Match Effects[1]

Simon D. Woodcock[2]

Simon Fraser University

simon_woodcock@sfu.ca

February 2011

[1]This document reports the results of research and analysis undertaken by the U.S. Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This document is released to inform interested parties of ongoing research and to encourage discussion of work in progress. This research is a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grants SES-9978093 and SES-0427889 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant R01~AG018854, and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author(s) and do not represent the views of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau supports external researchers' use of these data through the Research Data Centers (see www.ces.census.gov). For other questions regarding the data, please contact Jeremy S. Wu, Manager, U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233, USA. (Jeremy.S.Wu@census.gov http://lehd.dsd.census.gov).

[2]Correspondence to: Department of Economics, Simon Fraser University, 8888 University Dr., Burnaby, BC V5A 1S6, Canada. I thank John Abowd, Krishna Pendakur, Julia Lane, and participants at CAFE 2006 for helpful comments and suggestions. I also thank the US Census Bureau and LEHD Program for facilitating data access. This research was partially supported by the SSHRC Institutional Grants program and NSF Grant SES-0339191 to Cornell University.

Abstract

We present direct evidence of the importance of matching in wage determination. It is based on an empirical specification that estimates the returns to person-, firm-, and match-specific determinants of match productivity. We call these person, firm, and match effects. The distinction between these components is important, because they have different implications for the persistence of individual earnings and the returns to employment mobility. We find that match effects, which have been ignored in previous work, are an important determinant of earnings dispersion. They explain 16 percent of variation in earnings, and much of the change in earnings when workers change employer. Specifications that omit match effects substantially over-estimate the returns to experience, attribute too much variation to personal heterogeneity, and underestimate the extent to which good workers sort into employment at good firms.

JEL Classification: J20, C23

Keywords: matching, sorting, linked employer-employee data, wage dispersion, person and firm effects, fixed effects, random effects, human capital

1 Introduction

A primary function of the labor market is to allocate workers to jobs. However, which workers match with which firms, and the consequences of matching for wage determination, remains poorly understood. Intuition suggests that "good" workers will match with "good" firms. Theory supports this intuition (in the presence of complementarity), but recent evidence based on wage data does not.[1] The idea that there are "good" matches and "bad" matches is well-established, but quantifying this in wages is hampered by a lack of direct measures of match quality, and the potentially confounding effects of unobserved worker and firm heterogeneity.

We present direct evidence of the importance of matching in wage determination. It is based on an empirical model that controls for observable and unobservable characteristics of workers and firms (person and firm effects), and an interaction effect between the worker and firm. We call this the match effect. In a simple model of wage determination, person effects measure the value of worker-specific determinants of match productivity; firm effects reflect firm-specific determinants of productivity, product market conditions, and the firm's compensation policy; and match effects measure the value of match quality.

The primary contribution of the match effects model is to measure the relative importance of worker-, firm-, and match-specific heterogeneity in labor earnings. The relative magnitude of these components is of substantive economic interest. If wage variation primarily reflects workers' measured and unmeasured productive characteristics, then individual wages will be highly persistent, largely invariant to where individuals work, and the potential returns to employment mobility will be small. On the other hand, if firm- and match-specific heterogeneity are important, then the cost of involuntary displacement from high-paying firms and good matches will be large, but so will the potential returns to search.

We estimate the match effects model on the US Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) database. Match effects explain 16 percent of observed varia-

[1] See Abowd et al. (2004) in particular.

tion in log earnings. Personal heterogeneity accounts for more than half of observed variation, and firm-level heterogeneity in compensation explains another 22 percent. Our results imply considerable persistence in individual earnings, coupled with substantial potential returns to search.

We use the match effects model to decompose earnings growth when individuals change employer. It is well known that a large portion of lifetime earnings growth occurs when individuals change job (e.g., Bartel and Borjas (1981), Altonji and Shakotko (1987), Topel and Ward (1992), and others). This could reflect moving from lower-paying firms to higher-paying firms or sorting into better matches. We find that the relative importance of these two factors depends on whether there is an intervening period of non-employment between jobs. Workers who transit directly from one employer to another experience year-over-year earnings growth nearly 3 times larger than job stayers. About 60 percent of the excess growth is due to sorting into higher-paying firms, and nearly 30 percent to sorting into better matches. In contrast, individuals who experience an intervening period of non-employment have much lower wage growth than individuals who do not change jobs, and the difference is almost entirely due to sorting into worse matches.

We find direct evidence that matching is positively assortative. That is, we find a positive correlation between person- and firm-specific components of log earnings, which indicates that "good" workers match with "good" firms on average. This finding is in contrast to previous work that ignores match effects. Abowd et al. (2004), for example, find a near-zero correlation between person and firm effects in the US, and a negative correlation in France. The difference between our finding and previous work is attributable to bias from omitted match effects. In fact, estimated person and firm effects are unbiased only if all excluded match effects are zero. We easily reject this hypothesis.

Omitted match effects also bias the estimated returns to observable characteristics that are correlated with match quality. This matters if workers with certain characteristics are more successful at generating good matches than others. We find evidence of this bias in the estimated returns to experience. A specification that omits match effects over-estimates the returns to 25 years of experience by 26 percent for men and 23 percent for women. This is evidence that some of the returns traditionally attributed to the accumulation of general human capital are actually attributable matching, and that workers sort into better matches over the course of a career.

The remainder of the paper is organized as follows. In Section 2, we present the match effects model, develop our estimators, and derive the bias due to omitted match effects. Section 3 describes the data used in the empirical application, and Section 4 presents the estimation results. We conclude in Section 5.

2 The Match Effects Model

A simple model of match productivity and wage determination helps fix ideas. Suppose that worker i has productive characteristics (e.g., ability, human capital, and other "portable" determinants of productivity) indexed by $L_i > 0$. Firm j has productive characteristics (e.g., technology, capital intensity, and organizational capital) represented by an index $K_j > 0$. When worker i is employed at firm j in period t, match productivity Q_{ijt} is given by the Cobb-Douglas function:

$$Q_{ijt} = \mu L_i^{\theta} K_j^{\psi} M_{ij}^{\phi} e_{ijt} \tag{1}$$

where μ is a scale factor; θ, ψ, and ϕ are parameters; $M_{ij} > 0$ is match-specific productivity shifter; and e_{ijt} is an idiosyncratic productivity shock with geometric mean one. We call M_{ij} match quality; it can be interpreted as an index of complementarity between the worker's and firm's productive attributes. Good matches are those that are more productive (i.e., M_{ij} is larger) for given values of L_i and K_j.

In the production function (1), an individual who consistently generates good matches (i.e., for whom expected match quality is above average) is indistinguishable from an individual whose productivity index L_i is above average. In both cases, the worker has above-average expected productivity at any firm. The same is true of firms that consistently generate good matches. Hence we assume that all workers and firms face the same distribution of match quality, and we normalize its geometric mean to one.[2] This is intuitive: consistently generating good matches is a skill that increases an agent's expected productivity in any match, and is consequently no different than other productive attributes embodied in L_i and K_j. We similarly normalize the geometric means of L_i and K_j to one, because we cannot distinguish an economy with high average worker productivity from one with high average firm productivity.

Suppose that firms face price p_j for their output, normalized to have geometric mean one, and that wages w_{ijt} are determined by a Nash bargain. When employees of firm j have bargaining strength γ_j, the outside option is zero for both parties, the worker maximizes w_{ijt}, and the firm maximizes $p_j Q_{ijt} - w_{ijt}$, the bargaining solution is $w_{ijt} = \gamma_j p_j Q_{ijt}$.[3] Taking logarithms, we have:

$$\ln w_{ijt} = \ln \gamma\mu + \theta \ln L_i + \left(\psi \ln K_j + \ln \gamma_j/\gamma + \ln p_j\right) + \phi \ln M_{ij} + \ln e_{ijt} \qquad (2)$$

where γ is the geometric mean of γ_j. The log wage is additively-separable in worker-, firm-, and match-specific components. Our normalizations imply that all three components have zero mean; they are identified up to location. They measure relative wage differences due to productivity differences between workers, firms, and matches, due to product market conditions as reflected in p_j, and due to the firm's compensation policy as reflected in γ_j.

Our empirical model is based on (2). Parameters of the production function (θ, ψ, ϕ) are not separately identified from the productivity indices, so we define a pure match effect, $\phi_{ij} = \phi \ln M_{ij}$, that measures the component of wages due to match quality. Similarly, and because we do not observe output prices or firms' compensation policies, we define a pure

[2] That is, without loss of generality, we normalize $E[\ln M_{ij}] = 0$.

[3] Other wage-setting institutions deliver the same result, e.g., if the worker and firm split the output from a match, with share γ_j going to the worker, as in Pissarides (1994). An observationally equivalent result arises if workers are paid the value of their marginal product, $w_{ijt} = p_j Q_{ijt}$.

firm effect $\psi_j = \psi \ln K_j + \ln \gamma_j/\gamma + \ln p_j$. Because workers' productivity may vary over time as they accumulate human capital, the worker-specific component of log wages is defined as $x_{it}'\beta + \theta_i = \theta \ln L_i$, where x_{it} is a vector of observable time-varying personal characteristics that determine productivity (e.g., experience), β is a parameter vector, and θ_i is a pure person effect that measures the returns to time-invariant characteristics. Hence $x_{it}'\beta + \theta_i$ is the portable component of an individual's wage, reflecting the market value of her productive attributes. Given these definitions, and defining $\varepsilon_{ijt} = \ln e_{ijt}$, our empirical specification is:

$$y_{ijt} = \mu + x_{it}'\beta + \theta_i + \psi_j + \phi_{ij} + \varepsilon_{ijt} \tag{3}$$

where y_{ijt} is log compensation of worker i at firm j in period t, and μ is the grand mean. Equation (3) is structurally identical to (2), and hence yields structural estimates of the worker-, firm-, and match-specific determinants of log wages under our simple model of wage determination.

Person, firm, and match effects may comprise observed and unobserved components. In our data, we observe time-invariant personal characteristics, u_i, that may be related to productivity and wages (e.g., sex, race, education). We therefore decompose the pure person effect into its observed and unobserved components via $\theta_i = u_i'\eta + \alpha_i$, where η measures the returns to time-invariant personal characteristics and α_i is the unobserved component. With appropriate data, we could decompose the firm and match effects similarly, but we do not do so here.[4]

In the absence of match effects, equation (3) reduces to the person and firm effects model considered by Abowd et al. (1999) and others. Given its growing use, we consider the person and firm effects model as a point of departure in our empirical application, and we test the restrictions it imposes on (3), namely $\phi_{ij} = 0$ for all i, j.

[4] We do not observe any match-level characteristics except job duration (tenure) in our data. We estimate specifications with tenure controls in Section 4.3. Observable firm characteristics in our data include industry, county, and size. Abowd et al. (1999) show that industry and size effects are aggregates of ψ_j; the same is true of county effects. Omitting these characteristics therefore has no consequences for the estimation of pure person, firm, and match effects.

Most previous work based on the person and firm effects model has ignored the role of matching. An exception is Woodcock (2010), who develops a dynamic model with a production technology similar to (1), but where match quality is not directly observed by workers and firms. Instead, they learn its value slowly. The equilibrium wage function is similar to (2), but does not include a time-invariant match effect. Rather, wages depend on agents' current beliefs about match quality. Empirically, this implies a model with person and firm effects where errors are correlated over time within a match. Correlation in the errors reflects intertemporal correlation in beliefs about match quality. The match effects model (3) is a limiting case of such a learning model where signals about match quality are perfectly informative. This implies restrictions on the within-match covariance of wages; we test these in our empirical application.

Let N^* denote the total number of observations; N is the number of individuals; J is the number of firms; $M \leq NJ$ is the number of worker-firm matches; k and q are the number of time-varying and time-invariant covariates, respectively; T_i is the number of observations on worker i; N_j is the number of observations on firm j; and T_{ij} is the number of observations on the match between worker i and firm j. We rewrite the model in matrix notation:

$$y = \mu + X\beta + D\theta + F\psi + G\phi + \varepsilon \tag{4}$$

$$\theta = \alpha + U\eta \tag{5}$$

where y is the $N^* \times 1$ vector of log compensation; μ is the $N^* \times 1$ mean vector; X is the $N^* \times k$ matrix of time-varying covariates; β is a $k \times 1$ parameter vector; D is the $N^* \times N$ design matrix of indicator variables for each worker; θ is the $N \times 1$ vector of person effects; F is the $N^* \times J$ design matrix of indicators for each firm; ψ is the $J \times 1$ vector of firm effects; G is the $N^* \times M$ design matrix of indicators for each match; ϕ is the $M \times 1$ vector of match effects; α is the $N \times 1$ vector of unobserved components of the person effect; U is the $N \times q$ matrix of time-invariant personal characteristics; η is a $q \times 1$ parameter vector; and ε is the

$N^* \times 1$ error vector.

2.1 Identification and Estimation

The intuition governing identification is straightforward. Person effects are portable, so identification is based on conditional covariation between an individual's wage at different employers. Likewise, identification of the firm effect is based on the common component of wages amongst the firm's employees. Identification of the match effect is based on conditional covariation in wages within the match that is not explained by person and firm effects. Separately identifying the three components requires mobility of workers between firms. As we shall see, some kinds of mobility are admissible, and some are not.

Even without specifying a complete model of employment mobility, our model of match productivity makes clear that mobility may depend on person, firm, and match effects. For example, suppose that employed workers quit when they receive an outside offer that exceeds their current wage. Workers will be less likely to quit when firm and match effects in the current job are large, since it is less likely that outside offers will exceed the current wage.

Our primary identifying assumption is that errors have zero conditional mean,

$$E\left[\varepsilon|D, F, G, X\right] = 0 \tag{6}$$

(a slightly weaker assumption based on orthogonality would also suffice). This assumption implies that employment mobility (or equivalently, selection into matches) is conditionally exogenous. Mobility (selection) may depend on observable characteristics, person effects, firm effects, and match effects, but not ε_{ijt}.[5] Note this is considerably weaker than the assumption required to identify the person and firm effects model: $E\left[\varepsilon|D, F, X\right] = 0$, see Abowd et al. (1999). That is, the match effects model is identified when mobility (selection into matches) depends on unobserved match-specific heterogeneity because (6) conditions

[5] A caveat applies for our fixed effect estimator, see below.

on G, but the person and firm effects model is not.

The match effects model is identified under fixed or random effects assumptions about the unobserved heterogeneity components. We consider both below. We discuss the fixed effect case first, because most prior work based on the person and firm effects model has taken this approach, and because economists often prefer these estimators over random effect alternatives. However, limitations of the fixed effect estimator lead us to prefer a random effect alternative.

We make minimal assumptions about second moments of the errors. Our fixed effect estimator is based on ordinary least squares (OLS). It is well known that non-spherical errors only affect the efficiency properties of OLS. When efficiency is a concern, straightforward GLS extensions of our OLS estimator are available. In the interest of brevity, we do not develop those here. In the random effects case, however, our primary interest is the variance and covariance of random person, firm, and match effects. Estimates of the variance parameters may be sensitive to mis-specification of the error distribution. Consequently, we develop our random effect estimator in the case of general non-spherical errors, $E\left[\varepsilon\varepsilon'|X,D,F,G\right]=R$, where R is an $N^*\times N^*$ positive definite symmetric matrix. Our main empirical results are based on $R=\sigma_\varepsilon^2 I_{N^*}$, where I_A is the identity matrix of order A, but we subsequently relax this assumption.[6]

2.1.1 A Fixed Effect Estimator

Estimating β in the presence of fixed person, firm, and match effects is straightforward. Applying standard results for partitioned regression, the OLS estimator of β in (4) is the within-match estimator:

$$\hat{\beta}=\left(X'M_{[D\ F\ G]}X\right)^{-1}X'M_{[D\ F\ G]}y \tag{7}$$

[6] In general, we must impose some structure on R. Aside from the incidental parameters problem, some specifications of R will render some random effects unidentified. For example, if we allow arbitrary correlation between errors corresponding to the same individual, we cannot separately identify a random person effect without a scale restriction.

where $M_A \equiv I - A(A'A)^{-} A'$ projects onto the column null space of A, and A^{-} is a generalized inverse of A.[7] Some algebra verifies that $M_{[D\ F\ G]}$ takes deviations from means within each match.[8] So we can recover $\hat{\beta}$ from the regression:

$$y_{ijt} - \bar{y}_{ij\cdot} = (x_{it} - \bar{x}_{ij\cdot})' \beta + \nu_{ijt} \tag{8}$$

where $\bar{y}_{ij\cdot}$ and $\bar{x}_{ij\cdot}$ are sample means of y_{ijt} and x_{it}, respectively, in the match between worker i and firm j; and $\nu_{ijt} = \varepsilon_{ijt} - \bar{\varepsilon}_{ij\cdot}$ is the within-match error.

The fundamental identification problem is to distinguish good matches from workers and firms with large person/firm effects. Complicating this, the fixed effect specification is over-parameterized. There are $N + J + M + 1$ person effects, firm effects, match effects, and a constant term to estimate, but only M matches from which to estimate them.[9]

In light of this, our fixed effect estimator assumes that match effects are orthogonal to person and firm effects. This is restrictive, and is one of the reasons we prefer a random effect alternative.[10] In particular, orthogonality requires $E\left[\phi_{ij} | D_i, F_j\right] = 0$, where D_i is the indicator for worker i and F_j is the indicator for firm j. This condition will be violated if employment mobility (or selection into matches) depends on ϕ_{ij}.[11] However, an orthogonal match effect is identified whenever the corresponding person and firm effects are identified

[7] We assume X has full column rank. D and F do not generally have full column rank without further restrictions, e.g., excluding one column per connected group of workers and firms. See Searle (1987, Ch. 5) for a general discussion of connected data, or Abowd et al. (2002) for a discussion in the context of linked employer-employee data. Briefly, workers are connected by a common employer, and firms are connected by common employees.

[8] $M_{[D\ F\ G]}$ projects onto the column null space of $[D\ F\ G]$. It is a block diagonal matrix with N^* rows and columns, where the M diagonal blocks correspond to each of the M matches. The diagonal block corresponding to the match between worker i and firm j is a $T_{ij} \times T_{ij}$ submatrix $M^{ij}_{[D\ F\ G]} = I_{T_{ij}} - T_{ij}^{-1} \iota_{T_{ij}} \iota'_{T_{ij}}$, where ι_A is an $A \times 1$ vector of ones.

[9] There are M estimable functions of the person, firm, and match effects, the overall constant, and a mean in each connected group. For clarity of exposition, discussion in the main text presumes the sample comprises a single connected group.

[10] This assumption is innocuous in the "all cells filled" case where each worker is empoyed at each firm, because variance decompositions are invariant to the identifying normalization. This invariance does not hold when some cells are empty; see Searle (1987, Ch. 9).

[11] That is, $E\left[\phi_{ij} | D_i, F_j\right] = 0$ rules out any systematic relationship between ϕ_{ij} and employer identity. Note that violations of this condition do not affect identification or consistency of $\hat{\beta}$.

in a model without match effects. See Abowd et al. (2002) for discussion of identification in that model. The orthogonality assumption also normalizes the (duration-weighted) sample mean of estimated match effects to zero for each person and firm. Consequently, if a worker (or firm) enters only one match, the associated match effect is normalized to zero. This is another reason to prefer the random effect alternative, since it imposes no such normalization.

The orthogonal match effects estimator is implemented as follows. We estimate β using the within-match estimator (7), and calculate the sample cell means, $\bar{\mu}_{ij\cdot} = T_{ij}^{-1} \sum_t \left(y_{ijt} - x_{it}'\hat{\beta} \right)$. Let $\bar{\mu}$ denote the $N^* \times 1$ vector of sample cell means. The OLS estimator of the intercept is $\hat{\mu} = \frac{1}{N^*} \sum \bar{\mu}_{ij\cdot}$, and the estimated person and firm effects solve:

$$\begin{bmatrix} D'D & D'F \\ F'D & F'F \end{bmatrix} \begin{bmatrix} \hat{\theta} \\ \hat{\psi} \end{bmatrix} = \begin{bmatrix} D' \\ F' \end{bmatrix} (\bar{\mu} - \hat{\mu}) \tag{9}$$

subject to the Abowd et al. (2002) grouping conditions.[12] The OLS estimator of the orthogonal match effect is $\hat{\phi} = M_{[D\ F]} (\bar{\mu} - \hat{\mu}) = \bar{\mu} - \hat{\mu} - D\hat{\theta} - F\hat{\psi}$, which is the vector of residuals in the OLS regression of $\bar{\mu}$ on D, F, and an intercept. We subsequently decompose the person effect into its observable and unobservable components via least squares regression of $\hat{\theta}_i$ on u_i. Residuals in this regression define an estimator of α_i that is orthogonal to u_i.

The orthogonal match effect estimator has the following properties. As usual, $\hat{\beta}$ is unbiased and consistent as $N^* \to \infty$. It is BLUE and asymptotically efficient under spherical errors. The estimated person, firm, and match effects are unbiased (subject to the orthogonality restriction), and they are BLUE under spherical errors. However, there is an incidental parameters problem: $\hat{\theta}_i, \hat{\psi}_j$, and $\hat{\phi}_{ij}$ are only consistent as T_i, N_j, and T_{ij}, respectively, tend to infinity. Furthermore, Andrews et al. (2008) show, in the context of the person and firm effects model, that the sample variances and covariance of $\hat{\theta}_i$ and $\hat{\psi}_j$ are biased. The same

[12] Abowd et al. (2002) derive necessary and sufficient conditions to identify $\hat{\theta}$ and $\hat{\psi}$ in the person and firm effects model. They are identified up to location in each connected group of workers and firms. A sufficient condition is $\sum_{i \in g} \hat{\theta}_i = 0$ and $\sum_{j \in g} \hat{\psi}_j = 0$ in each group g.

bias is present in the orthogonal match effects estimator.[13] The bias is due to least squares estimation error in the presence of limited sample mobility. Bias in the estimated variance of person and firm effects vanishes asymptotically as $N^*/N \to \infty$ and $N^*/J \to \infty$, respectively, and bias in the estimated covariance vanishes as $(N^*/N)^{1/2} (N^*/J)^{1/2} \to \infty$. There is a similar bias in the estimated variance of $\hat{\phi}_{ij}$ that vanishes as $N^*/M \to \infty$. It is worth noting that Abowd et al. (2004) find that limited mobility bias is small in the context of the person and firm effects model, and simulations in Appendix A find similarly small bias in the orthogonal match effect estimator.

2.1.2 A Mixed Effect Estimator

Limitations of the orthogonal match effects estimator lead us to prefer a random effect specification of the unobserved heterogeneity components. However, a traditional random effect estimator imposes restrictions on the relationship between observables and unobservables. This is problematic, because we are interested in the relationship between personal heterogeneity in wages, $x_{it}'\beta + \theta_i$, and firm heterogeneity, ψ_j, to assess whether matching is assortative. Consequently, we develop a novel "hybrid" mixed effect estimator that combines features of traditional fixed and mixed (random) effect estimators. Similar to the Hausman and Taylor (1981) correlated random effects estimator, our estimator allows arbitrary correlation between the random effects and time-varying observable characteristics, x_{it}. Estimation proceeds in three stages, as follows.

In the first stage, we estimate β using the within-match estimator (7). This imposes no restrictions on the relationship between x_{it} and the unobserved heterogeneity components. Second, we estimate the variance of the random effects, $\sigma_\alpha^2, \sigma_\psi^2, \sigma_\phi^2$, and the error covariance R by Restricted Maximum Likelihood (REML) on $y_{ijt} - x_{it}'\hat{\beta}$.[14] REML estimates $\tilde{\sigma}_\alpha^2, \tilde{\sigma}_\psi^2, \tilde{\sigma}_\phi^2$,

[13] The proof parallels Andrews et al. (2008) closely and so is omitted but available on request. The intuition is straightforward: the orthogonal match effect estimator uses the same least squares estimator to decompose $\bar{\mu}_{ij\cdot}$ and consequently suffers a similar bias.

[14] REML's description as a maximum likelihood estimator is historical. Just as OLS is the MLE under normality but normality is not required to establish its main properties, REML does not require normality either; see Jiang (1996). It is akin to partitioned regression, and is equivalent to maximum likelihood

and $\tilde{R}$ are identified under the conditional moment assumptions:

$$E\left[\alpha|U,\hat{\beta}\right] = E\left[\psi|U,\hat{\beta}\right] = E\left[\phi|U,\hat{\beta}\right] = 0 \quad (10)$$

$$Cov\left[\begin{array}{c|c} \alpha & \\ \psi & U,\hat{\beta} \\ \phi & \end{array}\right] = \left[\begin{array}{ccc} \sigma_\alpha^2 I_N & 0 & 0 \\ 0 & \sigma_\psi^2 I_J & 0 \\ 0 & 0 & \sigma_\phi^2 I_M \end{array}\right] \equiv W \quad (11)$$

and assuming the random effects are uncorrelated with ε_{ijt}. Unlike a traditional random effect estimator, (10) and (11) condition on $\hat{\beta}$ but not X. This is weaker than the identifying assumptions of a traditional random effect estimator. Specifically, a traditional random effect specification requires that unobserved heterogeneity components have zero mean at each sample value of x_{it}. In contrast, our estimator only requires the random effects to have zero mean in the entire sample, given the sample realization of $\hat{\beta}$.[15] In large samples such as ours, $\hat{\beta}$ will carry almost the same information as the population parameter, so this is a very weak assumption.

In the final stage, we solve for the Best Linear Unbiased Estimator (BLUE) of η and Best Linear Unbiased Predictor (BLUP) of the random effects.[16] These solve an equation system

on linear combinations of the data under normality, where the linear combinations are invariant to η. In our application, the linear combinations $K'\left(y - X\hat{\beta}\right)$ satisfy $K'U\eta = 0$ for all values of η, which implies $K'U = 0$. Thus K' projects onto the column null space of U and is of the form $K' = C'M_U$ for arbitrary C'.

[15] In some cases, we may be willing to accept the more restrictive moment assumptions of a traditional random effect estimator to increase efficiency. We consider this alternative in Section 4.3, but reject the hypothesis that random effects are uncorrelated with x_{it}.

[16] BLUPs are *best* in the sense of minimizing the mean square error of prediction among linear unbiased estimators, and *unbiased* in the sense $E\left[\tilde{\alpha}\right] = E\left[\alpha\right]$, $E\left[\tilde{\psi}\right] = E\left[\psi\right]$, and $E\left[\tilde{\phi}\right] = E\left[\phi\right]$. See Goldberger (1962) and Robinson (1991) for details.

based on the Henderson et al. (1959) mixed model equations:

$$\begin{bmatrix} U'\tilde{R}^{-1}U & U'\tilde{R}^{-1}D & U'\tilde{R}^{-1}F & U'\tilde{R}^{-1}G \\ D'\tilde{R}^{-1}U & D'\tilde{R}^{-1}D+\tilde{\sigma}_{\alpha}^{-2}I_N & D'\tilde{R}^{-1}F & D'\tilde{R}^{-1}G \\ F'\tilde{R}^{-1}U & F'\tilde{R}^{-1}D & F'\tilde{R}^{-1}F+\tilde{\sigma}_{\psi}^{-2}I_J & F'\tilde{R}^{-1}G \\ G'\tilde{R}^{-1}U & G'\tilde{R}^{-1}D & G'\tilde{R}^{-1}F & G'\tilde{R}^{-1}G+\tilde{\sigma}_{\phi}^{-2}I_M \end{bmatrix} \begin{bmatrix} \tilde{\eta} \\ \tilde{\alpha} \\ \tilde{\psi} \\ \tilde{\phi} \end{bmatrix} = \begin{bmatrix} U' \\ D' \\ F' \\ G' \end{bmatrix} \tilde{R}^{-1}\left(y-X\hat{\beta}\right) \tag{12}$$

When $\tilde{R}=\tilde{\sigma}_{\varepsilon}^{2}I_{N^*}$ and $\left(\tilde{\sigma}_{\alpha}^{-2},\tilde{\sigma}_{\psi}^{-2},\tilde{\sigma}_{\phi}^{-2}\right)\rightarrow 0$, the mixed model equations converge to the OLS normal equations solved by a fixed effect estimator. Hence the OLS (fixed effect) estimator is a special case.[17]

From (12), the exact solution for the BLUPs is $\tilde{\Theta}=\tilde{W}Z'\tilde{V}^{-1}\left(y-X\hat{\beta}-U\tilde{\eta}\right)$, where $\tilde{\Theta}=[\ \tilde{\alpha}'\ \ \tilde{\psi}'\ \ \tilde{\phi}'\ ]'$, $Z=[\ D\ \ F\ \ G\ ]$, $V=\sigma_{\alpha}^{2}DD'+\sigma_{\psi}^{2}FF'+\sigma_{\phi}^{2}GG'+R$, and $\tilde{V}$ and $\tilde{W}$ are REML estimates of V and W. This solution does not mechanically normalize match effects to have zero mean for each worker and firm. Instead, wage variation is apportioned between the person, firm, and match effects to minimize mean squared error.[18] This is true even in the case where an individual or firm enters only one match.

BLUPs have a useful Bayesian interpretation. We can think of the moment restrictions (10) and (11) as prior information that identifies the person, firm, and match effects. Specifically, under a normal likelihood for $y-X\hat{\beta}$, an uninformative prior for η, and a normal prior for the random effects with mean (10) and covariance (11), the posterior distribution of (α,ψ,ϕ) is normal with mean $\tilde{\Theta}$ and covariance $\Sigma=W-WZ'\left(M_U^*-P_X^*\right)V\left(M_U^*-P_X^*\right)'ZW$,

[17] That is, the usual OLS fixed effect estimator minimizes the sum of squared residuals in (4). The first order conditions for this minimzation problem are the least squares normal equations. The mixed model equations (12) converge to the least squares normal equations as $\left(\tilde{\sigma}_{\alpha}^{-2},\tilde{\sigma}_{\psi}^{-2},\tilde{\sigma}_{\phi}^{-2}\right)\rightarrow 0$ when $\tilde{R}=\tilde{\sigma}_{\varepsilon}^{2}I_{N^*}$. Hence the usual fixed effect estimator is a special (limiting) case of the hybrid mixed model estimator. For the reasons given in Section 2.1.1, however, (θ,ψ,ϕ) are not identified via a fixed effect estimator without additional identifying restrictions. For example, our orthogonal match effect estimator identifies (θ,ψ,ϕ) by assuming that ϕ is orthogonal to θ and ψ.

[18] Over a long horizon, mean match effects will tend to zero for each worker and firm, because all agents face the same distribution of match quality. In a short panel, however, some agents will be "lucky" or "unlucky," and their true average match effect will depart from zero. The orthogonal match effect estimator assigns such departures to the person and firm effects. The BLUP solution accounts for sample uncertainty over the correct attribution of wage variation to the various effects, and apportions variation among them to minimize MSE.

where $M_U^* = V^{-1} - V^{-1}U\left(U'V^{-1}U\right)^{-1}U'V^{-1}$ is the usual GLS projection matrix and $P_X^* = V^{-1}X\left(X'M_{[D\ F\ G]}X\right)^{-1}X'M_{[D\ F\ G]}$ reflects the fact that $\hat{\beta}$ is estimated in the first stage.[19] The important thing to note is that the posterior covariance Σ is non-diagonal because D, F, and G are non-orthogonal.[20] Consequently, even when priors specify that person, firm, and match effects are uncorrelated, their posterior covariance is explicitly non-zero. Hence "orthogonal priors" as expressed in (11) do not imply orthogonal BLUPs. Indeed, we see this directly from a frequentist expression for the covariance of the BLUPs, $Var\left(\tilde{\Theta}\right) = WZ'\left(M_U^* - P_X^*\right)V\left(M_U^* - P_X^*\right)'ZW$, which is non-diagonal because D, F, and G are non-orthogonal. From a Bayesian perspective, covariation between BLUPs reflects posterior uncertainty over the correct attribution of wage variation to the person, firm, and match effects, conditional on observables.

The hybrid mixed effect estimator has the following properties. $\hat{\beta}$ is consistent and unbiased, and $\tilde{\eta}$ is consistent and the BLUE of η. The REML estimates are consistent, asymptotically normal, and asymptotically efficient in the Cramer-Rao sense; see Jiang (1996). In particular, $\tilde{\sigma}_\alpha^2$ is consistent as $N \to \infty$, $\tilde{\sigma}_\psi^2$ is consistent as $J \to \infty$, and $\tilde{\sigma}_\phi^2$ is consistent as $M \to \infty$, even if the number of observations on each worker, firm, and/or match is fixed. Thus we achieve consistency of variance estimates under much weaker conditions than in the fixed effect case. This solves the incidental parameters problem insofar as it affects variance estimation, and allows us to rely on asymptotic tests for the presence of match effects as $M \to \infty$. The realized random effects $\left(\tilde{\alpha}, \tilde{\psi}, \tilde{\phi}\right)$ are BLUPs, $\tilde{\theta}_i = \tilde{\alpha}_i + u_i'\tilde{\eta}$ is the BLUP of θ_i, and $x_{it}'\hat{\beta} + \tilde{\theta}_i$ is an unbiased predictor of $x_{it}'\beta + \theta_i$ (it is not BLUP because $\hat{\beta}$ is inefficient when errors are non-spherical). Results in Jiang (1998) imply that $\tilde{\alpha}_i, \tilde{\psi}_j$, and $\tilde{\phi}_{ij}$ are consistent as T_i, N_j, and T_{ij} tend to infinity.

[19] The full derivation is available from the author. When β is known or zero, the posterior variance reduces to the usual $\Sigma = W - WZ'M_U^*ZW$, and $Var\left(\tilde{\Theta}\right) = WZ'M_U^*ZW$.

[20] It is unlikely that D and F will be orthogonal in the typical application: $D'F$ is an $N \times J$ matrix of match durations. D and F are never orthogonal to G, because they lie within the column space of G: the column of G corresponding to the match between worker i and firm j is the elementwise product of the i^{th} column of D and the j^{th} column of F.

2.2 Bias Due To Omitted Match Effects

We find substantial differences between estimates of the match effects model and the person and firm effects model in our application. Given the growing use of the person and firm effects model in applied work, it is important to understand the nature of these differences. Our primary interest is how omitted match effects bias least squares estimates of β, θ, and ψ. We derive and interpret that bias here. Omitted match effects also bias sample estimates of the variance and covariance of θ and ψ; we derive that bias in Appendix A.

When the data generating process is given by equation (3) but the estimated equation excludes match effects, least squares estimates of the parameters, β^*, θ_i^*, and ψ_j^*, are biased:

$$E[\beta^*] = \beta + \left(X' M_{[D\ F]} X\right)^{-1} X' M_{[D\ F]} G\phi \quad (13)$$

$$E[\theta^*] = \theta + \left(D' M_{[X\ F]} D\right)^{-} D' M_{[X\ F]} G\phi \quad (14)$$

$$E[\psi^*] = \psi + \left(F' M_{[X\ D]} F\right)^{-} F' M_{[X\ D]} G\phi. \quad (15)$$

The expected value of estimated returns to observables, β^*, equals true returns plus a bias term whose sign and magnitude depends on the conditional covariance between X and match effects, conditional on D and F. Intuitively, if workers with particular characteristics (e.g., more experience) sort into better matches than others, the estimated returns to those characteristics reflect true returns plus the returns to sorting. Our within-match estimator (7) corrects this bias.

Estimated person and firm effects are biased whenever omitted match effects are nonzero, because D and F lie within the column space of G. This is intuitive: D contains information on worker identities ("who you are"), F contains information on firm identities ("where you work"), and G contains information on match identities ("who you are and where you work"). The bias term in (14) is an employment duration-weighted average of omitted match effects, conditional on X and F. In the simplest case where X and F are orthogonal to D, the bias is $E[\theta_i^*] - \theta_i = T_i^{-1} \sum_t \phi_{i\mathcal{J}(i,t)}$, where $\mathcal{J}(i,t) = j$ indicates worker i's employer in period t. Similarly, omitted variable bias in ψ^* is a duration-weighted average of omitted match effects, conditional on X and D. When X and D are orthogonal to F, the bias is $E\left[\psi_j^*\right] - \psi_j = N_j^{-1} \sum_{i \in \mathcal{I}_j} T_{ij} \phi_{ij}$ where $\mathcal{I}_j = \{i : \mathcal{J}(i,t) = j \text{ for some } t\}$ is the set of all employees of firm j.

3 Data

Our application uses data from the US Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) database. These data span forty-six states that represent the majority of American employment. We use data from two participating states (whose identity is confidential) that are broadly representative of the LEHD database.[21]

The LEHD data are administrative, constructed from Unemployment Insurance (UI) system employment reports. These are collected by each state's Employment Security agency to manage the unemployment insurance program. Employers are required to report total payments to all employees on a quarterly basis. These payments (earnings) include gross wages and salary, bonuses, stock options, tips and gratuities, and the value of meals and lodging when these are supplied (Bureau of Labor Statistics (1997, p. 44)). The UI reports contain only limited information: worker and firm identifiers and earnings. The LEHD database integrates these with internal Census Bureau data sources to add demographic and firm characteristics.

The coverage of UI data varies slightly from state to state, though the Bureau of Labor Statistics (1997, p. 42) claims that UI coverage is "broad and basically comparable from state to state" and that "over 96 percent of total wage and salary civilian jobs" were covered in 1994. See Abowd et al. (forthcoming) for further details. With the UI employment records as its frame, the LEHD data comprise the universe of employment at firms required to file UI reports: all employment potentially covered by the UI system in participating states.

[21] Computational considerations dictate that we restrict our analysis to a subset of the data. In a small sample drawn from all states, work histories are not sufficiently connected to precisely estimate the person, firm, and match effects. Hence we focus on two representative states.

We aggregate the quarterly data to the annual level. The full sample consists of over 49 million annualized employment records on full-time workers between 25 and 65 years of age who were employed at a private-sector non-agricultural firm between 1990 and 1999.

Solving the mixed model equations (12) is computationally intensive, so we are obliged to estimate the hybrid mixed model on a subsample. Sampling from linked employer-employee data is nontrivial because the sample must be highly connected to precisely estimate the person, firm, and match effects. We therefore draw a ten percent "dense" subsample of individuals employed in 1997 using the procedure described in Appendix B. Our procedure ensures that each worker is connected to at least five others by a common employer, but is otherwise representative of the population of individuals employed in 1997. That is, all individuals employed in 1997 have an equal probability of being sampled. To facilitate comparisons between the fixed and mixed effect specifications, we estimate all specifications on the dense subsample. In Appendix B, we investigate whether our results are sensitive to sampling, and find no evidence that they are.

Table 1 presents characteristics of the samples (see Appendix Table 1 for variable definitions). Individuals employed in 1997 are largely representative of the full sample. Minor differences indicate that they have a slightly stronger labor force attachment than the universe of all individuals employed between 1990 and 1999: males are slightly over-represented, as are individuals with higher educational attainment and those who work four full quarters in an average year. The ten percent subsample has characteristics virtually identical to the sample of all individuals employed in 1997, although our procedure over-samples large firms.

4 Estimation Results

Table 2 reports regression coefficients in a fixed effect specification of the person and firm effects model and our two specifications of the match effects model. Estimates of the person and firm effects model are consistent with earlier work. There are important differences, however, between these estimates and those of the match effects model. Notably, the person and firm effects model over-estimates the returns to experience; see Figure 1. This specification estimates that a male worker with 25 years of labor market experience earns 0.76 log points (113 percent) more than a labor market entrant, all else equal. The comparable differential for women is 0.57 log points (76 percent). The within-match estimator (7) yields a much flatter experience profile: the earnings gap associated with 25 years of experience is 0.53 log points (69 percent) for men and 0.36 log points (43 percent) for women. Hence the person and firm effects model over-estimates the returns to 25 years of experience by 0.23 log points (44 percent of the earnings of a labor market entrant) for men, and 0.21 log points (33 percent of an entrant's earnings) for women.

The discrepancy arises because (7) identifies the returns to experience entirely from within-match variation. In contrast, the person and firm effects model attributes some of the earnings growth that occurs when individuals change employer to labor market experience. Consequently, the estimated return to labor market experience (i.e., the accumulation of general human capital) partly reflects wage growth due to employment mobility. Workers could experience higher wage growth when they change jobs because they sort into higher-paying firms or better matches. Sorting into higher-paying firms can not account for the bias, because the person and firm effects model includes firm effects. Rather, (13) makes clear that the bias is due to conditional covariation between experience and omitted match effects, i.e., because more experienced individuals sort into better matches on average, all else equal. When match effects are omitted, the returns to sorting are incorrectly attributed to labor market experience.

Table 3 presents the estimated variance of log earnings components. In all three specifications, person effects exhibit the greatest dispersion and the market value of time-varying characteristics ($x_{it}'\hat{\beta}$) exhibits the least. This is consistent with prior research based on the person and firm effects model. The two fixed effect specifications deliver nearly identical estimates of the variance of person and firm effects: about 0.29 squared log points and 0.08

squared log points, respectively. The hybrid mixed model attributes less variation to person effects (0.198 squared log points), slightly more to firm effects (0.102 squared log points), and much more variation to match effects than the orthogonal estimator: 0.079 versus 0.016 squared log points.

A more intuitive measure of the relative importance of these components is the proportional decomposition of the variance of log earnings in the lower panel of the table.[22] The person and firm effects model attributes 73 percent of the variance of log earnings to person-specific determinants of match productivity ($x_{it}'\beta+\theta_i$) and 16 percent to firm-specific factors. Results for the orthogonal match effects model are very similar, with about 5 percent of the variance of log earnings attributed to match effects and the unexplained component reduced by a corresponding amount. Results for the hybrid mixed model are quite different. It attributes much less variation to the person-specific component (53 percent), and considerably more to firm and match effects (22 and 16 percent, respectively). Clearly, matching is an important determinant of earnings: firm- and match-specific factors collectively account for nearly 40 percent of its variance. In particular, the fraction of earnings variation attributed to match effects is about three times that of time-varying observables such as experience.

Table 3 also reports a test for the presence of match effects. For the orthogonal match effects estimator, the null hypothesis is $H_0 : \phi_{ij} = 0$ for each i, j pair in the data. This implies $M - N - J = 323,476$ linear restrictions.[23] Given the huge number of restrictions, it is no surprise that we easily reject the null of no match effects at conventional levels.[24] In the hybrid mixed model, the null of no match effects is $H_0 : \sigma_\phi^2 = 0$. We test this hypothesis with a likelihood ratio test based on the REML log-likelihoods (a REML ratio test, or REMLRT)

[22] The decomposition is based on $Var(y_{ijt}) = Cov\left(y_{ijt}, x_{ijt}'\hat{\beta}\right) + Cov\left(y_{ijt}, \hat{\theta}_i\right) + Cov\left(y_{ijt}, \hat{\psi}_j\right) + Cov\left(y_{ijt}, \hat{\phi}_{ij}\right) + Cov(y_{ijt}, e_{ijt})$, where $\hat{\beta}, \hat{\theta}_i, \hat{\psi}_j, \hat{\phi}_{ij}$ are sample estimates and e_{ijt} is the corresponding residual. Dividing both sides by $Var(y_{ijt})$ gives a proportional decomposition similar to one in Gruetter and Lalive (2004) for the person and firm effects model.

[23] When the data consist of $\mathcal{G}$ connected groups of observations, there are $N^* + \mathcal{G} - N - J - k - 1$ degrees of freedom for β in the person and firm effects model, and $N^* + \mathcal{G} - M - k - 1$ degrees of freedom in the match effects model.

[24] The value of the Wald statistic is around 1.47 million.

of specifications with and without match effects. Once again, we easily reject the null of no match effects at conventional significance levels.[25]

Matching models predict a strong relationship between match-specific heterogeneity in wages and employment mobility. In the Jovanovic (1979) model, for example, wages are higher in longer duration jobs, because good matches survive longer than bad ones. It is therefore natural to ask how match effects (and other wage components) are related to employment duration. Figure 2 plots average person, firm, and match effects from the hybrid mixed model specification by completed job duration (tenure). Reported means are for the subset of employment spells where job duration is uncensored, i.e., spells that begin and end during the sample period.[26] The well known cross-sectional relationship between earnings and tenure is clearly evident in all three components: the mean of each effect increases with the eventual duration of the job. Consistent with search and matching models, and with intuition based on our model of match productivity, jobs last longer when workers receive a high wage relative to their alternatives (i.e., when either the firm or match effect is large). Person effects exhibit the strongest unconditional relationship with job duration, which is in line with the high variance of person effects relative to firm and match effects. Intuitively, individuals whose characteristics are highly valued in the labor market are less likely to experience layoffs or quits. These relationships also hold conditionally.[27]

4.1 Matching and Sorting

A long-standing question is whether "good" workers sort into employment at "good" firms. The recent literature on search and matching with heterogeneous agents (e.g., Shimer and Smith (2000), Shimer (2005), Woodcock (2010)) has renewed interest in this question, and

[25] The null distribution of the REMLRT statistic is a $50:50$ mixture of a χ^2_0 and a χ^2_1; see Stram and Lee (1994). The REMLRT statistic exceeds 35 thousand.

[26] Tenure is left-censored for spells that were active in the first quarter of the sample.

[27] We estimate a linear probability model to predict match dissolution conditional on x_{it}, u_i, a quartic in current job tenure, and BLUPS from the wage equation ($\tilde{\theta}_i, \tilde{\psi}_j$, and $\tilde{\phi}_{ij}$). The dependent variable is a binary indicator $s_{it} = 1$ if the match dissolves in period t and zero otherwise. The estimated coefficients on $\tilde{\theta}_i, \tilde{\psi}_j$, and $\tilde{\phi}_{ij}$ are -0.076, -0.133, and -0.072, respectively, with very small bootstrap standard errors.

researchers have sought an answer in the estimated correlation between person and firm effects. Perhaps surprisingly, Abowd et al. (2004) find a near-zero correlation in US data, and a moderate negative correlation in France. Other authors report similar findings in other countries.[28]

Sample correlations between personal heterogeneity in earnings (both θ_i and $x_{it}'\beta + \theta_i$) and firm heterogeneity (ψ_j) are given in Table 3. Like prior US studies, we find a near-zero correlation in both fixed effect specifications. In contrast, we find a modest positive correlation (0.185) in the hybrid mixed model.[29] This is evidence that high-productivity workers sort into employment at high-productivity firms on average, i.e., that matching is positively assortative. This result is consistent with our specification of match productivity in (1).[30]

It is important to understand why previous work has failed to find evidence of assortative matching.[31] We have already shown that omitted match effects bias estimated person and firm effects in equations (14) and (15). In Appendix A, we show that omitted match effects also bias the estimated correlation between person and firm effects. The sign of the bias is ambiguous, but our estimates imply that the bias is toward zero. In Appendix A, we give a sufficient condition for this result, and some supporting simulations. Compounding the bias from omitted match effects is the limited mobility bias noted by Andrews et al. (2008), which biases the sample correlation between θ_i and ψ_j toward zero in the person and firm effects model. Our simulations indicage that limited mobility bias is modest in the person

[28] Andrews et al. (2008) and Cornelißen and Hübler (2007) both find a negative correlation in Germany, as do Barth and Dale-Olsen (2003) in Norway and Gruetter and Lalive (2004) in Austria; Maré and Hyslop (2006) report a modest positive correlation in New Zealand.

[29] The non-zero correlation is not a feature of random versus fixed effects *per se*: random effect estimates of the person and firm effects model also yield a near-zero correlation.

[30] Becker (1973) shows that positively assortative matching (PAM) is optimal in a frictionless environment when the production function is strictly supermodular, i.e., $\partial^2 Q_{ij}/\partial L_i \partial K_j > 0$. When there are search frictions, Shimer and Smith (2000) show that PAM also requires supermodularity of $\ln \partial Q_{ij}/\partial L_i$ and $\ln \partial^2 Q_{ij}/\partial L_i \partial K_j$.

[31] Lopes de Melo (2009) provides a complementary explanation to that advanced here. Specifically, in a frictional search environment where there are production complementarities between workers and firms, firms face capacity constraints in posting new vacancies, and agents bargain over wages, he shows that wages are non-monotone in firm productivity.

and firm effects model, and we find none in the hybrid mixed model.

Intuition might lead us to expect that omitted match effects would bias the estimated correlation between person and firm effects upward. When match effects are omitted, for example, a good match $(\phi_{ij} > 0)$ will impart positive bias to both $\hat{\theta}_i$ and $\hat{\psi}_j$ and thereby inflate the estimated correlation between person and firm effects. In fact we observe the reverse. The main reason is that bias from omitted match effects is highly asymmetric: it loads predominantly onto $\hat{\theta}_i$. We see evidence of this in Table 3, where the person and firm effects model substantially overestimates the variance of person effects but not firm effects. We find the same in our simulations. It is also apparent in the sample correlation between BLUPs of ϕ_{ij} and estimates of θ_i and ψ_j from the person and firm effects model: 0.49 and 0.04, respectively. The underlying cause is the imbalance between N and J. Firms in our sample match with many workers, but workers only match with a few firms.[32] For the typical firm, positive and negative omitted match effects average out over many matches, imparting no systematic bias to $\hat{\psi}_j$; but there is no such tendency for workers who only enter a few matches. It is clear that when matches are perfectly assortative (i.e., $\theta_i = \psi_j$ for all matches) then *any* asymmetry in bias to $\hat{\theta}_i$ versus $\hat{\psi}_j$ will bias the estimated correlation toward zero. Outside of this case, it is easy to show that when bias from omitted match effects loads entirely onto $\hat{\theta}_i$, the estimated correlation between person and firm effects is biased toward zero.[33] In this case, omitted match effects are like measurement error, attenuating the estimated relationship between person and firm effects.

Unlike classical measurement error, however, the magnitude of bias from omitted match effects depends on job duration. The best matches (where $\phi_{ij} > 0$ and large) last longest. In a short panel, individuals who experience a long employment spell enter fewer matches overall. Because the bias is a weighted average of omitted match effects, it will therefore

[32] We have at most ten observations per worker; the median worker is observed in 2 matches. Some large firms in our sample enter many thousands of matches; the mean is about 8.

[33] Suppose ϕ is uncorrelated with θ and ψ and $Cov(\theta, \psi) > 0$. In the extreme case where $E\left[\hat{\theta}\right] = \theta + \phi$ and $E\left[\hat{\psi}\right] = \psi$, we have $Corr\left(\hat{\theta}, \hat{\psi}\right) = Cov(\theta, \psi) / \sqrt{(Var(\theta) + Var(\phi)) Var(\psi)} < Cov(\theta, \psi) / \sqrt{Var(\theta) Var(\psi)} = Corr(\theta, \psi)$.

tend to be larger in long spells. Compounding this, note that correlations reported in Table 3 are implicitly duration-weighted. Each wage record receives equal weight, so longer-lasting spells receive greater weight than short spells. This further amplifies the bias from omitted match effects. Again, this is confirmed in our simulations.

To summarize, our estimates of the hybrid mixed model indicate that good workers do, on average, sort into employment at good firms. Complex interactions between employment duration, mobility, and sample structure combine to impart downward bias on the estimated correlation between person and firm effects when match effects are omitted. Coupled with the limited mobility bias noted by Andrews et al. (2008), it is little wonder that previous research has found little evidence of assortative matching in the labor market.

4.2 Employment Mobility and the Importance of Matching

Bartel and Borjas (1981), Topel and Ward (1992), and others have shown that workers experience above-average wage growth when they change jobs. This could reflect sorting into higher-paying firms or better matches. In Table 4, we decompose the mean change in log earnings when individuals change employer by noting that when a worker moves from employer j to employer n (in periods t and s, respectively), the total change in log earnings is:

$$y_{ins} - y_{ijt} = (x'_{is} - x'_{it})\,\hat{\beta} + \left(\hat{\psi}_n - \hat{\psi}_j\right) + \left(\hat{\phi}_{in} - \hat{\phi}_{ij}\right) + (e_{ins} - e_{ijt}) \tag{16}$$

where e is the residual.

The average annual change in real log earnings is 0.032 log points. Year-over-year earnings growth is slightly lower when individuals do not change employer (0.029 log points), and considerably larger when they do (0.049 log points). In all specifications, the component of earnings growth due to time-varying characteristics is similar to the overall annual growth rate. In Panel A, the average change in match effects among all job changers is near zero; implying that most of the excess wage growth upon job change is due to sorting into higher-paying firms. At first glance, this result might seem surprising. However, the estimates in Panel A mask considerable heterogeneity. In Panels B and C, we divide employment transitions into two groups: job-to-job transitions, and those that include a period of non-employment. Because we do not observe the exact start and end dates of employment spells, we define job-to-job transitions as employment spells that overlap in at least one quarter. With the exception of transitions where the start/end dates of spells coincide exactly with the beginning/end of quarters, the remaining transitions include a spell of non-employment. Most of these will reflect unemployment, but note that non-employment could be voluntary, could reflect employment not covered by the UI reporting system, or in a state outside our sample.

Year-over-year log earnings growth in the subset of job-to-job transitions is large: 0.082 log points, which is nearly triple the annual wage growth of job stayers. Of this, the person and firm effects model attributes about equal shares to the change in time-varying covariates and firm effects. The orthogonal match effects specification delivers similar estimates and attributes virtually all of the remainder (0.022 log points, or 27 percent) to match effects. Hybrid mixed model estimates are similar, but attribute slightly less wage growth to match effects (0.015 log points, or 18 percent of the total). Of the 0.053 log point difference between earnings growth of job-to-job movers and job stayers, the hybrid mixed model estimates that 60 percent is due to sorting into higher-paying firms and 29 percent to sorting into better matches.

In contrast, individuals who experience a spell of non-employment have much lower earnings growth when they change jobs: 0.02 log points, or only 68 percent of the annual wage growth of job stayers. The person and firm effects model attributes most of the lower wage growth to the residual component. Both match effects specifications attribute almost all of the lower earnings growth to changes in match effects, implying that when individuals experience a spell of non-employment, their wage growth suffers because they sort into lower-paying matches. Presumably, unemployed job seekers are less concerned about finding good matches than employed job seekers.

4.3 Robustness and Alternate Explanations

In this section, we assess the robustness of our results and evaluate some potential alternative explanations for our findings.

Are there genuine person and firm effects? Our simple model in Section 2 posits that productivity varies across workers and firms, leading to earnings differences measured by θ_i and ψ_j. An extreme possibility, however, is that empirical person and firm effects are simply an estimation artifact. Consider a special case of the model where all workers and firms are equally productive ($L_i = 1$ and $K_j = 1$ for all i, j) and where p_j and γ_j are the same for all firms. In this model, all productivity and wage variation is due to match quality, M_{ij}, and there are no genuine person and firm effects in wages. However, empirical person and firm effects are identified given match effects: empirical person effects are worker-specific aggregates of match effects, and empirical firm effects are firm-specific aggregates of match effects. A similar concern arises when there is a genuine person or firm effect in productivity, but not both.

It is clear that we can distinguish genuine person and firm effects from aggregates of match effects in a long panel, because our estimators of θ_i and ψ_j are consistent as $T_i \to \infty$ and $N_j \to \infty$. The question is whether we can do so in a short panel. The answer is yes, in the case of the hybrid mixed model. REML estimates of the variance of person and firm effects are consistent as $N \to \infty$ and $J \to \infty$. If there are no genuine person and firm effects, REML estimates of σ^2_α and σ^2_ψ will approach zero when N and J are large. We test this hypothesis with a REMLRT, and easily reject the hypothesis that $\sigma^2_\alpha = 0$, the hypothesis that $\sigma^2_\psi = 0$, and the joint hypothesis that both are zero (in each case the p-value is less that 10^{-5}).

Is there a genuine match effect? We have rejected the hypothesis that $\sigma_\phi^2 = 0$. However, matching may imply richer wage dynamics than are captured by a match effect. Match effects could vary over time as workers acquire match-specific human capital, either as a consequence of learning-by-doing or learning about match quality. In either case, the within-match intertemporal correlation in wages is more complex than implied by the match effects model. We test the match effects model against more general alternatives by estimating a model with random person and firm effects, where errors have an unstructured within-match covariance. The within-match error covariance is a $T \times T$ symmetric matrix, where $T = \max_{ij} \{T_{ij}\} = 10$ and the element in row r and column c, σ_{rc}, is the within-match covariance between errors at r and c years of job tenure. This nests error structures implied by learning about match quality under various distributional assumptions (see Woodcock (2010) for an example), error structures implied by time-varying match effects, and cases where errors are heteroskedastic and serially correlated within matches. It also nests the match effects model, where $\sigma_{rc} = \sigma_\phi^2$ for all $r \neq c$, and $\sigma_{rc} = \sigma_\phi^2 + \sigma_\varepsilon^2$ for $r = c$.

The general model increases computational burden, because we need to estimate 55 distinct covariance terms instead of two variance components (σ_ϕ^2 and σ_ε^2). We are consequently obliged to estimate the model on a smaller (one percent) dense sample. We further restrict the sample to spells that began during the sample period, because the covariance structure depends on job tenure (not observed for left-censored spells). The sample remains large, $N^* = 228,386$. Parameter estimates in the general model are very similar to those in Table 3: $\tilde{\sigma}_\alpha^2 = 0.18$, $\tilde{\sigma}_\psi^2 = 0.08$, and $Corr\left(\tilde{\theta}_i, \tilde{\psi}_j\right) = 0.22$.[34] We fail to reject the restrictions implied by the match effects model using a REMLRT (p-value$= 0.117$). This indicates that a constant match effect adequately captures wage dynamics. Possible explanations are that learning about match quality is very fast (Woodcock (2010) and Nagypal (2007) find that most learning occurs in the first 2 years) and/or there is not much systematic variation in match effects over time.

[34] See Woodcock (2010), Appendix Tables 1 and 2, for complete estimates of this specification.

Is a more efficient estimator available? Our hybrid mixed model imposes no restrictions on the relationship between time-varying observable characteristics and unobservables. However, if x_{it} is uncorrelated with person, firm, and match effects, then a traditional random effect estimator is a more efficient alternative. This does not seem implausible, since sample correlations between $x_{it}'\hat{\beta}$ and BLUPs are 0.05 or less in absolute value. Column 1 of Table 5 gives estimates of a traditional random effects specification. They are very similar to our hybrid mixed model estimates. However, a Hausman test easily rejects the hypothesis that x_{it} is uncorrelated with unobservables, leading us to prefer the hybrid mixed model.

Is the error covariance correctly specified? Our hybrid mixed model estimates in Table 3 assume spherical errors, $R = \sigma_{\varepsilon}^2 I_{N^*}$. If individual errors are serially correlated, we might erroneously attribute wage persistence arising from serial correlation to match effects. Column 2 of Table 5 reports estimates where individual errors follow an ARMA(1,1) process: $\varepsilon_{ijt} = \rho\varepsilon_{i\mathcal{J}(i,t-1)t-1} + \nu u_{i\mathcal{J}(i,t-1)t-1} + u_{ijt}$, where u_{ijt} is an iid error, ρ is the autoregressive coefficient, and ν is the moving average coefficient. We have also estimated AR(1), AR(2), MA(1), and MA(2) specifications (not reported, but available on request). The ARMA(1,1) specification gives a better fit to the data than these alternatives (as measured by AIC and BIC) and yields the most significant departure in parameter estimates versus the spherical errors case. That said, the estimated variance of firm and match effects are virtually unchanged. However, the estimated variance of person effects is much reduced because errors are highly persistent at the individual level: $\rho = 0.91$ and $\nu = 0.22$. This does not affect our conclusions about the importance of matching, but it does provide a more flexible representation of earnings persistence.

Is there a selection bias? We only observe log earnings, y_{ijt}, if we observe a match between worker i and firm j in period t. There may be a selection bias, because not all matches are observed in all periods and matching is probably not random. In particular, our identification condition (6) is violated if matching depends on ε_{ijt}.

There are two potential sources of selection bias. First, observed matches probably comprise a nonrandom subset of all potential matches. A match only forms if productivity is large enough to compensate agents for foregoing their next best alternative. There are over 288 billion potential matches between workers in our 10 percent sample and firms active during the sample period, however, so it is intractable to model which matches form and correct possible bias from nonrandom match formation.[35] However, bias from nonrandom match formation is probably small, since we would expect match formation to depend on the permanent component of match productivity (measured by $\mu + x'_{it}\beta + \theta_i + \psi_j + \phi_{ij}$), rather than transitory factors in ε_{ijt}. Nevertheless, our estimates should be interpreted as representative of matches that actually form, as opposed to all potential matches.

The second source of bias is nonrandom match duration. Conditional on formation, matches only endure if they are "good enough." This may incidentally truncate the empirical distribution of match effects. There is an additional bias if mobility depends on ε_{ijt}. We correct these biases following Heckman (1979). We estimate a probit model of the probability that the match between worker i and firm j continues to period t. The probit model controls for observable characteristics x_{it} and u_i, indicators for job tenure, the annual average of the quarterly separation rate from firms in the same industry and county, and a random match effect. We use probit estimates to construct the familiar inverse Mill's ratio term, λ_{ijt}, and include it among the covariates in (8). The selection-corrected estimates are reported in column 4 of Table 5.[36] They are based on the subset of spells that are not left-censored, because the probit model includes controls for job tenure. The inverse Mill's ratio term has

[35] An emerging literature on match formation in social networks, e.g., Snijders (2001), models networks as a bipartite random graph. Similar methods could be applied to the labor market: workers and firms are nodes, an employment relationship is the edge that connects them, and their characteristics determine the probability of edge (match) formation. These methods hold promise for the future. However, the computational demands remain formidable, and current applications are limited to networks of several thousand agents, e.g., Goodreau (2007).

[36] Probit estimates are available on request. The separation rate is based on the restricted access version of the US Census Bureau's Quarterly Workforce Indicators (QWI) at the county by two-digit SIC level. The probit excludes person and firm effects for computational reasons; the match effect should be sufficient to capture unobserved heterogeneity in match duration. Note λ_{ijt} is based on the predicted probability of match continuation, which includes the realization of the probit match effect.

the expected sign: it is negative, indicating that wages are lower in matches that have a low probability of continuing (where λ_{ijt} is large). The appropriate comparison for other parameters is column 3, which gives hybrid mixed model estimates after excluding left-censored spells. Estimates in columns 3 and 4 are virtually identical, so there is no evidence that nonrandom match duration causes bias to the parameters of interest.

Do match effects measure the returns to job tenure? We have already seen that match effects are larger, on average, in longer-lasting jobs. Our empirical specification does not control for job tenure, so it is natural to ask whether match effects simply measure the average return to tenure in the match. Column 5 of Table 5 reports estimates of our hybrid mixed model with controls for job tenure, again excluding left-censored spells. To avoid functional form restrictions on the returns to tenure, this specification includes a complete set tenure indicators. In line with Abraham and Farber (1987) and Altonji and Shakotko (1987), we find modest returns to tenure. They reach a maximum of 0.06 log points at six years, and decline slightly thereafter. Other parameter estimates, including the variance of match effects, are virtually unchanged from those reported previously.

Endogeneity is a central concern in most studies of the returns to tenure, because tenure is likely correlated with person, firm, and match effects. Our within-match estimator of the returns to tenure is consistent in the presence of such correlation. However, it is biased if mobility (and hence tenure) depends on ε_{ijt}. To account for this, we apply the selection correction described above. The selection-corrected estimates are reported in column 6. They are very similar to previous estimates, although the coefficient on the selection term is considerably larger, as are the estimated returns to tenure. However, the variance of the match effect is virtually unchanged, so there remains no evidence that match effects measure the average return to tenure.

Can we discriminate between heterogeneity and time–varying dependence? An alternate explanation for the dynamics captured by the match effects model is state dependence. For instance, firms might set wages as a function of past wages. Individuals' reservation wages might also depend on past wages, introducing time-varying dependence on earnings. Both cases could induce spurious person, firm, and match effects. Thus we augment the match effects model with a control for lag earnings, $y_{i\mathcal{J}(i,t-1)t-1}$. The within-match estimator of its coefficient is biased and inconsistent, because the deviation of lag earnings from its match-specific mean, $y_{i\mathcal{J}(i,t-1)t-1} - \bar{y}_{i\mathcal{J}(i,\cdot-1)\cdot}$, is correlated with the within-match error, $\nu_{ijt} = \varepsilon_{ijt} - \bar{\varepsilon}_{ij\cdot}$. In place of the within-match estimator, we therefore use an IV estimator based on first differences within the match. First-differencing eliminates person, firm, and match effects, but the differenced lag value, $\Delta y_{ijt-1} = y_{ijt-1} - y_{ijt-2}$, is correlated with the differenced error term, $\Delta\varepsilon_{ijt} = \varepsilon_{ijt} - \varepsilon_{ijt-1}$. As in Anderson and Hsiao (1981), we therefore use y_{ijt-2} as an instrument for Δy_{ijt-1}. This limits the sample to spells that last three or more years, but yields consistent estimates. They are presented in column 7 of Table 5. The coefficient on y_{ijt-1} is quite small, presumably reflecting the importance of experience and other time-varying characteristics. The estimated variance of the person effect is smaller than we found before, indicating that some persistence in individual earnings reflects dynamic dependence. The estimated variance of firm and match effects are marginally smaller, but they still explain 20 and 15 percent of the variance of log earnings, respectively.

5 Conclusion

We have shown that match effects are an important source of earnings dispersion. They explain about 16 percent of variation in log earnings, and much of the wage change that occurs when individuals change job. Worker-specific determinants of match productivity explain more than half of the variation in log earnings, which implies considerable persistence in individual earnings. However, the potential returns to job search are large, as is the loss associated with displacement from a high-paying firm or a good match: 0.32 log points for a firm one standard deviation above the mean, and 0.28 log points for a match one standard deviation above the mean.

Omitted match effects cause substantial bias in parameter estimates. Our orthogonal match effects estimator provides a partial solution. It corrects bias in β, which we found was substantial in the estimated returns to experience. However, its ability to correct bias in the estimated person and firm effects is limited: if match effects are truly orthogonal to person and firm effects, then there is no bias to correct; otherwise, the orthogonal match effects estimator is mis-specified. The hybrid mixed model is therefore a compelling alternative. It corrects bias in β and yields consistent estimates of the variance components under minimal assumptions. Using this estimator, we find evidence that good workers sort into employment at good firms, contrary to most previous findings. Together, these results demonstrate that ignoring the role of matching in wage determination produces misleading conclusions about sorting in the labor market.

There remain some open questions. Of particular interest is how the relative importance of person, firm, and match heterogeneity in wages varies over the business cycle, between industries, and geographically. Is matching more important in some industries than others? During layoff events, do firms terminate the "bad" matches first? We leave these and other important questions for future research.

A Appendix: Bias in the Estimated Variance and Covariance of Person and Firm Effects

Full derivation of the following bias expressions is available on request. Let $A = I_{N^*} - \frac{1}{N^*}\iota_{N^*}\iota'_{N^*}$ be the matrix that takes deviations from sample means. When match effects are omitted, the expected value of the sample variance of the least squares estimate of person effects is:

$$\begin{aligned} E\left[\widehat{Var}\left(\hat{\theta}\right)\right] &= \frac{1}{N^*-1}\theta D'AD\theta + \frac{\sigma_\varepsilon^2}{N^*-1}tr\left[\left(D'M_{[X\ F]}D\right)^- D'AD\right] \\ &+\frac{1}{N^*-1}tr\left[\phi'G'M_{[X\ F]}D\left(D'M_{[X\ F]}D\right)^- D'AD\left(D'M_{[X\ F]}D\right)^- D'M_{[X\ F]}G\phi\right] \\ &+\frac{2}{N^*-1}tr\left[\theta'D'AD\left(D'M_{[X\ F]}D\right)^- D'M_{[X\ F]}G\phi\right]. \end{aligned} \tag{17}$$

The first term is the sample variance of the true person effects. The second term is least squares estimation error. It is identical to the bias discussed by Andrews et al. (2008) in the absence of match effects, is positive, and proportional to N/N^*. The third term is the average within-person variance of ϕ, conditional on X and F and weighted by employment duration. It is positive. The final term is twice the duration-weighted covariance between θ and ϕ, conditional on X and F. Its sign is indeterminate. However, when the conditional covariance between θ and ϕ is zero, $\widehat{Var}\left(\hat{\theta}\right)$ is biased upward.

The expected value of the sample variance of least squares estimates of firm effects is:

$$\begin{aligned} E\left[\widehat{Var}\left(\hat{\psi}\right)\right] &= \frac{1}{N^*-1}\psi'F'AF\psi + \frac{\sigma_\varepsilon^2}{N^*-1}tr\left[\left(F'M_{[X\ D]}F\right)^- F'AF\right] \\ &+\frac{1}{N^*-1}tr\left[\phi'G'M_{[X\ D]}F\left(F'M_{[X\ D]}F\right)^- F'AF\left(F'M_{[X\ D]}F\right)^- F'M_{[X\ D]}G\phi\right] \\ &+\frac{2}{N^*-1}tr\left[\psi'F'AF\left(F'M_{[X\ D]}F\right)^- F'M_{[X\ D]}G\phi\right]. \end{aligned} \tag{18}$$

Again, the first term is the sample variance of true firm effects and the second term is least squares estimation error that also arises in the absence of match effects. The estimation error is positive and proportional to J/N^*. The third term is the average within-firm variance of ϕ, conditional on X and D and weighted by employment duration, and is positive. The final term is twice the duration-weighted covariance between ψ and ϕ, conditional on X and D. Its sign is indeterminate. However, as in the case of person effects, $\widehat{Var}\left(\hat{\psi}\right)$ is biased upward when ψ and ϕ are conditionally uncorrelated.

When match effects are omitted, the sample covariance between $\hat{\theta}$ and $\hat{\psi}$ satisfies:

$$\begin{aligned} E\left[\widehat{Cov}\left(\hat{\theta},\hat{\psi}\right)\right] = & \frac{1}{N^*-1}\theta D'AF\psi - \frac{\sigma_\varepsilon^2}{N^*-1}tr\left[\left(I_{N^*}-M_{[X\ D]}\right)AF\left(F'M_{[X\ D]}F\right)^{-}F'\right] \\ & +\frac{1}{N^*-1}tr\left[\phi'G'M_{[X\ F]}D\left(D'M_{[X\ F]}D\right)^{-}D'AF\left(F'M_{[X\ D]}F\right)^{-}F'M_{[X\ D]}G\phi\right] \\ & +\frac{1}{N^*-1}tr\left[\theta'D'AF\left(F'M_{[X\ D]}F\right)^{-}F'M_{[X\ D]}G\phi\right] \\ & +\frac{1}{N^*-1}tr\left[\psi'F'AD\left(D'M_{[X\ F]}D\right)^{-}D'M_{[X\ F]}G\phi\right]. \end{aligned} \tag{19}$$

The first term is the true sample covariance between person and firm effects, and the second term is least squares estimation error. The estimation error can be interpreted as the conditional variance of employment duration, weighted by T_i and N_j. It is zero when D and F are orthogonal, and in the balanced data case where each worker is employed at each firm and all spells have equal duration. Otherwise, this term imparts downward bias on $\widehat{Cov}\left(\hat{\theta},\hat{\psi}\right)$ that is proportional to $(N/N^*)^{1/2}(J/N^*)^{1/2}$. The third term is the duration-weighted covariance between person-average match effects $(T_i^{-1}\sum_j T_{ij}\phi_{ij})$ and firm-average match effects $(N_j^{-1}\sum_i T_{ij}\phi_{ij})$, conditional on X. When there is no covariation between match effects, this reduces to the conditional variance of ϕ_{ij}, and biases $\widehat{Cov}\left(\hat{\theta},\hat{\psi}\right)$ upward. The fourth term is the duration-weighted covariance between firm-average person effects $(N_j^{-1}\sum_i T_{ij}\theta_i)$ and match effects, conditional on X. The final term is the duration-weighted covariance between person-average firm effects $(T_i^{-1}\sum_j T_{ij}\psi_j)$ and match effects, conditional on X. In general, both are of indeterminate sign but they are zero when match effects are conditionally uncorrelated with person and firm effects.

Because we cannot sign the bias in the estimated covariance when match effects are omitted, we cannot sign the bias in the sample correlation either. A sufficient condition for the correlation to be biased toward zero is that match effects are uncorrelated with one another, person effects, and firm effects; and that the weighted conditional variance of employment duration exceeds the conditional variance of match effects.

A.1 Simulations

Some simulations illustrate how omitted match effects bias the estimated correlation between θ and ψ. In each simulation, $N = 250$ workers are employed over $T = 10$ periods at one of $J = 25$ firms. The imbalance between N and J reflects the structure of the LEHD data. Each worker is endowed with a person effect $\theta_i \sim N(0, \sigma_\theta^2)$ and each firm is endowed with a firm effect $\psi_j \sim N\left(0, \sigma_\psi^2\right)$. In the first period, each person is randomly matched to a firm with uniform probabilities. At the end of each period, matches dissolve with probability p. When a match dissolves, the worker is randomly matched to a new firm, again with uniform probabilities. Because matching probabilities do not depend on θ_i and ψ_j, the population correlation between θ_i and ψ_j is zero in new matches. Each match is endowed with a match effect $\phi_{ij} \sim N\left(0, \sigma_\phi^2\right)$. In each period, wages are $y_{it} = \mu + \theta_i + \psi_j + \phi_{ij} + \varepsilon_{ijt}$ where $\mu = 10$ and $\varepsilon_{ijt} \sim N(0, \sigma_\varepsilon^2)$. We set $\sigma_\theta^2, \sigma_\psi^2$, and σ_ε^2 equal to the hybrid mixed model estimates in Table 3. We estimate the person and firm effects model and a mixed model with random person, firm, and match effects on the realized employment and wage histories in each simulation.

Appendix Table 2 reports variance and covariance parameters, averaged over 1000 simulations, under four different scenarios. In Panels A and B, there are no match effects ($\sigma_\phi^2 = 0$). In Panels C and D, $\sigma_\phi^2 = 0.079$ (our hybrid mixed model estimate in Table 3). In Panels A and C, matches have equal probability of dissolution, $p = 0.2$. In Panels B and D, separation probabilities are decreasing in the value of person, firm, and match effects, $p = 0.2 - \theta_i - \psi_j - \phi_{ij}$. This reflects our empirical finding that average person, firm, and match effects increase with completed job duration.

There is no apparent bias in the estimated variance and covariance parameters in the match effects model. There is also no substantive bias in the person and firm effects model in the absence of match effects. In particular, downward bias in the estimated correlation between person and firm effects, as predicted by Andrews et al. (2008), is small.

Note that the duration-weighted correlation between actual person and firm effects (where each wage record is given equal weight) is less than the unweighted correlation (where each match is given equal weight) when separation probabilities are heterogeneous. This is because some "mismatches" (i.e., where θ_i and ψ_j have opposite signs) are less likely to dissolve when p is decreasing in person and firm effects than in the case where all matches end with equal probability.[37] In the absence of match effects, where $p = 0.2 - \theta_i - \psi_j$, we have $p < 0.2$ whenever $\max\left(\theta_i, \psi_j\right) > -\min\left(\theta_i, \psi_j\right)$. This includes mild mismatches where θ_i and ψ_j have opposite signs but their sum is positive. Because these mild mismatches dissolve with below-average probability, they tend to survive longer than average and hence receive greater weight in the duration-weighted correlation between θ_i and ψ_j than in the unweighted correlation. This may partly explain why most studies find a low correlation between θ_i and ψ_j (prior studies report duration-weighted correlations, as we do in throughout), and is independent of any estimation bias. The persistence of mismatches is consistent with search and matching models (e.g., Shimer and Smith (2000), Shimer (2005), and Woodcock (2010)), provided the mismatch is not too great and match productivity does not exhibit too much complementarity.

The lower panel of the table gives simulation results when wages depend on match effects. The bias from omitted match effects loads predominantly onto $\hat{\theta}_i$. We see this indirectly in the sample variance of $\hat{\theta}_i$, which is substantially inflated. It is also apparent in the mean biases. It is due to the imbalance between N and J: firms enter into many more matches than workers, so omitted match effects tend to average to zero for firms, but not workers.

Match effects have no effect on the prevalence of mismatch: true correlations between

[37] This result is quite generic. When p is a decreasing convex function of θ and ψ, Jensen's inequality implies that for all $\theta > 0$ there is a $\psi < 0$ such that $p < E[p]$. Even when p is strictly concave, for any $\theta > 0$ there is a $\psi \neq \theta$ such that $p < E[p]$. This is a weaker form of "mismatch."

θ_i and ψ_j are about the same in the lower and upper panels. However, omitted match effects bias the estimated correlation between θ_i and ψ_j downward. The bias is small when separation probabilities are homogeneous, but large when separation probabilities are not. The downward bias is disproportionately large when $\phi_{ij} < 0$. Here, average bias in θ_i is greatest when both θ_i and ψ_j are positive. Because p is decreasing in θ_i and ψ_j, these matches are highly durable. Consequently, the bias loads almost entirely onto $\hat{\theta}_i$ because these workers enter few (if any) other matches. This imparts substantial downward bias in $\hat{\theta}_i$ but little bias in $\hat{\psi}_j$, and makes the match appear less assortative. Because these spells endure, they receive an above-average weight in the duration-weighted correlation, inducing a disproportionate downward bias. The case is similar when $\phi_{ij} > 0$, $\theta_i > 0$, and $\psi_j < 0$.

B Appendix: The Dense Sampling Procedure

We draw dense samples in disjoint pairs, denoted by $s \in \{1, 2\}$. Our main results are based on one sample; the other sample in the pair is used in robustness checks below. Sample characteristics depend on three parameters: n determines the degree of connectedness; $p \in [0, 1]$ determines firms' probability of being sampled; and $m \in [0, 1]$ determines the relative size of the two disjoint samples. Sample $s = 1$ is a $100mpn$ percent random sample of workers, and $s = 2$ is a $100(1 - m)\,pn$ percent sample. In our application, $n = 5$, $p = 1/25$, and $m = 1/2$.

Let $\mathbf{D}$ denote the LEHD universe data; t is the reference period, and N_j is firm j's employment in t. In our application, $t = 1997$. We make two assumptions:

Assumption 1 *Each worker i is employed at only one firm $j = \mathcal{J}(i, t)$ in t.*

Assumption 2 *All firms have employment $N_j \geq n$ in t.*

To satisfy Assumption 1, we restrict $\mathbf{D}$ to the universe of dominant jobs. If a worker held more than one job some year, the one where her total earnings were greatest is her dominant job that year. To satisfy Assumption 2, we restrict $\mathbf{D}$ to the universe of individuals employed at firms with $N_j \geq 5$ employees in 1997. Let $\mathcal{J}_{\mathbf{D}}$ denote this set of firms, and $\mathcal{J}_s$ the set of firms in sample s. Similarly, $\mathcal{I}_{\mathbf{D}}$ is the set of individuals employed at firms in $\mathcal{J}_{\mathbf{D}}$ in 1997, and $\mathcal{I}_s$ is the set of workers in sample s.

Sampling proceeds in two stages. First, we sample firms from $\mathcal{J}_{\mathbf{D}}$ with probability proportional to N_j. Firm j is sampled with probability $\pi(j) = \min\{1, pN_j\}$. Conditional on being sampled, we sample j into $\mathcal{J}_1$ with probability m, and into $\mathcal{J}_2$ with probability $1-m$. Hence the probability that $j \in \mathcal{J}_1$ is $\pi_1(j) = m \min\{1, pN_j\}$, and the probability that $j \in \mathcal{J}_2$ is $\pi_2(j) = (1-m)\min\{1, pN_j\}$. Note $\mathcal{J}_1$ and $\mathcal{J}_2$ are disjoint.

Second, we sample employees of firms in $\mathcal{J}_1$ and $\mathcal{J}_2$. Let $n_j = \max\{n, npN_j\}$. If $j \in \mathcal{J}_s$, we sample n_j individuals such that $\mathcal{J}(i,t) = j$ into $\mathcal{I}_s$, each with equal probability. Hence the probability that individual i is sampled into $\mathcal{I}_s$, given that $\mathcal{J}(i,t) = j$ and $j \in \mathcal{J}_s$, is $\pi_s(i|j) = n_j/N_j = np \max\left\{p^{-1}N_j^{-1}, 1\right\}$. Because $\mathcal{J}_1$ and $\mathcal{J}_2$ are disjoint, Assumption 1 implies that $\mathcal{I}_1$ and $\mathcal{I}_2$ are disjoint also.

By Bayes' rule, the probability that i is sampled into $\mathcal{I}_s$ is $\pi_s(i) = \pi_s(i|j)\,\pi_s(j)/\pi_s(j|i)$, where $\pi_s(j|i)$ is the probability that $j \in \mathcal{J}_s$, given that $i \in \mathcal{I}_s$ and $j = \mathcal{J}(i,t)$. Assumption 1 implies $\pi_s(j|i) = 1$, so that $\pi_s(i) = \pi_s(i|j)\,\pi_s(j)$. It follows that $\pi_1(i) = mnp$ and $\pi_2(i) = (1-m)\,np$. Hence all individuals in $\mathcal{I}_{\mathbf{D}}$ have an equal probability of being sampled into each $\mathcal{I}_s$. Furthermore, each sampled worker is connected to at least $n_j \geq n$ others: the other employees sampled from their common employer in t. To complete the sampling procedure, subsample s consist of the complete work histories (at dominant jobs) of all $i \in \mathcal{I}_s$.

We assess the robustness of our results to sampling in several ways. First, we take five additional dense subsamples and estimate the three specifications reported in Tables 2-4 on each. The additional samples are taken in disjoint pairs: dense sample 2 is disjoint from the main sample (on which Tables 2-5 are based); dense samples 3 and 4 are disjoint from one another, as are dense samples 5 and 6. Disjoint sampling ensures that very large firms are not included in every sample. The first three columns of Appendix Table 3 present the mean and standard deviation of parameter estimates in the additional 5 samples. Estimates are very similar to those presented in the main text, and there is little variation across samples.

There is therefore no evidence that our results are specific to a particular sample.

Second, using estimates of the hybrid mixed model on all six samples, we estimate the between-sample correlation of BLUPs for those individuals, firms, and matches that appear in multiple samples. The pairwise correlation between BLUPs of a given effect in different samples is very large: 0.91 for person effects, 0.72 for firm effects, and 0.87 for match effects.

Finally, since least squares is computationally feasible on very large samples, we estimate the person and firm effects model and the orthogonal match effects model on all individuals employed in 1997 ($\mathcal{I}_{\mathbf{D}}$). Estimates are in the final two columns of Appendix Table 3. They are very similar to estimates based on the dense samples. We slightly overestimate the variance of θ_i and ψ_j in the dense samples, and slightly underestimate the variance of ϕ_{ij}, but discrepancies are all in the second or third decimal place. The estimated correlation between θ_i and ψ_j is near zero throughout, though it is slightly positive in the full sample and slightly negative in the dense subsamples.

References

Abowd, J. M., R. H. Creecy, and F. Kramarz (2002). Computing person and firm effects using linked longitudinal employer-employee data. Mimeo.

Abowd, J. M., F. Kramarz, P. Lengermann, and S. Perez-Duarte (2004). Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States. Mimeo.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–334.

Abowd, J. M., B. E. Stephens, L. Vilhuber, F. Andersson, K. L. McKinney, M. Roemer, and S. D. Woodcock (2009). The LEHD infrastructure files and the creation of the Quarterly Workforce Indicators. In T. Dunne, J. B. Jensen, and M. J. Roberts (Eds.), *Producer Dynamics: New Evidence from Micro Data*, pp. 149–236. Cambridge, MA: National Bureau of Economic Research.

Abraham, K. G. and H. S. Farber (1987). Job duration, seniority, and earnings. *The American Economic Review 77*(3), 278–297.

Altonji, J. G. and R. A. Shakotko (1987). Do wages rise with job seniority? *Review of Economic Studies 54*(3), 437–459.

Anderson, T. W. and C. Hsiao (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association 76*, 598–606.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008). High wage workers and low wage firms: Negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A 171*(3), 673–697.

Bartel, A. P. and G. J. Borjas (1981). Wage growth and job turnover: An empirical analysis. In S. Rosen (Ed.), *Studies in Labor Markets*, pp. 65–90. Chicago: National Bureau of Economic Research.

Barth, E. and H. Dale-Olsen (2003). Assortative matching in the labour market? Stylised facts about workers and plants. Institute for Social Research Working Paper.

Becker, G. (1973). A theory of marriage: Part I. *Journal of Political Economy 81*, 813–846.

Bureau of Labor Statistics (1997). *BLS Handbook of Methods*. U.S. Department of Labor.

Cornelißen, T. and O. Hübler (2007). Unobserved individual and firm heterogeneity in wage and tenure functions: Evidence from German linked employer-employee data. IZA Discussion Paper 2741.

Goldberger, A. S. (1962). Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association 57*, 369–375.

Goodreau, S. M. (2007). Advances in exponential random graph (p*) models applied to a large social network. *Social Networks 29*(2), 231–248.

Gruetter, M. and R. Lalive (2004). The importance of firms in wage determination. IZA Discussion Paper No. 1367.

Hausman, J. A. and W. E. Taylor (1981). Panel data and unobservable individual effects. *Econometrica 49*(6), 1377–1398.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Henderson, C., O. Kempthorne, S. Searle, and C. V. Krosigk (1959). The estimation of environmental and genetic trends from records subject to culling. *Biometrics 15*(2), 192–218.

Jiang, J. (1996). REML estimation: Asymptotic behavior and related topics. *The Annals of Statistics 24*(1), 255–286.

Jiang, J. (1998). Asymptotic properties of the empirical BLUP and BLUE in mixed linear models. *Statistica Sinica 8*, 861–885.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy 87*(5), 972–990.

Lopes de Melo, R. (2009). Sorting in the labor market: Theory and measurement. Mimeo.

Maré, D. C. and D. R. Hyslop (2006). Worker-firm heterogeneity and matching: An analysis using worker and firm fixed effects estimated from LEED. Statistics New Zealand LEED Report.

Nagypal, E. (2007). Learning by doing vs. learning about match quality: Can we tell them apart? *Review of Economic Studies 74*, 537–566.

Pissarides, C. (1994). Search unemployment with on-the-job search. *The Review of Economic Studies 61*(3), 457–475.

Robinson, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statistical Science 6*(1), 15–32.

Searle, S. R. (1987). *Linear Models for Unbalanced Data*. New York: John Wiley and Sons.

Shimer, R. (2005). The assigment of workers to jobs in an economy with coordination frictions. *Journal of Political Economy 113*(5), 996–1025.

Shimer, R. and L. Smith (2000). Assortative matching and search. *Econometrica 68*(2), 343–369.

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology 30*, 361–395.

Stram, D. O. and J. W. Lee (1994). Variance component testing in the longitudinal mixed effects model. *Biometrics 50*, 1171–1177.

Topel, R. H. and M. P. Ward (1992). Job mobility and the careers of young men. *Quarterly Journal of Economics 107*(2), 439–479.

Woodcock, S. D. (2010). Heterogeneity and learning in labor markets. *The B.E. Journal of Economic Analysis and Policy (Advances) 10*(1), Article 85.

FIGURE 1
Estimated Returns to Experience

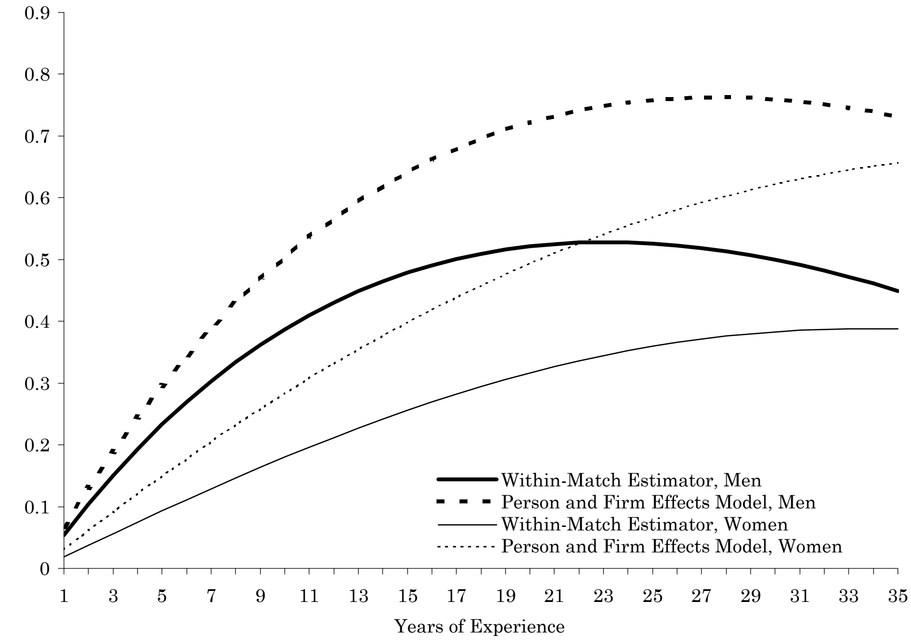
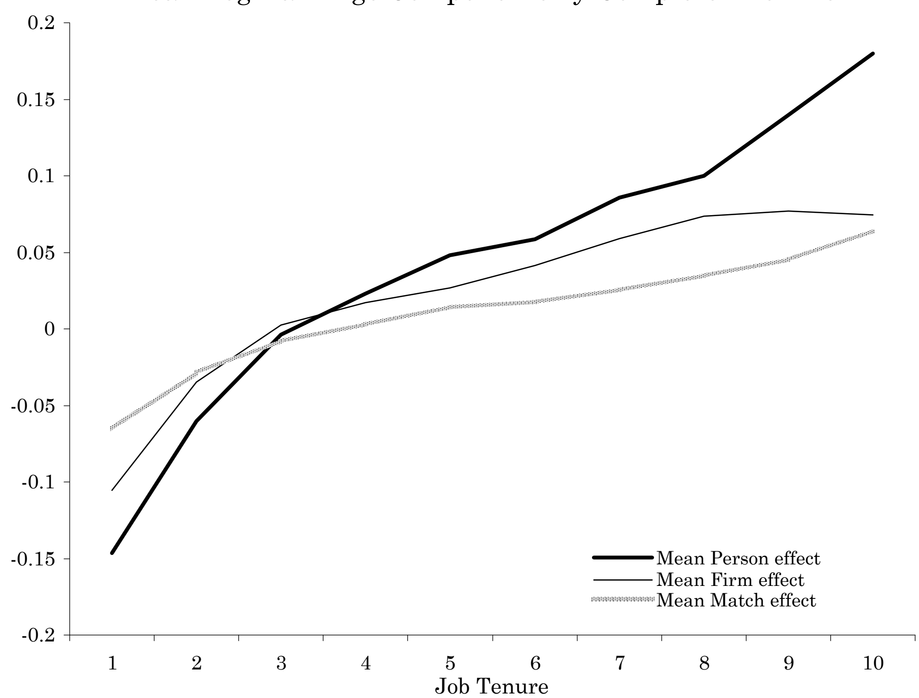


FIGURE 2
Mean Log Earnings Components by Completed Tenure



Notes: Estimates in Figure 2 are based on the hybrid mixed model and exclude censored employment spells.

TABLE 1
SUMMARY STATISTICS (Sample Proportions Unless Otherwise Stated)

	FULL SAMPLE		ALL INDIVIDUALS EMPLOYED IN 1997		TEN PERCENT DENSE SUBSAMPLE	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Demographic Characteristics						
Male	.56	.50	.58	.49	.57	.50
Age (Years)	40.6	10.2	40.3	9.6	40.3	9.6
Men						
Nonwhite	.21	.57	.20	.55	.20	.56
Race Missing	.04	.25	.03	.24	.03	.24
Less Than High School	.12	.45	.11	.43	.11	.43
High School	.30	.67	.30	.65	.29	.66
Some College	.23	.60	.23	.59	.23	.59
Associate or Bachelor's Degree	.25	.62	.25	.61	.25	.62
Graduate or Professional Degree	.10	.42	.11	.42	.11	.42
Women						
Nonwhite	.24	.69	.24	.71	.25	.72
Race Missing	.02	.22	.02	.22	.02	.22
Less Than High School	.09	.45	.09	.45	.09	.44
High School	.31	.78	.30	.79	.30	.78
Some College	.25	.71	.25	.73	.25	.72
Associate or Bachelor's Degree	.26	.72	.27	.75	.27	.75
Graduate or Professional Degree	.08	.42	.09	.44	.09	.44
Work History Characteristics						
Real Annualized Earnings (1990 Dollars)	41,107	38,849	43,183	39,324	43,528	38,782
Men						
Labor Market Experience (Years)	11.8	13.1	11.9	12.7	11.8	12.7
Worked 0 Full Quarters in Calendar Year	.08	.36	.06	.32	.06	.32
Worked 1 Full Quarter in Calendar Year	.15	.49	.12	.44	.12	.44
Worked 2 Full Quarters in Calendar Year	.13	.47	.12	.44	.12	.44
Worked 3 Full Quarters in Calendar Year	.14	.48	.13	.46	.14	.47
Worked 4 Full Quarters in Calendar Year	.50	.80	.56	.81	.57	.00
Women						
Labor Market Experience (Years)	9.5	13.0	9.0	12.5	9.2	12.6
Worked 0 Full Quarters in Calendar Year	.07	.39	.06	.36	.05	.35
Worked 1 Full Quarter in Calendar Year	.14	.54	.11	.50	.11	.50
Worked 2 Full Quarters in Calendar Year	.13	.53	.12	.51	.11	.50
Worked 3 Full Quarters in Calendar Year	.14	.55	.13	.54	.13	.54
Worked 4 Full Quarters in Calendar Year	.52	.96	.58	1.02	.59	1.01
Year						
1990	.09	.29	.07	.26	.07	.26
1991	.09	.29	.08	.27	.08	.27
1992	.09	.29	.08	.27	.08	.28
1993	.10	.29	.09	.28	.09	.28
1994	.10	.30	.10	.29	.10	.29
1995	.10	.30	.10	.31	.10	.31
1996	.10	.31	.11	.32	.11	.32
1997	.11	.31	.14	.35	.14	.34
1998	.11	.31	.12	.32	.12	.32
1999	.11	.31	.11	.31	.11	.31
Average firm size*	23.8		27.0		98.7	
Average firm size, sample-weighted**	2376		2486		2946	
Number of Observations (N*)	49,291,205		37,688,492		3,652,544	
Number of Workers (N)	9,272,529		5,235,887		503,179	
Number of Firms (J)	573,307		476,745		121,227	
Number of Worker-Firm Matches (M)	15,309,134		9,889,502		947,883	
Number of Connected Groups	84,748		46,829		1,460	

* Annual average of the firm's monthly employment reports, each firm-year is given equal weight.

** Annual average of the firm's monthly employment reports, each wage record is given equal weight.

TABLE 2
ESTIMATED RETURNS TO OBSERVABLE CHARACTERISTICS

	(1) PERSON AND FIRM EFFECTS		(2) ORTHOGONAL MATCH EFFECTS		(3) HYBRID MIXED MODEL	
	Estimate	SE	Estimate	SE	Estimate	SE
Time-Varying Characteristics (β)						
Male x Experience	.069	.000	.056	.001	.056	.001
Male x Experience^2 / 100	-.215	.002	-.194	.005	-.194	.005
Male x Experience^3 / 1000	.028	.001	.027	.001	.027	.001
Male x Experience^4 / 10000	-.002	.000	-.002	.000	-.002	.000
Male x Worked 0 Full Quarters	.039	.001	.053	.001	.053	.001
Male x Worked 1 Full Quarters	-.001	.001	.008	.001	.008	.001
Male x Worked 2 Full Quarters	-.011	.001	-.007	.001	-.007	.001
Male x Worked 3 Full Quarters	-.015	.001	-.014	.001	-.014	.001
Female x Experience	.031	.000	.019	.001	.019	.001
Female x Experience^2 / 100	-.027	.002	.000	.005	.000	.005
Female x Experience^3 / 1000	-.005	.001	-.009	.001	-.009	.001
Female x Experience^4 / 10000	.001	.000	.001	.000	.001	.000
Female x Worked 0 Full Quarters	.011	.001	.027	.001	.027	.001
Female x Worked 1 Full Quarters	-.004	.001	.006	.001	.006	.001
Female x Worked 2 Full Quarters	-.015	.001	-.012	.001	-.012	.001
Female x Worked 3 Full Quarters	-.022	.001	-.020	.001	-.020	.001
Time-Invariant Characteristics (η)						
Male x High School	.079	.004	.050	.001	.051	.004
Male x Some College	.175	.004	.140	.001	.131	.004
Male x Associate or Bachelor's Degree	.333	.004	.285	.001	.265	.004
Male x Graduate or Professional Degree	.518	.005	.468	.002	.437	.005
Male x Nonwhite	-.340	.003	-.357	.001	-.360	.003
Male x Race Missing	.017	.006	.012	.002	-.062	.006
Male x First Period Potential Experience <0	-.085	.008	-.234	.003	-.186	.006
Female	-.261	.006	-.198	.002	-.236	.005
Female x High School	.180	.004	.074	.002	.073	.004
Female x Some College	.287	.005	.172	.002	.155	.004
Female x Bachelor or Associate's Degree	.468	.005	.336	.002	.299	.004
Female x Graduate or Professional Degree	.646	.006	.504	.002	.463	.005
Female x Nonwhite	-.120	.003	-.132	.001	-.130	.003
Female x Race Missing	.006	.010	.001	.003	-.038	.008
Female x First Period Potential Experience <0	.080	.009	-.064	.003	-.025	.007
Intercept	9.86	.005	9.86	.003	9.86	.003
Year Effects	YES		YES		YES	

Notes: Estimates are based on the ten percent dense subsample.

TABLE 3
VARIANCE OF ESTIMATED COMPONENTS OF LOG EARNINGS

	(1) PERSON AND FIRM EFFECTS	(2) ORTHOGONAL MATCH EFFECTS	(3) HYBRID MIXED MODEL
Variance of Log Real Annualized Earnings (y)	.410	.410	.410
Variance of Returns to Time-Varying Characteristics (Xβ)	.030	.017	.017
Variance of Pure Person Effect (θ)	.291	.290	.198
Returns to Time-Invariant Characteristics (Uη)	.044	.040	.039
Unobserved Heterogeneity (α)	.247	.250	.159
Variance of Firm Effect (ψ)	.080	.081	.102
Variance of Match Effect (ϕ)		.016	.079
Error Variance (ε)	.055	.036	.036
Proportional Decomposition of Variance			
Returns to Time-Varying Characteristics (Xβ)	.068	.051	.051
Pure Person Effect (θ)	.658	.673	.482
Returns to Time-Invariant Characteristics (Uη)	.121	.120	.124
Unobserved Heterogeneity (α)	.537	.553	.358
Firm Effect (ψ)	.164	.164	.222
Match Effect (ϕ)		.039	.157
Residual (e)	.111	.072	.087
TOTAL	1.000	1.000	1.000
Corr(θ,ψ)	-0.013	-0.008	0.185
Corr(Xβ+θ,ψ)	0.007	0.010	0.185
H_0: No Match Effects (p-value)		<0.00001	<0.00001
R^2	.889	.928	.933
Model Degrees of Freedom	3,029,572	2,706,095	2,706,095

Notes: Estimates are based on the ten percent dense subsample. Rows labeled α, ψ, ϕ, ε in column 3 are REML estimates. All other estimates are sample variances and correlations.

TABLE 4
EMPLOYMENT MOBILITY AND THE CHANGE IN LOG EARNINGS

	(1) PERSON AND FIRM EFFECTS	(2) ORTHOGONAL MATCH EFFECTS	(3) HYBRID MIXED MODEL
Mean Annual Change in Real Log Earnings[†]	.032	.032	.032
Mean Change in Log Earnings, Job Stayers[‡]	.029	.029	.029
A. All Employment Transitions			
Mean Change in Time-Varying Characteristics (Xβ)	.035	.038	.038
Mean Change in Firm Effect (ψ)	.014	.012	.016
Mean Change in Match Effect (ϕ)		.003	-.003
Mean Change in Residual (e)	.000	-.003	-.002
TOTAL	.049	.049	.049
N	461,397	461,397	461,397
B. Job-to-Job Employment Transitions			
Mean Change in Time-Varying Characteristics (Xβ)	.033	.034	.034
Mean Change in Firm Effect (ψ)	.032	.031	.032
Mean Change in Match Effect (ϕ)		.022	.015
Mean Change in Residual (e)	.017	-.005	.001*
TOTAL	.082	.082	.082
N	213,763	213,763	213,763
C. Non Job-to-Job Employment Transitions			
Mean Change in Time-Varying Characteristics (Xβ)	.036	.041	.041
Mean Change in Firm Effect (ψ)	-.002	-.005	.002
Mean Change in Match Effect (ϕ)		-.014	-.018
Mean Change in Residual (e)	-.014	-.002	-.005
TOTAL	.020	.020	.020
N	247,634	247,634	247,634

[†] Average change in log earnings between year t and t+1, all observations.

[‡] Average change in log earnings between year t and t+1, all observations where the employer is the same in years t and t+1.

Notes: Estimates are based on the ten percent dense subsample of individuals employed in 1997. Asterisk (*) indicates estimate is **not** statistically significant at the 5% level.

TABLE 5

ROBUSTNESS AND ALTERNATE EXPLANATIONS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		HYBRID MIXED MODEL					
	RANDOM EFFECTS MODEL	WITH ARMA(1,1) ERRORS	EXCLUDING LEFT-CENSORED SPELLS	SELECTION CORRECTED	WITH TENURE CONTROLS	TENURE CONTROLS, SELECTION CORRECTED	LAG EARNINGS CONTROL
Variance of y	.410	.410	.432	.432	.432	.432	.391
Variance of Xβ	.027	.017	.022	.026	.027	.026	.033
Variance of θ	.189	.105	.200	.208	.206	.205	.121
Variance of Uη	.036	.040	.031	.030	.030	.029	.023
Variance of α	.153	.065	.170	.177	.176	.175	.098
Variance of ψ	.104	.100	.107	.106	.106	.105	.084
Variance of ϕ	.079	.074	.080	.080	.079	.078	.063
Variance of ε	.036	.146	.040	.040	.040	.040	.031
AR(1) coefficient		.910					
MA(1) coefficient		.222					
Corr(θ,ψ)	.185	.185	.215	.210	.209	.211	.166
Corr(Xβ+θ,ψ)	.186	.176	.219	.218	.222	.223	.242
Coefficients on Additional Controls*							
Inverse Mills Ratio (λ)				-.013		-.028	
Tenure = 2 years					.006	.008	
Tenure = 3 years					.032	.033	
Tenure = 4 years					.048	.059	
Tenure = 5 years					.056	.065	
Tenure = 6 years					.060	.068	
Tenure = 7 years					.058	.072	
Tenure = 8 years					.054	.067	
Tenure = 9 years					.050	.063	
Tenure = 10 years					.044	.057	
y_{ijt-1}							.163
Hausman Test (p-value)	<0.00001						
H_0: No Match Effects (p-val)	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001
R^2	.933	.814	.934	.931	.934	.934	.940
Number of Observations	3,652,544	3,652,544	2,335,623	2,335,623	2,335,623	2,335,623	1,029,204

* All reported coefficients are statistically significant at the 1% level.

Notes: Estimates are based on the ten percent dense subsample of individuals employed in 1997. Columns 3-7 exclude left-censored spells. Values are based on sample variances and correlations of estimated effects, except reported variances of α, ψ, ϕ, ε are REML estimates of variance components; AR and MA coefficients are also REML estimates. Column 7 is based on the IV estimator described in the text and excludes the first two observations on each job.

APPENDIX TABLE 1
VARIABLE DEFINITIONS

VARIABLE	DEFINITION
Log Real Annualized Earnings	Natural logarithm of real annualized earnings, in 1990 dollars. Annualized earnings equal the annual average of reported UI earnings in full quarters* of employment, multiplied by four. Deflated using the CPI.
Labor Market Experience	Quartic in experience and dummy variable for negative potential experience in first quarter of employment, all interacted with sex. Potential experience is start-of-quarter age minus years of education minus six in the first quarter an individual appears in the sample. Experience increases by 0.25 in each subsequent quarter of employment.
Race	Dummy variables for Nonwhite and Race Missing, interacted with sex.
Education	Dummy variables for educational attainment, interacted with sex. Categories are Less than High School (omitted), High School, Some College or Vocational Training, College Degee, and Graduate or Professional Degree.
Year Effects	Dummy variables for year. Omitted category is 1990.
Sex	Dummy variable for female.
Full Quarters Worked	Dummy variables for the number of full quarters* worked during the calendar year, interacted with sex. Omitted category is four full quarters.

* An individual is defined to have worked a full quarter at firm j in quarter q if she was employed at firm j in quarter q-1 and quarter q+1.

APPENDIX TABLE 2
Simulation Results

NO MATCH EFFECTS (σ^2_ϕ=0)

	A: Homogeneous Separations			B: Heterogeneous Separations		
	Realized Values	Estimates from Person and Firm Effects Model	Estimates from Match Effects Model	Realized Values	Estimates from Person and Firm Effects Model	Estimates from Match Effects Model
Var(θ)	.198	.199	.198	.198	.199	.198
Var(ψ)	.102	.096	.103	.102	.094	.106
Var(ϕ)			.001			.002
Unweighted Corr(θ,ψ)	.002	.002	.003	.002	-.002	-.001
Duration-weighted Corr(θ,ψ)	.001	.001	.002	-.080	-.086	-.084

MATCH EFFECTS (σ^2_ϕ=0.079)

	C: Homogeneous Separations			D: Heterogeneous Separations		
	Realized Values	Estimates from Person and Firm Effects Model	Estimates from Match Effects Model	Realized Values	Estimates from Person and Firm Effects Model	Estimates from Match Effects Model
Var(θ)	.198	.243	.198	.198	.226	.197
Var(ψ)	.102	.100	.102	.102	.101	.105
Var(ϕ)	.079		.082	.079		.085
Unweighted Corr(θ,ψ)	.002	-.011	.012	-.002	-.025	.002
Duration-weighted Corr(θ,ψ)	.002	-.015	.014	-.070	-.120	-.072
Duration-weighted Corr(θ,ψ), ϕ>0	.002	-.022	.009	-.119	-.131	-.099
Duration-weighted Corr(θ,ψ), ϕ<0	.003	-.020	.011	-.011	-.101	-.028

Mean bias in estimated effects when match effects are omitted

ϕ	θ	ψ	Sample Proportion	Mean bias in $\hat{\theta}$	Mean bias in $\hat{\psi}$	Sample Proportion	Mean bias in $\hat{\theta}$	Mean bias in $\hat{\psi}$
Neg	Neg	Neg	.125	-.121	-.007	.088	-.067	-.018
Neg	Neg	Pos	.125	-.121	-.008	.102	-.116	-.001
Neg	Pos	Neg	.124	-.120	-.007	.104	-.146	-.015
Neg	Pos	Pos	.125	-.119	-.009	.114	-.209	.001
Pos	Neg	Neg	.125	.120	.008	.140	.081	.000
Pos	Neg	Pos	.126	.120	.007	.170	.090	.013
Pos	Pos	Neg	.124	.119	.008	.159	.114	-.003
Pos	Pos	Pos	.125	.121	.007	.123	.098	.011

APPENDIX TABLE 3
ROBUSTNESS OF MAIN RESULTS TO SAMPLING

	(1)		(2)		(3)		(4)	(5)
	RESULTS IN 5 ADDITIONAL DENSE SAMPLES						ALL INDIVIDUALS EMPLOYED IN 1997	
	PERSON AND FIRM EFFECTS		ORTHOGONAL MATCH EFFECTS		HYBRID MIXED MODEL		PERSON AND FIRM EFFECTS	ORTHOGONAL MATCH EFFECTS
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev		
Variance of y	.408	.002	.408	.002	.408	.002	.422	.422
Variance of $X\beta$	.029	.003	.017	.002	.017	.002	.031	.017
Variance of θ	.286	.003	.287	.002	.196	.002	.274	.273
Variance of $U\eta$	.041	.001	.039	.001	.038	.002	.043	.041
Variance of α	.245	.002	.249	.002	.157	.002	.232	.233
Variance of ψ	.082	.002	.083	.002	.104	.001	.065	.066
Variance of ϕ			.016	.000	.079	.001		.022
Variance of ε	.056	.001	.037	.001	.037	.001	.066	.040
Corr(θ,ψ)	-.011	.001	-.011	.001	.187	.005	.003	.005
Corr($X\beta$+θ,ψ)	.011	.003	.008	.003	.186	.004	.026	.023
H_0: No Match Effects (p-value)			<0.00001		<0.00001			<0.00001
R^2	.886	.002	.926	.002	.931	.002	.867	.919
Number of Observations (N^*)	3,514,046		3,514,046		3,514,046		37,688,492	37,688,492
Number of Workers (N)	483,973		483,973		483,973		5,235,887	5,235,887
Number of Firms (J)	120,279		120,279		120,279		476,745	476,745
Number of Matches (M)	914,852		914,852		914,852		9,889,502	9,889,502
Number of Connected Groups	1,481		1,481		1,481		46,829	46,829

Notes: Mean and standard deviation in Columns 1-3 is over five 10% dense samples. Estimates in rows labeled α, ψ, ϕ, ε in column 3 are REML estimates of variance components. All other estimates are sample variances and covariances of estimated effects.