427,650.5	300,665.8	573,687.0	361,198.2	345,770.1	22,587.7	16,521.8

This table shows different versions of the paired movers test by reason of separation and time to re-employment. We estimate a 2SLS regression where the log earnings gap in the destination firm is regressed on the log earnings gap in the origin firm, instrumented by the lagged earnings gap in the origin firm. Panel A shows the second stage of the test, while Panel B shows the first stage. Columns 1 and 5 show the estimate using the sample of paired movers who did not separate with cause (Column 2 of Table A3) and additionally both workers take the same (discrete) amount of time to re-employment. Time to reemployment is defined as the difference between the hiring date at the destination firm and the separation date at the origin firm. Columns 1 and 5 show the results for involuntary and voluntary separations, respectively, benchmarking the results in Columns 3 and 4 of Table A3. Columns 2–4 and 6–8 show the results for involuntary and voluntary separations, respectively, split by whether both workers found a new job within one month of the month of separation between 2 and 6 months after separation, or after 6 months of separation.

Table 6: 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 $\hat{\psi}_1$	1.07	1.05	0.977	0.992	0.987
	(0.001)	(0.003)	(0.001)	(0.001)	(0.002)
Observations	3,574,659	390,178	3,124,563	3,158,568	900,960
B. Brazil, 1st Stage					
Change in $\hat{\psi}_2$	0.761	0.815	0.884	0.882	0.879
	(0.0003)	(0.0009)	(0.0003)	(0.0003)	(0.0005)
F-statistic	4,845,767.8	762,862.2	10,880,588.9	10,872,954.4	3,110,841.6
C. Veneto, 2nd Stage					
Change in $\hat{\psi}_1$	1.14	0.727	0.945	0.904	0.720
	(0.005)	(0.033)	(0.008)	(0.007)	(0.031)
Observations	490,693	5,972	116,711	137,500	21,959
D. Veneto, 1st Stage					
Change in $\hat{\psi}_2$	0.469	0.679	0.740	0.722	0.778
	(0.001)	(0.009)	(0.002)	(0.002)	(0.004)
F-statistic	143,994.6	5,344.9	138,337.3	135,006.7	42,479.4

This table shows the results of the standard event study test using AKM estimates. We first estimate two sets of firm fixed effects by randomly splitting the full data (Column 1, Table 1) in two samples. We then 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 the first sample to instrument the firm effects in the second sample. 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 7: 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	0.83	0.804	0.768	0.732	0.693	0.794	0.778	0.744	0.717	0.686
-1		0.956	0.906	0.862	0.813		0.947	0.884	0.842	0.803
-2			0.934	0.887	0.837			0.921	0.872	0.831
-3				0.929	0.875				0.92	0.87
-4					0.923					0.921
B. Origin Firm, AR(1) coef.										
Time lag	-1	-2	-3	-4	-5	-1	-2	-3	-4	-5
0		0.969	0.955	0.953	0.947		0.980	0.956	0.964	0.957
-1			0.948	0.951	0.943			0.933	0.952	0.954
-2				0.950	0.944				0.947	0.953
-3					0.942					0.946
C. Destination Firm, autocor.										
Time lead		2	3	4	5		2	3	4	5
1		0.918	0.871	0.835	0.778		0.847	0.815	0.789	0.767
2			0.933	0.89	0.827			0.88	0.846	0.819
3				0.934	0.865				0.893	0.86
4					0.904					0.903
D. Destination Firm, AR(1) coef.										
Time lead		2	3	4	5		2	3	4	5
1			0.949	0.959	0.932			0.962	0.968	0.972
2				0.954	0.929				0.961	0.968
3					0.926					0.963

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).

Table 8: Match Effects AR(1) Coefficients

	Full sample (1)	Destination (2)	Origin (3)
A. Balanced Panel			
k	-0.068	-0.065	-0.069
	(0.003)	(0.002)	(0.005)
$\exp(\beta)$	0.934	0.937	0.933
B. Full Sample			
k	-0.090	-0.088	-0.091
	(0.008)	(0.009)	(0.011)
$\exp(\beta)$	0.914	0.915	0.913
Observations	25	10	15

This table shows the results of the 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. The coefficient β can be interpreted as an estimate of $\ln(\nu)$ in the AR(1) process $y_t = \nu y_{t-1} + u_t$. Hence, $\exp(\beta)$ is the AR(1) coefficient of the log earnings gap. The regressed autocorrelations are the same as those shown in Table 7. Panel A shows the results for the balanced panel of paired movers, 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, respectively. All regressions are weighted by the number of observations used to calculate each autocorrelation. Standard errors are shown in parentheses.

Table 9: Model results

	(1)	(2)	(3)	(4)	(5)
A. Brazil					
	EE rate	β_1	Mean gap	$Var(\psi_j)$	$Var(\alpha_i)$ (untargeted)
Empirical Moments (Data)	0.038	0.699	0.364	0.071	0.271
Simulated Moments (Model)	0.040	0.701	0.380	0.073	0.074
	λ_0 (UE rate)	δ (EU rate)			
Calibrated Parameters	0.072	0.076			
	λ_1	$Var(m_{ij})$	$Var(\alpha_i)$	$Var(\psi_j)$	
Estimated Parameters	0.121	0.041	0.076	0.083	
Steady-state (employed)		m_{ij}	α_i	ψ_j	
Mean		0.069	0.000	0.111	
Variance		0.041	0.076	0.074	
Correlation (m_{ij},ψ_j)	-0.000				
B. Veneto					
	EE rate	β_1	$Var(\alpha_i)$	$Var(\psi_j)$	
Empirical Moments (Data)	0.0245	0.776	0.3174	0.1259	
Simulated Moments (Model)	0.0259	0.7433	0.2381	0.0953	
	λ_0	δ			
Calibrated Parameters	0.4000	0.0150			
	λ_1	$Var(m_{ij})$	$Var(\alpha_i)$	$Var(\psi_j)$	
Estimated Parameters	0.2389	0.3025	0.2669	0.1622	
Steady-state (employed)		m_{ij}	α_i	ψ_j	
Mean		0.7072	-0.0018	0.3686	
Variance		0.2377	0.2709	0.1427	

This table shows the results of the model estimation. We estimate the model using the simulated method of moments (SMM) on the empirical moments of the data. We fix the λ_0 and δ parameters to quarterly UE and EU rates (we use PNAD for Brazil). We estimate the four remaining parameters of the model, λ_1, $Var(m_{ij})$, $Var(\alpha_i)$, and $Var(\psi_j)$, to match empirical moments of the data: the EE rate, the β_1 coefficient in the paired movers test, and the variances of the firm and worker effects (Column 1 of Table A9). We search the parameter space using a Metropolis-Hastings algorithm with 10 chains of 250 iterations each. At each iteration, we simulate 5000 workers and 100 firms over 200 periods. The first two rows in each panel show the empirical moments in the data and the simulated moments in the estimated model. The next two rows show the estimated and calibrated parameters of the model. The last two rows show the steady-state moments of employed workers in the model.

Table 10: Event Study and Paired Mover Tests in Model Simulated Data

	Paired Mover test			Event Study test				
	All movers (1)	EE moves (2)	EUE moves (3)	All movers (4)	EE moves (5)	EUE moves (6)	Accepted (7)	Rejected (8)
A. Brazil								
Gap in origin	0.701	0.788	0.653					
	(0.0007)	(0.0010)	(0.0010)					
Change in $\hat{\psi}$ (IV)				0.991	0.563	0.988		
				(0.0007)	(0.001)	(0.0009)		
Observations	1,019,161	376,104	643,057	998,366	411,615	586,751		
B. Veneto								
Gap in origin	0.743	0.793	0.559					
	(0.007)	(0.007)	(0.022)					
Change in $\hat{\psi}$ (IV)				1.05	0.333	0.941	0.333	0.681
				(0.011)	(0.008)	(0.020)	(0.008)	(0.004)
Observations	8,418	6,466	1,952	17,704	12,361	5,343	12,361	102,533

This table shows the paired movers test (Columns 1 to 3) and the AKM-based event study test (Columns 4 to 8) in data simulated from the estimated model (Table 9). Since there is no measurement error in the model, we simply run OLS for the paired movers test. Column 1 uses all movers in the data, while columns 2 and 3 split moves by whether they are EE (workers accept an outside offer while employed) or EUE (workers accept an outside offer while unemployed). Columns 4 to 8 show the event study test, where we estimate a TSLS regression where we estimate firm effects in two randomly split samples of the data and run a 2SLS regression as in Table 6. Column 4 shows the results for all movers, while Columns 5 and 6 show the results for EE and EUE moves, respectively. Columns 7 and 8 show the results for accepted and rejected offers while employed, respectively. By construction, Columns 5 and 7 report the same regression.

Table 11: Simulated Borovičková and Shimer (2024) data

	(1)	(2)
A. Paired Mover and Event Study Test		
Gap in origin	0.622	
	(0.002)	
Change in $\hat{\psi}$ (IV)		0.958
		(0.002)
Observations	180,560	2,516,273
B. AKM estimation		
R^2	0.7825	
Adj. R^2	0.7388	
Adj. R^2, interactions	0.7600	
Observations	3,016,548	
# Firms	5,000	
# Workers	500,000	
# Movers	487,614	
# of Moves	975,580	

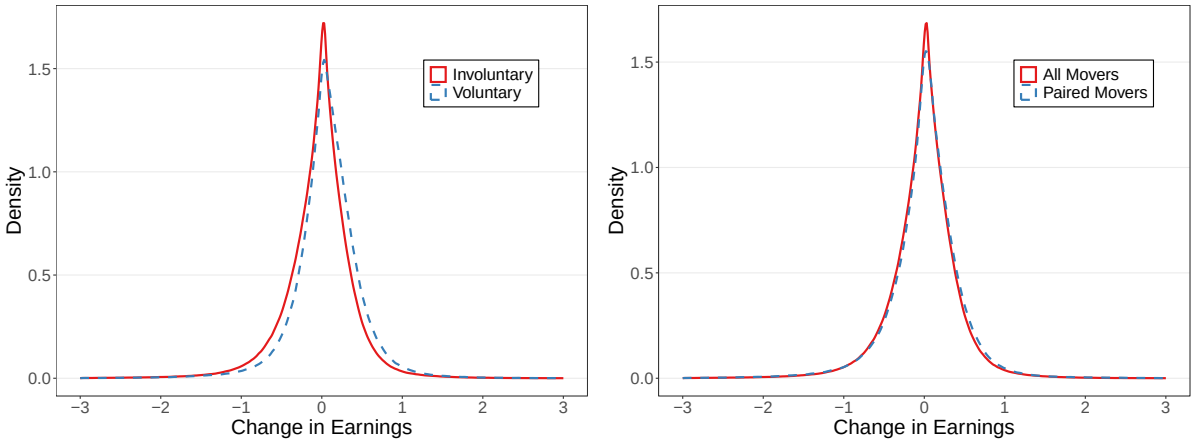
This table shows the results using data simulated from the model of Borovičková and Shimer (2024), with the same parameters as estimated in the original paper. In Panel A, we report the paired movers and event study tests. In Panel B, we report the (adjusted) R^2 of the AKM estimation, and the adjusted R^2 when we include interactions between the worker and firm effects. We also report the number of observations, firms, workers, movers, and moves in the simulated data.

Table 12: Paired Movers in Simulated BLM data

	Gap in origin / Gap in destination	
	All movers	Paired Movers
	(1)	(2)
A. Brazil, 2nd Stage		
Gap in origin	0.690	0.916
	(0.001)	(0.001)
Observations	1,760,034	2,052,400
B. Brazil, 1st Stage		
Gap in origin (lag)	0.628	0.625
	(0.0006)	(0.0005)
F-statistic	1,139,765.2	1,374,285.1
C. Veneto, 2nd Stage		
Gap in origin	0.923	0.503
	(0.002)	(0.001)
Observations	1,353,802	1,780,044
D. Veneto, 1st Stage		
Gap in origin (lag)	0.494	0.547
	(0.0008)	(0.0006)
F-statistic	415,355.9	776,818.1

This table shows the results of the paired mover test in paired movers datasets simulated from the Bonhomme, Lamadon, and Manresa (2019) model estimated on the data. After estimating the model, we compute the set of worker-type specific transition probabilities across firm classes and obtain the steady-state distribution implied by the model. We then simulate a panel dataset of workers and firms using the estimated parameters and the steady-state distribution. Finally, we construct a paired movers sample from the simulated data using the same procedure as in the data. We estimate a TSLS regression where the log earnings gap in the destination firm is regressed on the log earnings gap in the origin firm, instrumented by the lagged earnings gap in the origin firm. Column 1 shows the estimate using the model estimated on the full data. Column 2 shows the estimate using the model estimated with only movers in the paired movers sample.

Figure 1: Change in Earnings around the Move, Brazil

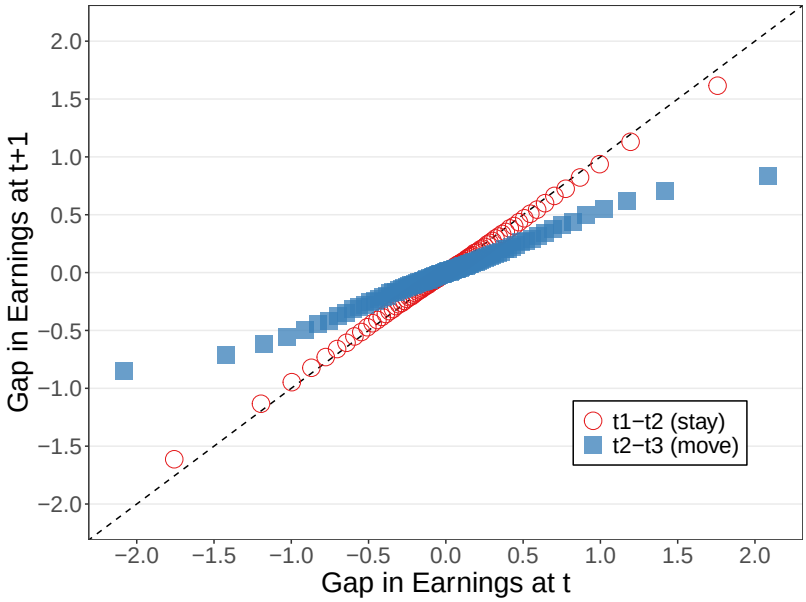


(a) Voluntary *vs.* Involuntary, Distribution

(b) All *vs.* Paired Movers, Distribution

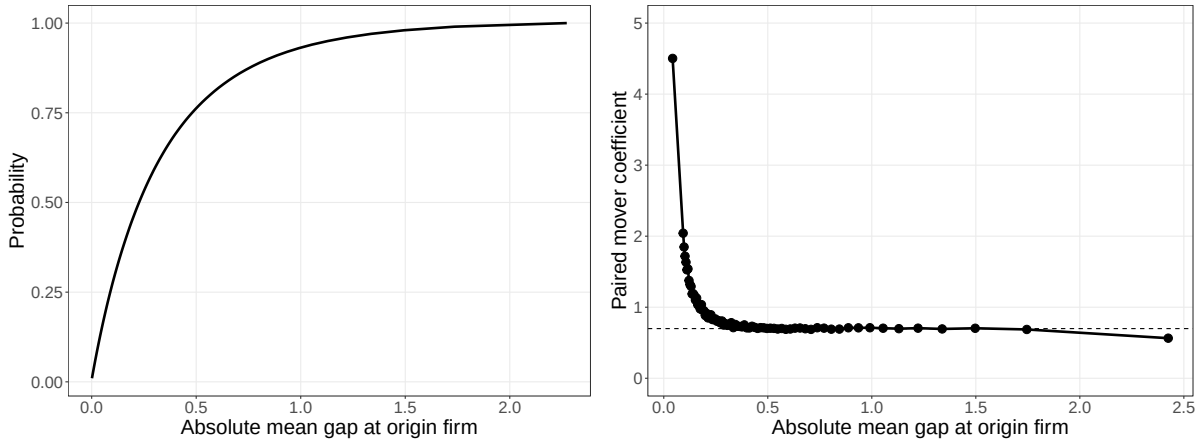
This figure shows the distribution of the change in log earnings between the destination and origin firms for different samples of movers. Panel A shows the distribution for voluntary (employee-initiated) and involuntary (employer-initiated) moves. Panel B shows the distribution for all movers (Column 2, Table 1) and the final sample of paired movers (Column 5, Table 1).

Figure 2: Changes in Gap in Earnings, Paired



This figure shows the plot version of the paired movers test. We plot the gap in earnings at t in the horizontal axis and the gap in earnings at $t+1$ in the vertical axis. We split the gap in earnings at t into 100 bins and plot the average gap in earnings at $t+1$ for each bin. The red circles show the first stage of the test, where we instrument the gap in earnings at the origin with the lagged gap in earnings at the origin. The blue squares show the reduced form of the test, where we regress the gap in earnings at the destination on the gap in earnings at the origin. The dashed line shows the 45-degree line.

Figure 3: Paired Movers Coefficient across the Gap Distribution

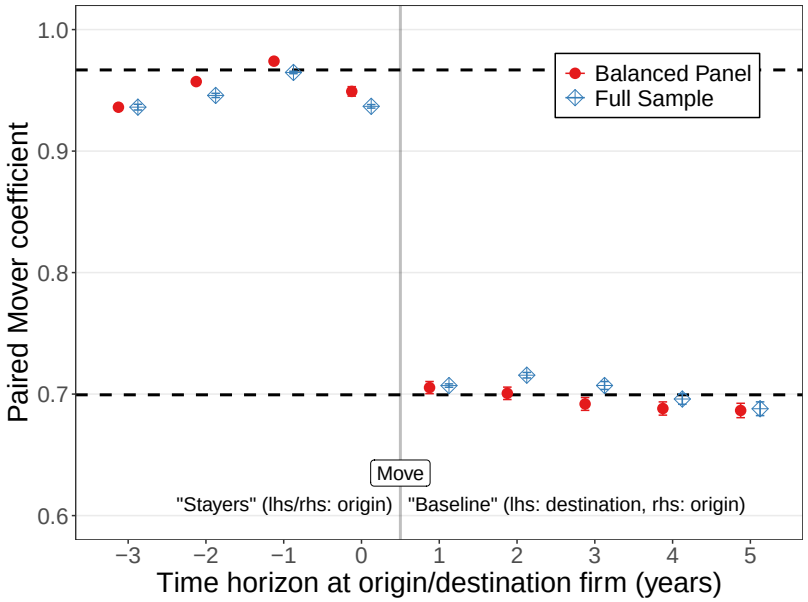


(a) Gap in earnings distribution

(b) Paired movers coefficient

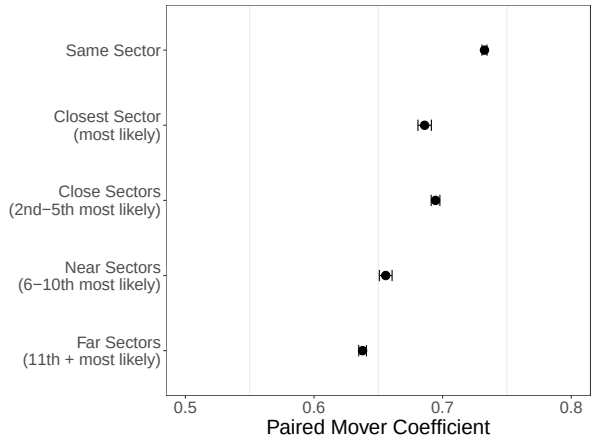
This figure shows the paired movers coefficient across the distribution of the gap in earnings at the origin firm. Panel A shows the cumulative distribution of the absolute gap in earnings at the origin firm at $t = 0$ (last year at the origin). Panel B shows the paired movers coefficient estimated using only movers within a given percentile of the gap in earnings at the origin firm. The first 25 percentiles are combined into a single bin, and then we show the coefficient for each subsequent percentile up to the 99th percentile. The dashed horizontal line shows the baseline estimate from Column 1 of Table 4.

Figure 4: Temporal Stability

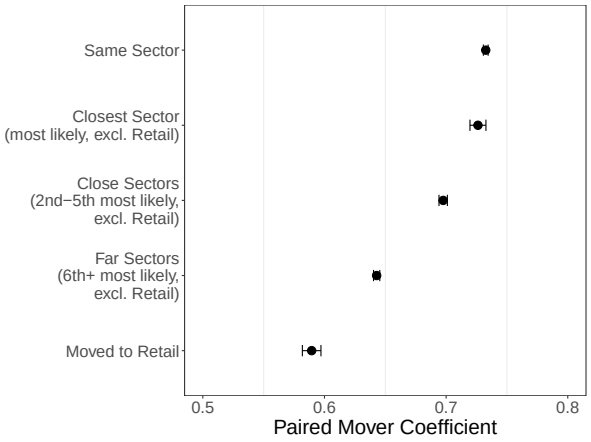


This figure shows different versions of the paired movers test. The dashed horizontal line shows the baseline estimate from Column 1 of Table 4. The dotted horizontal line shows the estimate from Column 2 of Table 4 ("stayers" placebo). The gray vertical line marks the timing of the move between origin and destination firms. To the left of the vertical line, the sample is paired movers who move between 2012 and 2019 that are observed for at least three years at the origin firm. We report the second-stage coefficient of a regression of the gap in earnings at the origin at t on the gap in earnings at the origin at $t-1$ instrumented by the lagged gap in earnings at the origin at $t-2$. Time t runs from $t=0$ (last year at the origin) to $t=-3$. Thus, these are "stayers" regressions similar to Column 2 of Table 4. To the right of the vertical line, the sample is paired movers who move between 2012 and 2019. Here, we report the second-stage coefficient of the baseline paired movers regression: the outcome is the gap in earnings at the destination at t and the endogenous variable is the gap in earnings at the origin at $t=0$, instrumented by the lagged gap in earnings at the origin at $t=-1$. At each period from $t=1$ (first year at the destination) to $t=5$, we change the outcome variable to the gap in earnings at the destination at t. At each period, both prior and post the move, 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).

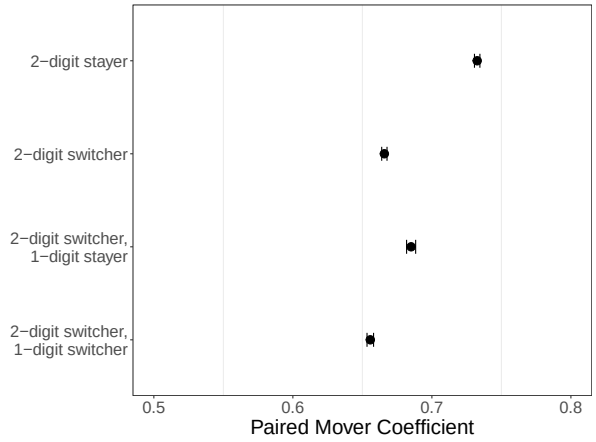
Figure 5: Large *vs* Small Moves



(a) By move likelihood



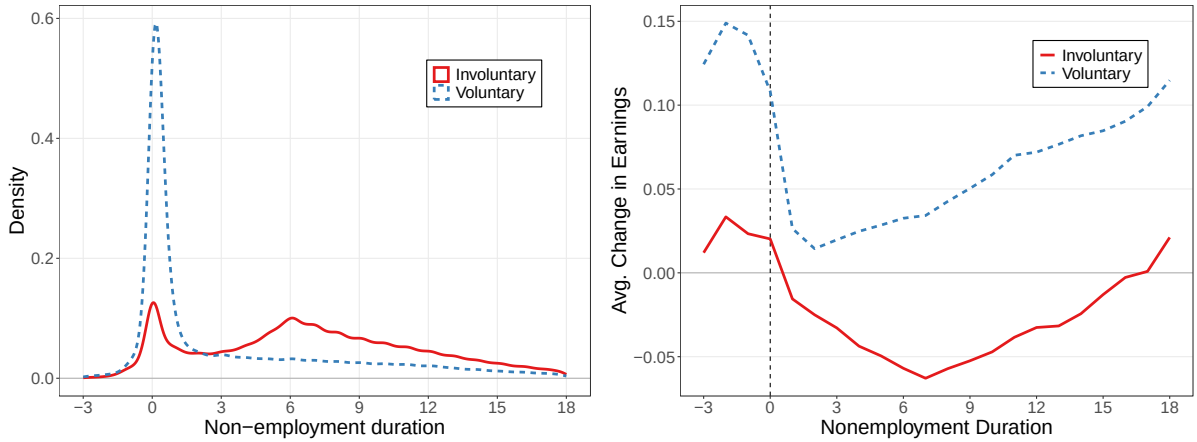
(b) By move likelihood, excl. Retail



(c) By move "distance"

This figure shows the second stage of the paired movers test for different types of moves in the Brazilian data. 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. In Panel C, we further split moves by the "distance" between sectors, distinguishing moves across 2-digit sectors that are within the same 1-digit sector or not.

Figure 6: Voluntary *vs* Involuntary Moves

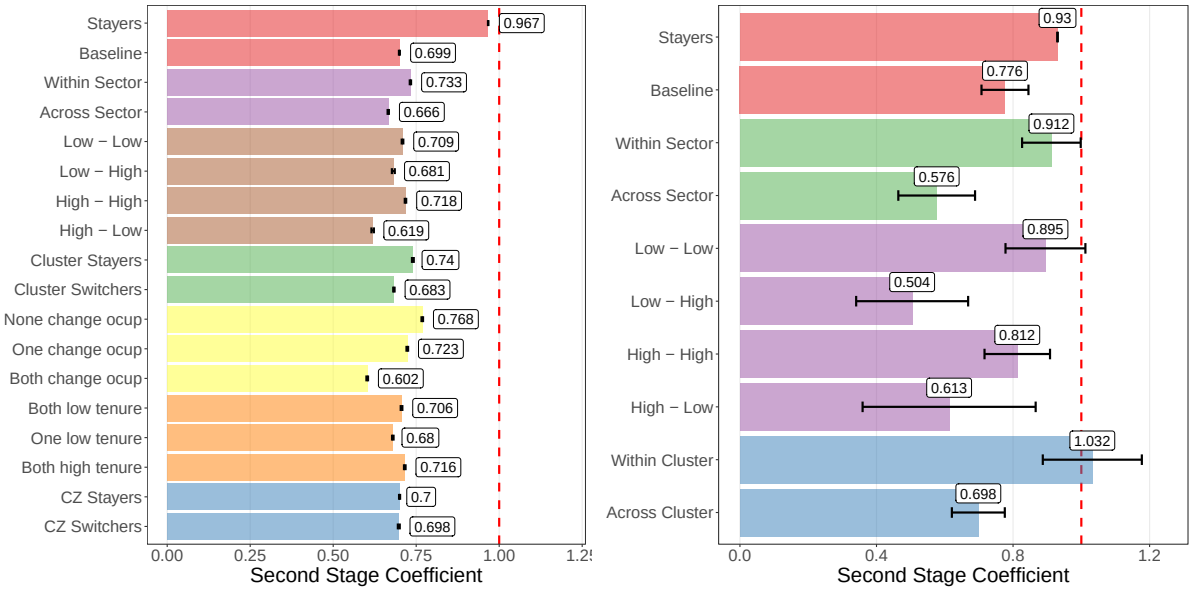


(a) Time between jobs

(b) Change in log earnings

This figure shows results distinguishing between voluntary (employee-initiated) and involuntary (employer-initiated) moves. Panel A shows the distribution of the time between jobs for voluntary and involuntary moves. Time between jobs is defined as the difference between the hiring date at the destination firm and the separation date at the origin firm. The sample is all paired movers who did not separate with cause (Column 2, Table A3). Panel B shows the average change in log earnings between the destination and origin firms for voluntary and involuntary moves, according to the time between jobs. The sample is defined as in Panel A.

Figure 7: Paired movers test

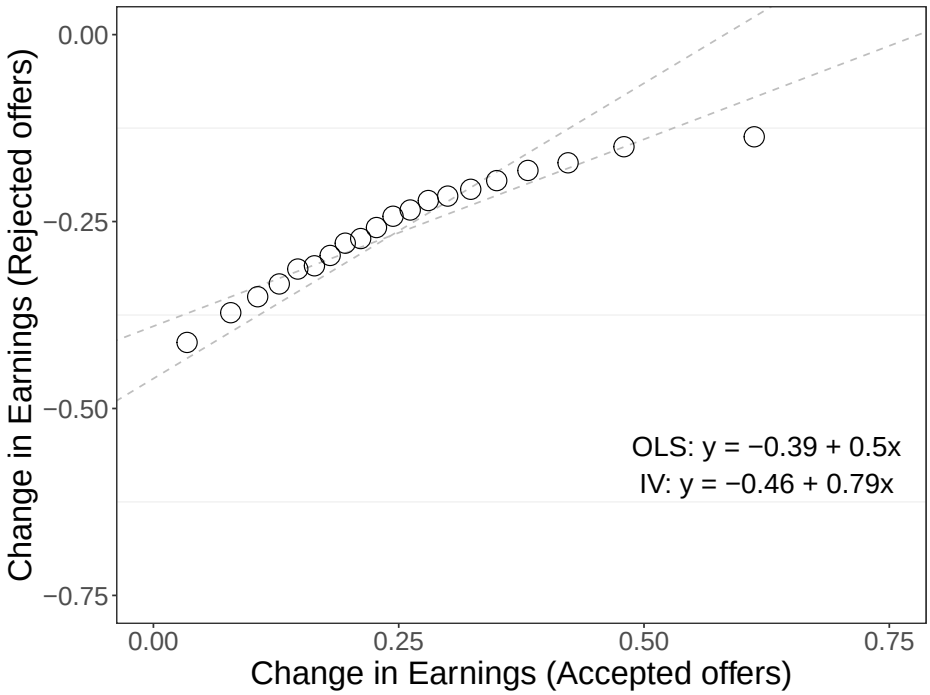


(a) Brazil, Baseline

(b) Veneto, Baseline

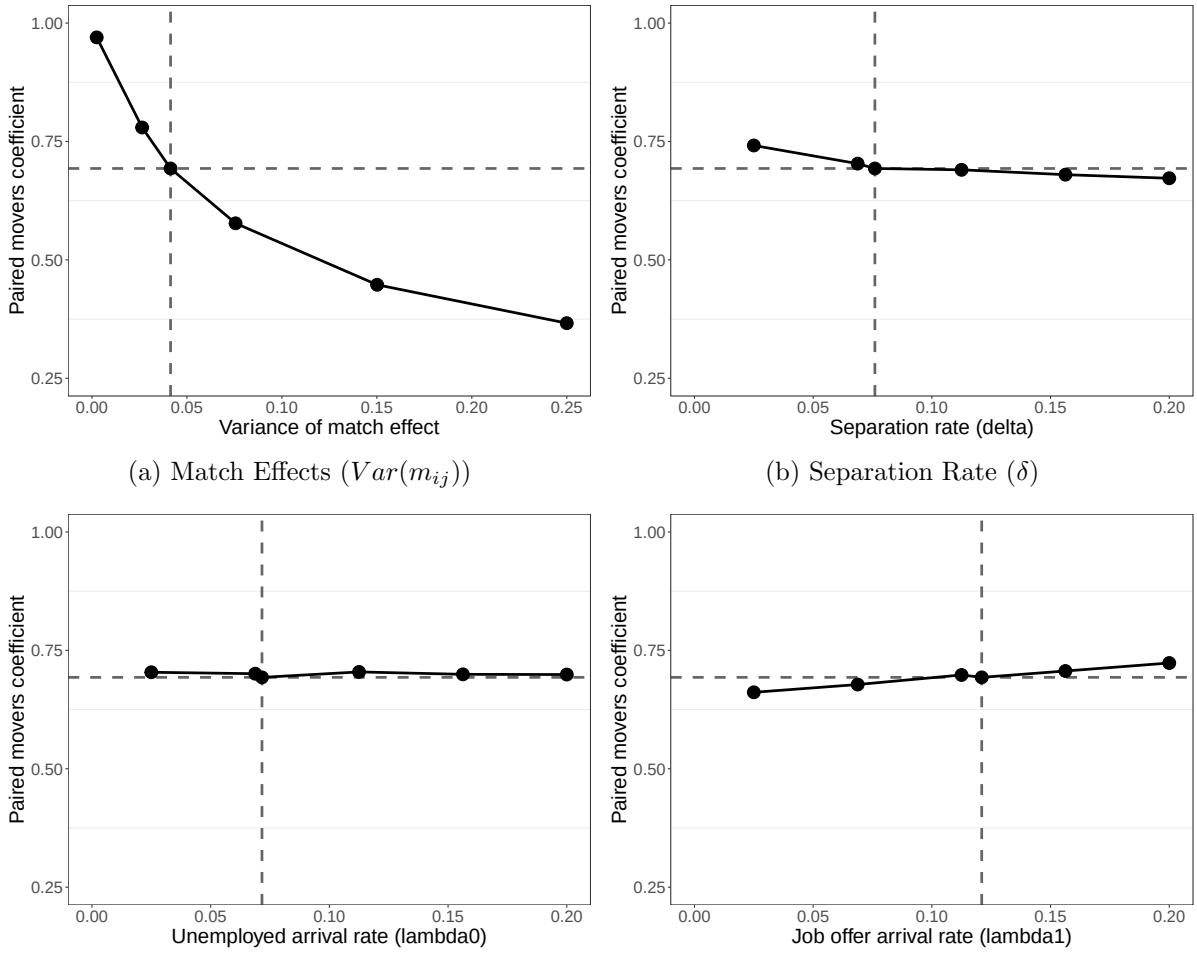
This figure shows the coefficient of the second stage of the paired movers test in Brazil (Panel A) and Veneto (Panel B). The vertical axis shows different samples reported in Table 4.

Figure 8: Model Simulated Data: Accepted vs. rejected offers



This figure shows results using data simulated from the model estimated in Brazilian data. We match workers who moved from firm i to firm j with workers who were in firm i and received an offer from firm j but did not accept it. We compute the change in log earnings between the destination and origin firms for accepted and rejected offers and take the means for each origin-destination pair. We then split the data in 20 bins based on the change for accepted offers and plot the average change for rejected offers in each bin. We also report the slope of the regression line in the underlying individual-level data. The IV coefficient comes from splitting the sample randomly.

Figure 9: Paired Movers coefficient and model parameters



(a) Match Effects ($Var(m_{ij})$)

(b) Separation Rate (δ)

(c) unemployed arrival rate (λ_0)

(d) On-the-job arrival rate (λ_1)

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.

Table A1: Summary statistics: paired movers

	All pairs (1)	Sector stayers (2)	Sector movers (3)	Low-Low (4)	Low-High (5)	High-High (6)	High-Low (7)	Cluster stayers (8)	Cluster movers (9)
A. Brazil									
Avg. age	32.37	33.02	31.84	32.21	31.26	32.72	33.01	32.59	31.97
Share female	0.31	0.34	0.29	0.38	0.31	0.25	0.27	0.30	0.32
Years of schooling	10.90	10.89	10.90	10.16	10.81	11.83	10.48	11.30	10.91
Tenure (months)	38.13	39.47	37.00	33.59	34.75	43.37	39.58	41.31	38.13
Avg. log earnings	2.39	2.41	2.37	1.99	2.18	2.86	2.37	2.61	2.37
Avg. log gap origin	0.35	0.36	0.35	0.27	0.31	0.46	0.36	0.40	0.35
Var. log gap origin	0.18	0.19	0.17	0.11	0.14	0.24	0.17	0.21	0.17
p90 - p10 log gap origin	0.85	0.88	0.83	0.64	0.74	1.05	0.84	0.95	0.85
# of Worker-years	3,962,382	1,801,954	2,160,428	1,447,180	534,096	1,446,994	534,112	836,668	1,802,796
Avg. origin firm size	78.82	98.83	107.68	68.62	117.02	144.43	174.41	181.25	121.35
Avg. origin firm pay	2.47	2.48	2.49	2.09	2.14	2.75	2.61	2.56	2.49
Avg. destination firm size	77.06	95.05	106.17	67.80	180.97	138.10	111.06	175.15	117.44
Avg. destination firm pay	2.49	2.49	2.52	2.11	2.70	2.79	2.16	2.58	2.52
# of Firm-Years	1,579,661	916,911	939,525	776,810	313,236	505,591	315,767	393,130	757,674
B. Veneto									
Avg. age	33.91	35.09	32.37	36.32	31.16	33.30	32.58	34.94	33.62
Share female	0.38	0.38	0.38	0.67	0.39	0.14	0.33	0.37	0.38
Avg. log earnings	9.99	10.06	9.91	9.54	9.77	10.40	10.20	10.04	9.98
Avg. log gap origin	0.33	0.33	0.35	0.36	0.37	0.30	0.35	0.33	0.34
Var. log gap origin	0.20	0.17	0.24	0.20	0.34	0.15	0.19	0.17	0.21
p90 - p10 log gap origin	0.77	0.75	0.81	0.84	0.82	0.68	0.84	0.73	0.80
# of Worker-years	10,508	5,822	4,686	3,452	1,543	3,868	1,645	2,742	7,400
Avg. origin firm size	263.82	302.68	267.00	200.77	181.47	354.83	372.77	383.00	266.39
Avg. origin firm pay	10.15	10.17	10.10	9.45	9.62	10.49	10.30	10.23	10.10
Avg. destination firm size	295.88	337.28	320.29	225.97	450.87	391.03	271.01	419.55	307.97
Avg. destination firm pay	10.23	10.25	10.24	9.57	10.48	10.57	9.86	10.29	10.22
# of Firm-years	6,495	3,585	3,418	2,048	1,184	2,522	1,259	1,846	4,881

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 to 7 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. Columns 8 and 9 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).

Table A1: Summary statistics: paired movers (cont.)

	No ocup. change (1)	1 ocup. change (2)	2 ocup. change (3)	Both Low Tenure (4)	Both High Tenure (5)	One Low Tenure (6)	CZ stayers (7)	CZ switchers (8)
C. Brazil, cont.								
Avg. age	33.98	32.59	30.50	31.41	34.22	31.92	32.00	33.55
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	11.35	11.11	11.04	10.44
Tenure (months)	36.85	39.44	38.49	16.88	67.94	37.89	39.71	33.11
Avg. log earnings	2.37	2.47	2.34	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	0.82	0.79	0.88	0.91	0.85	0.86
# of Worker-years	1,493,172	1,065,362	1,403,848	1,575,878	1,132,990	1,253,514	3,010,262	952,120
Avg. origin firm size	111.26	125.47	123.85	108.73	130.20	120.28	82.73	175.49
Avg. origin firm pay	2.48	2.50	2.50	2.45	2.53	2.49	2.47	2.52
Avg. destination firm size	108.99	122.03	120.85	111.37	122.02	119.04	80.83	172.19
Avg. destination firm pay	2.51	2.53	2.53	2.49	2.54	2.52	2.49	2.55
# of Firm-Years	762,490	638,164	706,312	779,079	650,597	696,432	1,382,675	394,431

This table shows descriptive statistics of the Brazilian paired movers sample. Columns 1 to 3 display movers split by whether they change occupation or not (3-digit CBO codes). Column 1 shows moves in which both workers in the pair remain in the same occupation, Column 2 shows moves in which one worker changes occupation and Column 3 shows moves in which both workers change occupation. Columns 4 to 6 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 4 shows pairs where both workers have low tenure, Column 5 shows pairs where one worker has low tenure and Column 6 shows pairs where both workers have high tenure. Finally, Columns 7 and 8 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 A2: Sectoral Moves in Brazil

CNAE	Category	Sector	% Total Employment	% Within Sector	Most Likely Destination	% Moves	2nd Most Likely Destination	% Moves
47	Retail	Retail	18.08	51.77	Wholesale	4.95	Restaurant	4.21
49	Transportation	Ground Transport	4.60	51.26	Retail	8.36	Wholesale	4.90
56	Accommodation	Restaurant	4.41	47.77	Retail	16.98	Services for Buildings	4.72
86	Health	Health	4.35	60.78	Retail	6.23	Associations	3.47
46	Retail	Wholesale	4.21	25.51	Retail	17.89	Ground Transport	6.87
81	Administrative Services	Services for Buildings	4.20	36.85	Retail	8.71	Human Resources	5.83
85	Education	Education	3.85	55.97	Retail	6.65	Associations	4.66
41	Construction	Construction	3.81	42.23	Construction Services	10.04	Infrastructure	9.14
10	Manufacturing	Food Products	3.65	27.51	Retail	15.37	Wholesale	6.62
82	Administrative Services	Services for Offices	3.59	26.67	Retail	13.30	Services for Buildings	6.80
45	Retail	Retail, Motor Vehicles	2.65	47.30	Retail	10.82	Ground Transport	4.88
42	Construction	Infrastructure	2.53	33.47	Construction	13.18	Construction Services	8.84
43	Construction	Construction Services	2.21	23.51	Construction	17.92	Infrastructure	10.36
94	Other Services	Associations	1.93	21.20	Education	12.85	Health	10.20
14	Manufacturing	Apparel	1.78	49.18	Retail	15.04	Wholesale	3.37
80	Administrative Services	Security	1.69	56.57	Services for Buildings	8.98	Retail	6.49
64	Finance	Financial Services	1.55	80.71	Services for Offices	2.25	Retail	2.12
25	Manufacturing	Metal Products, except Machinery	1.30	18.76	Retail	10.62	Machinery	4.91
84	Public Administration	Public Administration	1.20	27.39	Education	14.57	Health	14.31
23	Manufacturing	Non-Metallic Mineral Products	1.18	28.31	Retail	12.17	Construction	6.38
29	Manufacturing	Motor Vehicles	1.18	21.13	Retail	9.58	Metal Products, except Machinery	4.97
78	Administrative Services	Human Resources	1.15	13.99	Services for Buildings	18.42	Retail	9.18
22	Manufacturing	Rubber and Plastic	1.14	23.49	Retail	10.63	Wholesale	5.10
52	Transportation	Warehousing	1.11	28.41	Ground Transport	11.38	Retail	10.36
69	Professional Services	Law and Accounting	1.10	36.79	Retail	10.08	Services for Offices	9.35
1	Agriculture	Agriculture	1.09	33.79	Food Products	10.44	Retail	8.68
15	Manufacturing	Leather and Footwear	0.99	53.61	Retail	9.64	Food Products	3.97
55	Accommodation	Accommodation	0.98	30.47	Retail	13.15	Restaurant	12.60
28	Manufacturing	Machinery	0.97	21.62	Retail	8.64	Metal Products, except Machinery	7.47
62	Information and Communication	IT	0.86	41.52	Services for Offices	7.64	Retail	6.94
71	Professional Services	Architecture and Engineering	0.80	16.81	Infrastructure	12.85	Construction	10.65
13	Manufacturing	Textiles	0.74	26.94	Retail	12.36	Apparel	8.77
20	Manufacturing	Chemical Products	0.72	21.71	Retail	9.91	Wholesale	8.56
31	Manufacturing	Furniture	0.71	32.78	Retail	14.68	Food Products	4.25
96	Other Services	Personal Services	0.64	30.36	Retail	15.52	Services for Buildings	5.85
77	Administrative Services	Management	0.61	14.31	Retail	10.97	Ground Transport	9.95
24	Manufacturing	Metallurgy	0.58	21.38	Retail	7.90	Metal Products, except Machinery	6.45
93	Entertainment	Sports	0.56	28.41	Retail	13.35	Education	6.72
27	Manufacturing	Electrical Equipment	0.53	14.57	Retail	11.08	Wholesale	5.60
38	Utilities	Garbage Collection	0.52	19.86	Ground Transport	8.81	Retail	8.19
16	Manufacturing	Wood Products	0.51	31.70	Retail	11.39	Ground Transport	4.97
61	Information and Communication	Telecommunications	0.47	44.22	Retail	9.37	Services for Offices	6.58
33	Manufacturing	Maintenance	0.46	12.70	Retail	9.52	Infrastructure	8.10
17	Manufacturing	Paper Products	0.46	22.31	Retail	10.41	Wholesale	5.47
53	Transportation	Postal Services	0.46	57.28	Retail	7.68	Ground Transport	6.49
26	Manufacturing	Computer and Electronic Products	0.40	21.47	Retail	11.11	Wholesale	6.20
32	Manufacturing	Miscellaneous	0.40	16.82	Retail	15.26	Wholesale	6.58
68	Real State	Real State	0.37	16.38	Construction	12.56	Services for Buildings	11.16
65	Finance	Insurance	0.36	37.25	Health	17.27	Services for Offices	6.38
19	Manufacturing	Refined Oil Products and Biofuels	0.33	25.29	Food Products	13.74	Ground Transport	8.35
11	Manufacturing	Beverages	0.33	29.59	Wholesale	14.46	Retail	8.97
35	Utilities	Electricity and Gas	0.32	51.22	Infrastructure	6.35	Financial Services	3.55
18	Manufacturing	Media Printing	0.31	22.20	Retail	12.90	Printing	5.12
95	Other Services	Repairs and Maintenance	0.31	9.65	Retail	17.93	Telecommunications	5.74
63	Information and Communication	IT services	0.28	8.23	Services for Offices	17.72	Retail	10.40
87	Health	Social Assistance	0.27	17.04	Health	26.85	Associations	8.87
70	Professional Services	Consulting	0.27	13.63	Services for Offices	9.77	Retail	7.02
36	Utilities	Water Treatment	0.27	56.03	Infrastructure	6.16	Ground Transport	3.83
88	Health	Other in Health	0.26	19.56	Education	18.11	Associations	10.83
21	Manufacturing	Pharmaceutical Products	0.26	36.12	Wholesale	12.83	Retail	7.64
73	Professional Services	Marketing	0.26	19.60	Retail	12.95	Services for Offices	9.11
66	Finance	Other in Finance and Insurance	0.25	29.37	Financial Services	10.62	Services for Offices	10.29
30	Manufacturing	Other Transport Equipment	0.25	22.13	Retail	7.89	Infrastructure	7.71
58	Information and Communication	Printing	0.25	17.67	Retail	11.18	Media Printing	7.17
8	Extraction	Mining – non Metals	0.25	27.84	Ground Transport	9.16	Retail	7.98
2	Agriculture	Forestry	0.22	31.34	Retail	7.29	Ground Transport	6.95
74	Professional Services	Other Professional Services	0.22	8.00	Retail	12.70	Services for Offices	8.83
60	Information and Communication	Radio and TV	0.22	36.64	Retail	6.67	Associations	4.94
7	Extraction	Mining – Metals	0.21	53.49	Infrastructure	5.03	Ground Transport	4.83
79	Administrative Services	Tourism	0.20	43.65	Ground Transport	7.70	Retail	7.15
51	Transportation	Air Transport	0.17	55.02	Warehousing	6.44	Retail	4.20
50	Transportation	Water Transport	0.11	50.60	Warehousing	6.65	Ground Transport	5.00
72	Professional Services	Science	0.11	23.09	IT	9.23	Education	7.29
59	Information and Communication	Cinema, Music	0.08	15.48	Retail	19.12	Restaurant	7.96
9	Extraction	Other Mining	0.08	35.14	Infrastructure	6.08	Architecture and Engineering	4.52
6	Extraction	Oil and Gas	0.07	54.77	Refined Oil Products and Biofuels	12.47	Other Mining	8.67
90	Entertainment	Art	0.05	11.22	Retail	11.36	Services for Offices	7.19
37	Utilities	Sewage	0.04	5.98	Garbage Collection	10.64	Retail	9.20
3	Agriculture	Fishing	0.04	43.39	Food Products	9.73	Retail	7.30
12	Manufacturing	Tobacco Products	0.03	28.64	Wholesale	12.59	Retail	8.56
75	Professional Services	Veterinary	0.02	13.50	Retail	33.46	Health	5.88
5	Extraction	Coal Mining	0.01	24.44	Retail	6.73	Ground Transport	6.28
91	Entertainment	Culture and Environment	0.01	2.88	Retail	11.28	Education	7.52
99	International Services	International Services	0.01	18.58	Infrastructure	11.48	Construction	9.02
97	Domestic Services	Domestic Services	0.01	0.98	Retail	15.69	Services for Buildings	12.99
39	Utilities	Decontamination	0.00	14.04	Services for Offices	11.06	Retail	7.66
92	Entertainment	Gambling	0.00	7.69	Retail	16.78	Services for Offices	10.49

Table A3: Paired Mover Test by Reason of Separation

	Gap in Destination				
	Baseline	Without cause	Both Involuntary	Both Voluntary	Mixed
	(1)	(2)	(3)	(4)	(5)
A. Second Stage					
Gap in Origin	0.698	0.702	0.694	0.720	0.719
	(0.0009)	(0.0009)	(0.001)	(0.002)	(0.002)
Observations	1,412,592	1,359,919	995,902	163,055	200,962
B. First Stage					
Gap in Origin (lag)	0.915	0.915	0.912	0.909	0.928
	(0.0007)	(0.0007)	(0.0008)	(0.002)	(0.002)
F-statistic	1,943,164.0	1,901,164.3	1,294,847.9	306,186.5	306,509.7

This table shows different versions of the paired movers test by reason of separation. We estimate a 2SLS regression where the log earnings gap in the destination firm is regressed on the log earnings gap in the origin firm, instrumented by the lagged earnings gap in the origin firm. Panel A shows the second stage of the test, while Panel B shows the first stage. Column 1 shows replicates the baseline estimate using the full paired movers sample (Column 1, Table 4). Column 2 shows the estimate using the sample of paired movers where we exclude workers whose separation was with cause. Columns 3 to 5 show the estimate using the sample in Column 2 split by whether both workers were separated involuntarily (employer-initiated), both workers separated voluntarily (employee-initiated), or one worker separated voluntarily and the other involuntarily.

Table A4: Paired Movers Test, Residualization Approach

	Age and gender			Age, gender, and education		
	Baseline (1)	Sector Stayers (2)	Sector Switchers (3)	Baseline (4)	Sector Stayers (5)	Sector Switchers (6)
A. Second Stage						
Gap in Origin	0.690	0.724	0.655	0.719	0.741	0.698
	(0.0007)	(0.0010)	(0.0009)	(0.0007)	(0.0010)	(0.0009)
Observations	2,137,179	980,965	1,156,214	2,137,179	980,965	1,156,214
B. Brazil, 1st Stage						
Gap in Origin (lag)	0.923	0.922	0.923	0.933	0.929	0.936
	(0.0005)	(0.0007)	(0.0007)	(0.0005)	(0.0007)	(0.0007)
F-statistic	3,403,026.1	1,734,319.6	1,672,230.2	3,666,801.9	1,809,701.7	1,859,840.8

This table shows estimates based on different residualization approaches.

Table A5: The Paired Mover Test, Reweighted

	Gap in Origin								
	Baseline	Within Sector	Across Sector	Low - Low	Low - High	High - High	High - Low	Cluster Stayers	Cluster Switchers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Brazil, 2nd Stage									
Gap in Origin	0.689	0.724	0.650	0.674	0.648	0.719	0.603	0.734	0.667
	(0.0007)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Observations	1,412,592	655,097	757,495	521,582	184,756	521,473	184,781	369,052	720,723
B. Brazil, 1st Stage									
Gap in Origin (lag)	0.982	0.981	0.984	0.976	0.976	0.985	0.984	0.989	0.987
	(0.0007)	(0.0009)	(0.0010)	(0.001)	(0.002)	(0.0010)	(0.002)	(0.001)	(0.0010)
F-statistic	2,179,748.1	1,161,154.8	1,021,238.6	602,757.8	248,537.8	972,094.8	239,419.4	677,699.1	1,074,053.3
C. Veneto, 2nd Stage									
Gap in origin	0.728	0.858	0.542	0.836	0.469	0.759	0.610	0.964	0.655
	(0.029)	(0.036)	(0.047)	(0.049)	(0.067)	(0.041)	(0.110)	(0.062)	(0.032)
Observations	8,167	4,960	3,207	2,866	1,062	3,133	1,106	2,211	5,751
D. Veneto, 1st Stage									
Gap in Origin (lag)	0.405	0.420	0.385	0.369	0.435	0.463	0.379	0.408	0.409
	(0.010)	(0.012)	(0.017)	(0.015)	(0.033)	(0.016)	(0.028)	(0.018)	(0.012)
F-statistic	1,686.7	1,211.4	537.67	622.48	176.75	795.90	178.25	510.96	1,160.3

This table shows different versions of the paired movers test. All Columns and both Panels are analogous to the ones in Table 4, but we reweight the sample to match the distribution of the gap in earnings of Benchmark 2 in Table 3.

Table A6: Reverse Paired Mover Test

	LHS: Destination, RHS: Origin			LHS: Origin, RHS: Destination		
	Baseline (1)	Sector Stayers (2)	Sector Movers (3)	Baseline (4)	Sector Stayers (5)	Sector Movers (6)
A. Brazil, 2nd Stage						
Gap in Origin	0.704	0.740	0.666			
	(0.0010)	(0.001)	(0.001)			
Gap in Destination				0.819	0.839	0.797
				(0.001)	(0.002)	(0.002)
Observations	1,004,685	466,798	537,887	1,004,685	466,798	537,887
B. Brazil, 1st Stage						
Gap in Origin (lag)	0.922	0.923	0.920			
	(0.0008)	(0.001)	(0.001)			
Gap in Destination (lead)				0.590	0.625	0.555
				(0.0006)	(0.0009)	(0.0009)
F-statistic	1,390,506.7	728,122.4	664,035.2	896,561.1	483,624.4	414,839.4
C. Veneto, 2nd Stage						
Gap in origin	0.716	0.874	0.504			
	(0.041)	(0.052)	(0.063)			
Gap in destination				0.745	0.788	0.644
				(0.035)	(0.038)	(0.072)
Observations	4,698	2,719	1,979	4,698	2,719	1,979
D. Veneto, 1st Stage						
Gap in origin (lag)	0.316	0.327	0.302			
	(0.012)	(0.015)	(0.020)			
Gap in destination (lead)				0.437	0.461	0.394
				(0.013)	(0.015)	(0.023)
F-statistic	695.01	487.10	235.33	1,131.7	878.73	295.87

This table shows versions of the paired movers test in a sample in which, in addition to observing workers for at least two periods at the origin firm, we also observe them for at least two periods at the destination firm. Columns 1 to 3 show the baseline test, where the log earnings gap in the destination firm is regressed on the log earnings gap in the origin firm, instrumented by the lagged earnings gap in the origin firm. Columns 4 to 6 show the "reverse" test, where the log earnings gap in the origin firm is regressed on the log earnings gap in the destination firm, instrumented by the lead earnings gap in the destination firm. The first stage is shown in Panels B and D. Panel A shows the second stage of the test in Brazilian data, while Panel C shows the second stage in Veneto data. Columns 1 and 4 show the baseline estimate using the all paired movers. Columns 2 and 5 show movers who stay in the same sector (2-digit CNAE code for Brazil and ATECO 91 code for Veneto), while Columns 3 and 6 show movers who move to a different sector.

Table A7: Reverse Paired Mover Test II

	Baseline (1 lead/lag)		Stacked (up to 5 leads/lags)		
	LHS: Destination	LHS: Origin	LHS: Destination	LHS: Origin	Combined
	(1)	(2)	(3)	(4)	(5)
A. Brazil, 2nd Stage					
Gap in Origin/Destination	0.705	0.800	0.701	0.766	0.722
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)
Observations	1,032,346	350,031	2,083,385	1,073,898	3,157,283
B. Brazil, 1st Stage					
Gap in Origin/Destination (lag/lead)	0.924	0.775	0.924	0.789	0.875
	(0.0008)	(0.0008)	(0.001)	(0.002)	(0.001)
F-statistic	1,310,908.5	863,744.9	1,416,158.4	923,334.5	1,774,558.0

This table shows versions of the baseline and "reverse" paired movers test in different samples. Column 1 reports the baseline test, where the log earnings gap in the destination firm is regressed on the log earnings gap in the origin firm, instrumented by the lagged earnings gap in the origin firm, in the sample of paired movers who move between 2007 and 2016. Column 2 shows the reverse test, where the log earnings gap in the origin firm is regressed on the log earnings gap in the destination firm, instrumented by the lead earnings gap in the destination firm, in the sample of paired movers who move between 2012 and 2019 and are observed for at least two periods at the destination firm. With the sample in Column 1, we construct additional leads of the earnings gap for paired workers who stay at the destination firm from two to five periods after the move. Column 3 reports the baseline test where we stacked all leads at the destination firm and run a single regression, where we cluster standard errors at the pair level. Column 4 reports the analogous exercise for the reverse test, where we stack lags at the origin firm using the sample in Column 2. Finally, Column 5 reports the combined test, where we stack all leads and lags at the destination and origin firms, respectively, and run a single regression. The first stage is shown in Panel B.

Table A8: Event Study Test, by Reason of Separation

	Change in $\hat{\psi}_1$ / Change in Log Earnings							
	Involuntary				Voluntary			
	Full sample	Within 1 month	Btw. 2–6 months	After 6 months	Full sample	Within 1 month	Btw. 2–6 months	After 6 months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Brazil, 2nd Stage								
Change in $\hat{\psi}_1$	1.09	1.13	1.04	1.02	1.00	1.09	1.04	0.921
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.006)	(0.006)	(0.004)
Observations	1,547,963	805,158	449,962	292,843	533,548	175,175	120,599	237,774
B. Brazil, 1st Stage								
Change in $\hat{\psi}_2$	0.711	0.708	0.706	0.714	0.691	0.677	0.666	0.707
	(0.0006)	(0.0008)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
F-statistic	1,579,098.3	808,215.0	446,606.1	312,171.9	490,430.3	150,112.5	95,430.0	240,959.7

This table shows the results of versions of the event study test using AKM estimates by reason of separation and time to re-employment. We first estimate two sets of firm fixed effects by randomly splitting the full data (Column 1, Table 1) in two samples. We then 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 the first sample to instrument the firm effects in the second sample. Columns 1 and 5 show the results using the sample of movers who did not separate with cause and both workers take the same (discrete) amount of time to re-employment (as in Table A3). Time to reemployment is defined as the difference between the hiring date at the destination firm and the separation date at the origin firm. Columns 1 and 5 show the results for involuntary and voluntary separations, respectively. Columns 2–4 and 6–8 show the results for involuntary and voluntary separations, respectively, split by whether both workers found a new job within one month of the month of separation between 2 and 6 months after separation, or after 6 months of separation.

Table A9: AKM Estimation (KSS-corrected)

	All Workers (1)	All workers, residualized (2)	Paired movers, residualized (3)	Model Simulated Data (4)
A. Brazil, 2012–2017				
Var(Y)	0.506	0.486	0.473	0.179
Var(α)	0.283	0.271	0.295	0.074
Var(ψ)	0.072	0.071	0.048	0.073
Cov(ψ, α)	0.050	0.047	0.035	0.001
Var(ε)	0.051	0.050	0.059	0.030
R^2	0.929	0.927	0.907	0.851
Adj. R^2	0.907	0.904	0.877	0.848
Adj. R^2, interactions	0.934	0.933	0.924	1.000
Observations	40,950,644	40,950,644	1,061,171	20,000,000
# Firms	911,755	911,755	49,927	200
# Workers	8,764,893	8,764,893	211,831	200,000
# Movers	4,598,015	4,598,015	211,831	199,045
# of Moves	14,585,164	14,585,164	644,314	1,783,852
B. Veneto, 1996–2001				
Var(Y)	0.7655	0.7182	0.8370	0.5762
Var(α)	0.3174	0.2308	0.3907	0.2381
Var(ψ)	0.1259	0.1118	0.1964	0.0953
Cov(ψ, α)	0.0438	0.0346	0.0228	-0.0066
Var(ε)	0.2347	0.3065	0.2044	0.2441
R^2	0.8200	0.7073	0.8566	0.7830
Adj. R^2	0.7525	0.6097	0.8074	0.7807
Adj. R^2, interactions	0.8037	0.7576	0.8282	0.9929
Observations	5,111,991	1,702,650	4,346	486,159
# Firms	61,638	61,638	179	100
# Workers	1,332,729	369,956	931	5,000
# Movers	363,959	369,956	931	4,391
# of Moves	988,088	988,088	2,362	35,408

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). Column 1 shows the reults 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 2 restricts the samples only to workers that move at least once in the period, and Column 3 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). Column 4 shows the same estimation in data simulated from the estimated model. 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 three rows show the number of observations, firms, workers, movers, and moves in the connected set of data used in the estimation. Following Kline, Saggio, and Sølvsten (2020), we construct the "leave-one-out-match" connected set of firms and workers.

Table A10: BLM Estimation

	All Workers (1)	Paired Movers (2)
A. Brazil, 2012–2017		
Cor(k, l)	0.260	0.616
Cov(k, l)	0.141	0.195
Var(k)	0.761	0.772
Var(l)	0.097	0.033
R^2	0.834	0.927
# of Workers	8,992,603	206,635
B. Veneto, 1996–2001		
Cor(k, l)	0.164	0.033
Cov(k, l)	0.042	0.008
Var(k)	0.940	0.976
Var(l)	0.018	0.016
R^2	0.901	0.962
# of Workers	947,928	4,351

This table shows moments of the BLM estimation relating to worker types (k) and firm classes (l) in Brazilian (Panel A) and Veneto (Panel B) data. The first row shows the correlation between worker types and firm classes, the second to fifth rows show the covariance and variance of worker types and firm classes as a share of the total explained variance. Finally, the sixth row shows the R^2 of a regression of the log earnings on worker types and firm classes (share of total variance explained). The last row reports the number of workers in the estimation sample. Column 1 shows the results for the estimation using all workers, while Column 2 shows the results for the estimation using only paired movers.

Table A11: β_1 Decomposition: Equation 17

Term	Value
$Var(\alpha_i - \alpha_{i'})$	0.152
$Cov(m_{iA} - m_{i'A}, m_{iB} - m_{i'B})$	0.008
$Var(m_{iA} - m_{i'A})$	0.076
$Cov(\alpha_i - \alpha_{i'}, m_{iA} - m_{i'A})$	0.000
$Cov(\alpha_i - \alpha_{i'}, m_{iB} - m_{i'B})$	0.000

This Table uses the model to decompose how the model arrives at its paired mover coefficient.

Table A12: Multi-sector Model results

	(1)	(2)	(3)	(4)	(5)
A. Brazil					
	EE rate	β_{Stayer}	β_{Switcher}	$Var(\alpha_i)$	$Var(\psi_j)$
Empirical Moments (Data)	0.0384	0.737	0.656	0.3100	0.0669
Simulated Moments (Model)	0.0354	0.776	0.675	0.3868	0.0706
	λ_0 (UE rate)	δ (EU rate)			
Calibrated Parameters	0.1074	0.0232			
	λ_1	$Var(m_{ij})$	$Var(m_{is})$	$Var(\alpha_i)$	$Var(\psi_j)$
Estimated Parameters	0.2869	1.3824	0.2094	0.5915	0.1090
Steady-state (employed workers)		m_{ij}	m_{is}	α_i	ψ_j
Mean		1.6095	0.2054	-0.0066	0.1691
Variance		1.0339	0.2073	0.5977	0.1220
Covariance (m_{ij},ψ_j)	-0.0280				
Covariance (m_{is},ψ_j)	-0.0050				
Mobility (EE)	% Offers Accepted	% Offers Cross-Sector	% Within-Sector Offers Accepted	% Cross-Sector Offers Accepted	% Moves Cross-Sector
	0.1179	0.478	0.2142	0.1126	0.9052

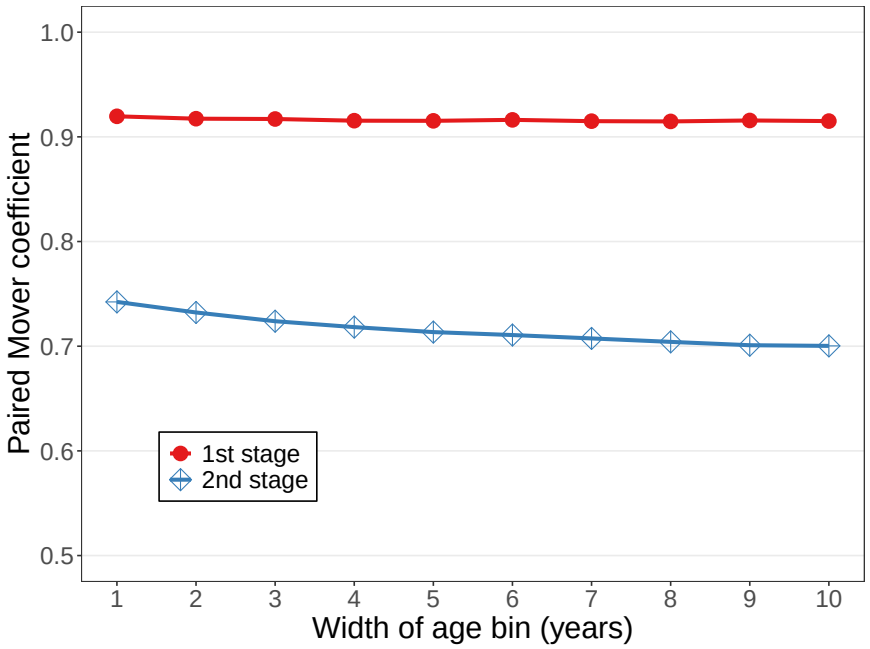
This table shows the results of the model estimation. We estimate the model using the simulated method of moments (SMM) on the empirical moments of the data. We fix the λ_0 and δ parameters to quarterly UE and EU rates (we use PNAD for Brazil and folow XXX for Veneto). We estimate the five remaining parameters of the model, λ_1, $Var(m_{ij})$, $Var(m_{is})$, $Var(\alpha_i)$, and $Var(\psi_j)$, to match empirical moments of the data: the EE rate, the β coefficient in the paired movers test split by sector stayers and switchers, and the variances of the firm and worker effects. We search the parameter space using a Metropolis-Hastings algorithm with 10 chains of 250 iterations each. At each iteration, we simulate 5000 workers, 200 firms in 20 sectors, over 200 periods. The first two rows in each panel show the empirical moments in the data and the simulated moments in the estimated model. The next two rows show the estimated and calibrated parameters of the model. The following four rows show the steady-state moments of employed workers in the model. The last row shows mobility statistics in the model. % Offers Accepted is the fraction of outside offers that are accepted by workers, % Cross-Sector Offers is the fraction of offers that are made across sectors, % Within-Sector Offers Accepted is the fraction of within-sector offers that are accepted, % Cross-Sector Offers Accepted is the fraction of cross-sector offers that are accepted, and % Moves Cross-Sector is the fraction of actual moves that occur across sectors.

Table A13: Paired Movers Test, Robustness

	Positive gap at $t = 0$				Gap in Destination > 0			
	Baseline (1)	Stayers (2)	Sector Stayers (3)	Sector Switchers (4)	Baseline (5)	Stayers (6)	Sector Stayers (7)	Sector Switchers (8)
A. Second Stage								
Abs. Gap in Origin	0.871	0.974	0.888	0.850				
	(0.001)	(0.0005)	(0.002)	(0.002)				
Gap in Origin > 0					0.632	0.925	0.686	0.586
					(0.001)	(0.0006)	(0.002)	(0.002)
Observations	1,981,191	1,302,157	900,977	1,080,214	1,981,191	1,302,157	900,977	1,080,214
B. Brazil, 1st Stage								
Abs. Gap in Origin (lag)	0.706	0.852	0.719	0.693				
	(0.0006)	(0.0004)	(0.0008)	(0.0008)				
Gap in Origin0$>$ 0 (lag)					0.605	0.810	0.619	0.593
					(0.0006)	(0.0005)	(0.0008)	(0.0008)
F-statistic	1,587,311.8	5,394,925.8	818,909.5	766,797.0	1,141,712.6	2,481,294.3	560,091.0	584,493.1

This table shows alternative ways of defining the gap in earnings at the origin firm.

Figure A1: Paired Movers Test, sensitivity to age bin width



This figure shows the paired mover coefficient using an alternative estimation approach compared to the baseline. Rather than residualizing earnings for covariates, we match workers on gender and age bins. The figure shows how our estimates vary with the size of the age bins.