Heterogeneity and Learning in Labor Markets[1]

Simon D. Woodcock[2]
Simon Fraser University
simon_woodcock@sfu.ca

September 2005

[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 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 Grant SES-9978093 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant R01∼AG018854-02, and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author 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, Program 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]I especially thank John M. Abowd, David Easley, George Jakubson, Robert Shimer, Dale Mortensen, James Heckman, Martha Stinson, and members of the LEHD program staff for helpful comments on earlier versions of this paper. I also thank seminar participants at the 2002 CEA meetings, the 2003 CESG meetings, the University of Chicago, Cornell University, the 2004 North American winter meeting of the Econometric Society, the Federal Reserve Bank of Atlanta, McGill University, the University of Montreal, the 2003 NBER Summer Institute, Queen's University, RAND, Simon Fraser University, 2003 SOLE meetings, Texas A&M University, the University of Washington, the University of Western Ontario, and the University of Wisconsin-Madison.

Abstract

This paper examines the role of agent heterogeneity and learning on wage dispersion and employment dynamics. In the first half of the paper, I present an equilibrium matching model where heterogeneous workers and firms learn about match quality and bargain over wages. The model generalizes Jovanovic (1979) to the case of heterogeneous workers and firms. Equilibrium wage dispersion arises due to productivity differences across workers, technological differences across firms, and heterogeneity in beliefs about match quality. Under a simple CRS technology, the equilibrium wage is additively separable in worker- and firm-specific components, and in the posterior mean of beliefs about match quality. This parallels the "person and firm effects" empirical specification of Abowd et. al. (1999, AKM) and others. It consequently provides a theoretical context for the AKM model, and a formal economic interpretation of their empirical person and firm effects. The model also yields an assortative matching result that predicts a negative correlation between estimated person and firm effects, which is consistent with most empirical evidence. Finally, the model makes novel predictions about the relationship between the person and firm effects and separation behavior, job duration, and firm size. In the second half of the paper, I test the model's empirical predictions. I estimate fixed and mixed effects specifications of the equilibrium wage function on the LEHD database. The mixed effect specifications generalize the earlier work of AKM and others. The learning component of the matching model implies a specific structure for the error covariance. I exploit this structure to test whether earnings residuals are consistent with Bayesian learning, and to estimate structural parameters of the matching model. I find considerable support for the matching model in these data.

1 Introduction

It is well known that otherwise identical workers often earn very different wages and have heterogeneous employment histories. Likewise, similar firms frequently pay dissimilar wages and exhibit great heterogeneity in turnover. However, despite the best efforts of many researchers, we have yet to fully explain this enormous variety of outcomes. A convincing explanation of why similar workers earn different wages, and how this is related to heterogeneity in job duration, unemployment, and the like, is central to our understanding of labor markets, and a necessary input to informed labor market policy.

Early work sought to explain wage differences across workers on the basis of variation in human capital and the non-pecuniary aspects of jobs. But observable characteristics of workers and firms usually only explain about 30 percent of wage variation. Attempts to explain the residual component of wage variation, often called *wage dispersion*, have proceeded along several dimensions. Search and matching models reveal that labor market frictions are one cause of equilibrium wage dispersion and unemployment. Learning models provide an explanation for wage dynamics and separation behavior, and show that heterogeneous beliefs about match productivity are another source of wage dispersion. And recent empirical work using linked employer-employee data shows that wage dispersion can be decomposed into a component attributable to unobserved characteristics of the worker – a "person effect" – and a component attributable to unobserved characteristics of the employer – a "firm effect." Each of these provides a partial explanation for the diversity of labor market outcomes. To date, however, they have remained distinct. This paper demonstrates that they are complementary. Together, they not only provide a comprehensive explanation for heterogeneity in labor market outcomes, they also provide important new insights.

The paper has both theoretical and empirical components, and makes contributions in both dimensions. In the first half of the paper, I present an equilibrium matching model where heterogeneous workers and firms learn about match quality and bargain over wages. The main theoretical innovation is to embed learning about match quality in a Mortensen-Pissarides style equilibrium matching model with heterogeneous agents. This delivers novel insights into the relationship between worker and firm heterogeneity, wages, and separation behavior.

One contribution of the matching model is to generalize the canonical Jovanovic (1979) model to the case of heterogeneous workers and firms. Traditional matching models provide an explanation for job duration and turnover: matches last as long as agents believe the match is good (highly productive). If they learn that match productivity is low, they prefer to separate. However, a fundamental limitation of models with homogeneous agents is that they cannot explain why some workers experience consistently longer job duration than others, and why some firms experience less turnover than others. In the model presented here, workers and firms vary in their marginal productivity. The productivity of a worker-firm match depends on worker and firm productivity, and on a match-specific interaction that I call match quality. Workers and firms learn the value of match quality slowly by observing production outcomes. Like traditional matching models, they terminate the employment relationship if they learn that match quality is poor, i.e., if beliefs about match quality fall below a reservation value. A key result is that the reservation value is decreasing in both the worker's and the firm's productivity. Consequently, more productive workers experience longer average job duration than less productive workers. Likewise, firms with more productive technologies experience less turnover than less productive firms. This is consistent with empirical evidence, and provides an explanation for heterogeneity in job duration and turnover.

Distinguishing between worker, firm, and match heterogeneity is an important departure from earlier research. It recognizes that workers are differently able, and hence some are more productive on average than others. Likewise, it recognizes that firms operate different production technologies, and consequently employee productivity varies across firms. It also recognizes that not all workers are equally suited to all production technologies. As a consequence, two workers that are equally able may be differently productive in a given firm, simply because one is well suited to the firm's production technology and the other is not. It is easy to construct real world examples of this phenomenon. For instance, two equally able academics may have different proclivities for teaching and research. One will thrive in a university that emphasizes research while the other's productivity suffers. The reverse will be true in a university that emphasizes teaching.

The very heterogeneous production environment considered here has important implications for equilibrium wages. When production is according to a simple CRS technology and wages are set by a Nash bargain, the equilibrium wage is additively-separable in a worker-specific component, a firm-specific component, and the mean of beliefs about match quality. The worker- and firm-specific components reflect worker and firm productivity, adjusted for bargaining strength and discounting. This result is important for several reasons. First, it provides a rich explanation for wage dispersion: equilibrium dispersion arises due to productivity differences between workers, technological differences between firms, and due to heterogeneity in beliefs about match quality. Second, the additively-separable structure parallels the empirical person and firm effect specification that Abowd et al. (1999, AKM, hereafter) and others have estimated on linked employer-employee data. This specification typically "explains" about 90 percent of observed wage variation. However, the AKM decomposition of wages into person and firm effects is purely statistical. It provides no formal economic interpretation of what the person and firm effects actually measure. Thus the matching model makes an important contribution to the empirical literature. It provides a theoretical context for the AKM specification, and consequently a formal economic interpretation of the person and firm effects: they reflect productivity differences across workers and firms, adjusted for bargaining strength and discounting.

The model also provides new insight into the sorting behavior of workers and firms. The optimal separation policy implies that worker-firm matches are negatively assortative. Consequently, the matching model predicts a negative correlation between estimated person and firm effects. This is in fact what most prior empirical studies have found, but it has been considered something of an empirical puzzle. The model provides an intuitive explanation for this finding. Because high-productivity workers have a high opportunity cost of unemployment, they are willing to match with low-productivity firms. Likewise, when highly-productive firms have an unfilled vacancy, they forego more output than low-productivity firms do. They are consequently willing to match with low-productivity workers to avoid an unfilled vacancy. The result is equilibrium mismatch.

In the second half of the paper, I estimate structural parameters of the matching model using linked employer-employee data, and test a variety of its predictions. The main econometric innovation is to estimate a mixed model of earnings that treats the person and firm effects as random. This specification generalizes the fixed effect estimator that AKM and others have used, and provides novel insight into several sources of bias in earlier estimates. The most general specification that I estimate allows a completely unrestricted within-match error covariance. This is in contrast to the matching model, where learning about match quality implies a structured error covariance. I exploit this result to test whether earnings residuals are consistent with Bayesian learning. We cannot reject this hypothesis with a high degree of confidence.

The empirical results show that worker heterogeneity contributes about twice as much to wage dispersion as firm heterogeneity and learning about match quality do. Learning is very rapid: about half of the uncertainty is resolved in the first year of an employment relationship. Consistent with the prediction of the matching model, I find that larger values of the person and firm effects are associated with longer average job duration.

The AKM estimator has been criticized for relying on an exogenous mobility assumption.[1] This requires that employment mobility depend only on observable characteristics and the person and firm effects, and be independent of wage errors. The matching model presented here makes an important contribution in this regard. It demonstrates that the additively-separable AKM structure can arise even in the presence of endogenous separations due to learning about match quality. However, these endogenous separations truncate the error distribution. Failing to correct for truncation may bias estimates of the person and firm effects. I therefore develop a simple correction based on the Heckman (1979) two-step estimator. I find the correction has little influence on parameter estimates, which implies the bias is small.

Barth and Dale-Olsen (2003) and Andrews et al. (2004b) have raised another criticism of the AKM estimator. They argue that the fixed effect estimator yields biased estimates of the correlation between person and firm effects. In particular, when the true correlation is positive, they demonstrate that the estimated correlation is biased downward. I find some corroborating empirical evidence. In particular, I find that fixed and mixed effect estimators that assume spherical errors show no economically significant correlation between estimated person and firm effects. However, the estimated correlation between person and firm effects increases dramatically when the specification admits correlated errors within worker-firm matches. This suggests the bias is not a consequence of the fixed effects estimator *per se*, but rather a consequence of mis-specifying the error distribution.

The remainder of the paper is structured as follows. I begin by briefly reviewing the related literature. I then present the matching model in Section 2, and develop the empirical specification in Section 3. I give a brief description of the data in Section 4 (more detail is provided in a Data Appendix). I present the empirical results in Section 5 and conclude in Section 6.

1.1 Related Literature

The model presented here brings together three distinct literatures: that on search and matching with heterogeneous agents, the literature on learning in labor markets, and the

[1] Recent instances of this criticism include Gruetter and Lalive (2004) and Gruetter (2005).

emerging empirical literature that seeks to explain wage dispersion using linked employer-employee data. In this section, I briefly review relevant research in each of these areas.

1.1.1 Search and Matching with Heterogeneous Agents

In general, the search and matching literature has focused on economies with heterogeneous workers and jobs.[2] In the typical model, firms employ only a single worker. There is therefore no distinction between heterogeneity at the level of the firm and at the level of the worker-firm match. In contrast, I model an economy where firms employ many workers, and distinguish between productive heterogeneity at the firm, which affects all employees, and productive heterogeneity that is specific to a worker-firm match. Postel-Vinay and Robin (2002) also consider an environment where firms employ many workers, though their workers are equally productive in every firm.

Recently, interest has focused on conditions under which "good" workers sort into "good" firms in frictional economies. Shimer and Smith (2000) develop conditions under which assortative matching arises in the presence of search frictions. Shimer (2005) considers the optimal assignment of workers to jobs in the presence of coordination frictions. In either case, positive assortative matching arises if production exhibits sufficient complementarity between worker and firm types, or if high-productivity workers enjoy sufficient comparative advantage in high-productivity firms. Negative assortative matching arises in the reverse case. In the matching model presented here, the production technology is additively separable in worker and firm productivities. This implies low-productivity workers have comparative advantage in high-productivity firms. Worker-firm matches are therefore negatively assortative in equilibrium, and the model predicts a negative correlation between person and firm effects. This is consistent with empirical studies based on European data, though American data typically yield no significant correlation between person and firm effects.[3]

1.1.2 Learning in Labor Markets

The learning literature has focused primarily on wage and turnover dynamics. The seminal Jovanovic (1979) model considered the case where identical workers and firms learn about the quality of a match. Flinn (1986) and Moscarini (2003) generalize the canonical model to the case of heterogeneous workers. Harris and Holmstrom (1982) and Farber and Gibbons (1996) present models where workers and firms learn about a worker's unobservable ability, which is correlated with observable characteristics. Gibbons et al. (2005) extend this framework to the case of an economy with heterogeneous sectors (e.g., occupation or industry), and where workers exhibit comparative advantage in some sectors. Felli and Harris (1996) present a model where workers learn about their aptitude for firm-specific tasks. None of these earlier works consider the case of worker, firm, and match heterogeneity.

[2] Examples include Stern (1990), Sattinger (1995), Shimer and Smith (2000), and Shimer and Smith (2001). Albrecht and Vroman (2002), Gautier (2002), and Kohns (2000) develop models with exogenous heterogeneity on one side of the market, and endogenous heterogeneity on the other.

[3] Abowd et al. (2002) and Abowd et al. (2004) report a negative correlation in French data, and approximately zero correlation in American data. Gruetter and Lalive (2004) find a negative correlation in Austrian data, Barth and Dale-Olsen (2003) find a negative correlation in Norwegian data, and Andrews et al. (2004a) find no significant correlation in German data.

1.1.3 Estimating Person and Firm Effects

I estimate both fixed and mixed effect specifications of the equilibrium wage function. All prior studies are based on the fixed effect estimator of person and firm effects. These include AKM, Abowd et al. (2002, ACK hereafter), Abowd et al. (2003), Abowd et al. (2004), Abowd et al. (2005), Andersson et al. (2005), Barth and Dale-Olsen (2003), Gruetter and Lalive (2004), and Andrews et al. (2004a). The empirical specification considered by these authors, which I refer to generically as the AKM model, is

$$y_{ijt} = \mu + x'_{it}\beta + \theta_i + \psi_j + \varepsilon_{ijt} \tag{1}$$

where i indexes workers and j indexes firms, y_{ijt} is log earnings, μ is the grand mean, x_{it} is a vector of covariates, β is a parameter vector, θ_i is the person effect, ψ_j is the firm effect, and ε_{ijt} is statistical error. The original AKM study relies on approximate solutions for the estimated person and firm effects. Most subsequent studies are based on exact solutions using computational methods developed in ACK.

Economists have historically preferred fixed effect estimators to mixed effect estimators that treat unobserved heterogeneity as random. Statisticians, on the other hand, generally prefer the mixed model because mixed effect estimates of the unobserved heterogeneity (here, the estimated person and firm effects) have better sampling properties. Robinson (1991) gives an extended discussion along these lines. Furthermore, it is well known (among statisticians) that the fixed effect estimator is a special case of the mixed effect estimator. This is demonstrated in most statistical references on mixed model theory, e.g., Searle et al. (1992) and McCulloch and Searle (2001). Contrary to popular belief among economists, the mixed effect estimator does not necessarily assume that the design of the random effects is orthogonal to observable characteristics. Abowd and Kramarz (1999) and AKM discuss the orthogonal design assumption in detail.

2 The Matching Model

The economy is populated by a continuum of infinitely-lived workers of measure one. There is a continuum of firms of measure ϕ. All agents are risk neutral and share the common discount factor $0 < \beta < 1$. Time is discrete.

In each period, workers are endowed with a single indivisible unit of labor that they supply to production at home or at a firm. Workers vary in their marginal productivity when employed, denoted $a \in [a_0, a_1]$. Conceptually, a represents the worker's ability, motivation, and the like. I refer to a as worker quality or ability. Let

$$a \sim F_a \text{ iid across workers} \tag{2}$$

where F_a is a probability distribution with zero mean, known to all agents. Worker quality a is exogenous, known to the worker, and observed by the firm when the worker and firm meet. Note a is not a choice variable. Unemployed workers receive income $h \in \mathbb{R}$ from home production.[4] For simplicity, h includes all search costs, the value of leisure, and the like.

[4] We can let h vary across individuals without changing any key theoretical results. However it changes the interpretation of the person-specific component of wages (Section 2.2.3), the person-specific term in the reservation level of beliefs (Section 2.2.4), and complicates the comparative statics in Section 2.3.

Workers seek to maximize the expected present value of wages.

Firms employ many workers. They operate in a competitive output market and produce a homogeneous good whose price is normalized to one. Firms can only produce output when matched with workers. They seek to maximize the expected net revenues of a match: the expected value of output minus a wage payment to the worker.

Firms vary in their technology, which determines the marginal productivity of all their employees, denoted $b \in [b_0, b_1]$. Let

$$b \sim F_b \text{ iid across firms} \tag{3}$$

where F_b is a probability distribution known to all agents, and with zero mean. I call b firm quality. Firms know their own value of b, and it is observed by the worker when the worker and firm meet. Note b is exogenous. Firms are free to open vacancies.[5] They incur cost $\kappa(l)$ to hire l workers in the current period. Assume κ is continuous, increasing, and convex.

The marginal productivity of a type a worker when employed at a type b firm depends not only on worker and firm quality, but also on a worker- and firm-specific interaction that I call match quality and denote c. "Good" matches are more productive than "bad" ones, all else equal. Let

$$c \sim N\left(0, \sigma_c^2\right) \text{ iid across matches.} \tag{4}$$

The normality assumption follows Jovanovic (1979) and others. It yields a convenient closed form for beliefs about match quality.

Match quality c is a pure experience good. It is unobserved by either the worker or the firm. They learn its value slowly. When the worker and firm meet, they observe the noisy signal $x = c + z$ where

$$z \sim N\left(0, \sigma_z^2\right) \text{ iid across matches.} \tag{5}$$

The worker and firm's initial beliefs about c are based on a common prior and the signal x. They subsequently update their beliefs about c on the basis of output realizations. Prior beliefs and the updating process are discussed in Section 2.1. Note that information is incomplete, since c is unobserved, but is symmetric. That is, the worker and firm both know a and b, and observe the same signals about c. They therefore share common beliefs about c at every point in time.

Output is produced according to the constant returns to scale production function:

$$q_\tau = \mu + a + b + c + e_\tau \tag{6}$$

where τ indexes tenure (the duration of the match), μ is the grand mean of productivity (known to all agents), and e_τ is a match- and tenure-specific idiosyncratic shock. Let

$$e_\tau \sim N\left(0, \sigma_e^2\right) \text{ iid across matches and tenure.} \tag{7}$$

Note the Jovanovic (1979) production technology is a (continuous time) special case of (6) where workers and firms are homogeneous. Note also that there are no aggregate shocks to

[5] Firm-specific vacancy-opening costs do not affect any of the main results, but introduce an additional source of firm-level heterogeneity in hiring and employment growth.

productivity, and no human capital accumulation over the life cycle.[6] Since a, b, and μ are known, agents extract the noisy signal of match quality $c + e_\tau$ from production outcomes q_τ.

Following Flinn (1986), I assume that q_τ is bounded. This implies that the random variables c, z, and e_τ have bounded support. Thus the distributional assumptions (4), (5), and (7) are approximate.

Unemployed workers are matched to firms with open vacancies. Search is undirected. The total number of matches formed in a period is given by $m(u, v)$ where u is the number of unemployed workers in the economy, and v is the number of open vacancies. Both u and v are determined endogenously. Assume m is non-decreasing in both u and v. The probability that a randomly selected unemployed worker will be matched to a firm in the current period is $\pi \equiv m(u, v)/u$. Similarly, the probability that a randomly selected vacancy will be filled is $\lambda \equiv m(u, v)/v$. With a large number of workers and firms, all agents take u and v as given.

As discussed in greater detail below, a match between a worker and firm terminates endogenously when their point estimate of match quality falls below a threshold value. In addition, matches terminate with exogenous probability $\delta > 0$ in each period.

I restrict attention to steady states of the economy. The economy is in steady state when the end-of-period distribution of type a workers across employment at type b firms and across unemployment is constant. The various flow-balance equations that characterize the steady state are given in Appendix B. An implication of these is that the steady state level of unemployment u and the steady state number of vacancies v are constant. Hence so are the steady state values of λ and π.

Within-period timing is as follows:

1. With probability π, unemployed workers are randomly matched to a firm with an open vacancy. Upon meeting, agents observe a, b, and the signal x.

2. Workers and firms decide whether or not to continue the match. The decision is based on all current information about the match: a, b, and current beliefs about c. The current period wage w_τ is simultaneously determined by a Nash bargain.

3a. If agents decide to terminate the match, the worker enters unemployment and receives h. There are no firing costs.

3b. If agents decide to continue the match, the negotiated wage is paid to the worker and output q_τ is produced. Agents update their beliefs about c.

[6] Introducing a publicly-observable aggregate shock to productivity is relatively straightforward. Likewise, introducing a deterministic trend to individual productivity (i.e., an "experience effect") presents no serious complication provided that it is observable by the worker and firm. I omit these generalizations since they complicate the exposition considerably – both require additional notation and an additional index of calendar time. However, there is little loss of generality in their omission. The production function (6) can be considered output net of additive aggregate shocks and deterministically accumulated human capital. The same is true of the equlibrium wage w_τ in (30) and the net value of output $q_\tau - w_\tau$. That is, in the more general model, the equilibrium wage (see Proposition 1) remains additively separable in person- and firm- specific components and in the posterior mean of beliefs, and is linear and additively separable in the productivity shock and the experience effect.

4. Exogenous separations occur with probability δ.

5. Firms open new vacancies υ.

Assume that reputational considerations preclude agents from reneging on the agreed-upon wage payment.

2.1 Beliefs About Match Quality

Assume agents' prior beliefs about a, b, c, z, and e_τ are rational. That is, they are governed by equations (2), (3), (4), (5), and (7). Recall that the worker's type a and the firm's type b are observed by both parties when the worker and firm meet. Agents update their beliefs about match quality c using Bayes' rule when they acquire new information, i.e., upon observing the signal x and production outcomes q_τ.

After observing the signal x, worker and firm posterior beliefs about c are normally distributed with mean m_1 and variance s_1^2 where

$$m_1 = x\left(\frac{\sigma_c^2}{\sigma_c^2+\sigma_z^2}\right) \text{ and } s_1^2 = \frac{\sigma_c^2\sigma_z^2}{\sigma_c^2+\sigma_z^2}. \tag{8}$$

In each subsequent period that the match continues, the worker and firm extract the signal $c + e_\tau$ from observed output q_τ. Hence at the beginning of the τ^{th} period of the match (that is, after observing $\tau - 1$ production outcomes), worker and firm posterior beliefs about match quality are normally distributed with mean m_τ and variance s_τ^2, where

$$m_\tau = \left(\frac{m_{\tau-1}}{s_{\tau-1}^2} + \frac{c+e_{\tau-1}}{\sigma_e^2}\right) / \left(\frac{1}{s_{\tau-1}^2} + \frac{1}{\sigma_e^2}\right) \tag{9}$$

$$\frac{1}{s_\tau^2} = \frac{1}{s_{\tau-1}^2} + \frac{1}{\sigma_e^2}. \tag{10}$$

Equation (10) demonstrates that the evolution of s_τ^2 is deterministic and does not depend on the value of the signals received. It also shows that $s_\tau^2 > s_{\tau+1}^2$ for each $\tau > 0$. Equation (9) demonstrates that the posterior mean of beliefs m_τ is a precision-weighted average of the prior mean $m_{\tau-1}$ and the most recent signal $c + e_{\tau-1}$. Since the precision of signals $(1/\sigma_e^2)$ is constant but the precision of beliefs $(1/s_\tau^2)$ increases with tenure, each new signal is given successively smaller weight in the update. Asymptotically, beliefs converge to unit mass at true match quality. That is,

$$\lim_{\tau\to\infty} m_\tau = c \text{ and } \lim_{\tau\to\infty} s_\tau^2 = 0. \tag{11}$$

This is a standard result for Bayesian learning with rational priors.

In what follows, it will be of interest to describe the distribution of beliefs in the population. It is easy to show that

$$m_\tau \sim N(0, V_\tau) \tag{12}$$

$$V_\tau = s_\tau^2\sigma_c^2\left(\frac{1}{\sigma_z^2} + \frac{\tau-1}{\sigma_e^2}\right). \tag{13}$$

With a little algebra, one can also show $V_{\tau+1} > V_\tau$ for all $\tau > 0$. That is, the variance of the posterior mean of beliefs about match quality increases with the number of signals received. Another standard result for Bayesian learning with normal priors and signals is

$$m_p | m_\tau \sim N(m_\tau, v_p) \tag{14}$$

$$v_p = \frac{s_\tau^4 (p-\tau)}{s_\tau^2 (p-\tau) + \sigma_e^2} \tag{15}$$

for any $p > \tau$. Equation (14) implies the posterior mean of beliefs is a martingale. Conditional on current information, the expected value of any future realization of m_τ equals its current value.

2.2 Match Formation, Duration, and Wages

In each period, wages are determined by a Nash bargain between the worker and the firm. They divide the expected match surplus. They take expectations with respect to tenure τ beliefs about match quality, given the worker's quality a and the firm's quality b. The equilibrium wage therefore maps tenure τ information about the match (a, b, m_τ, s_τ^2) into a payment from the firm to the worker. Because the Nash bargain is efficient, the match only continues if the expected surplus is non-negative. Otherwise, the worker and firm both prefer to separate.

Let J_τ denote the worker's value of employment at tenure τ. Let U denote the value of the worker's outside option (unemployment). Let Π_τ denote the firm's value of employment at tenure τ, and let V denote the value of the firm's outside option (a vacancy). In the steady state, U and V are constant. At tenure τ, the match continues if and only if

$$J_\tau + \Pi_\tau \geq U + V. \tag{16}$$

When (16) is satisfied, the equilibrium wage w_τ solves the Nash bargaining condition

$$J_\tau - U = \gamma [J_\tau + \Pi_\tau - U - V] \tag{17}$$

or equivalently,

$$(1-\gamma)(J_\tau - U) = \gamma (\Pi_\tau - V) \tag{18}$$

where γ is the worker's exogenous share of match surplus.

2.2.1 The Worker's Value of Employment and Unemployment

The worker's expected value of employment at wage w_τ is

$$J_\tau = w_\tau + \beta(1-\delta) E_\tau [\max\{J_{\tau+1}, U\}] + \beta\delta U \tag{19}$$

where E_τ denotes the expectation taken with respect to tenure τ information: (a, b, m_τ, s_τ^2). For what follows, it is convenient to rewrite J_τ net of the value of unemployment, i.e.,

$$J_\tau - U = w_\tau + \beta(1-\delta) E_\tau [\max\{J_{\tau+1} - U, 0\}] - (1-\beta) U. \tag{20}$$

The steady state value of being unemployed today and behaving optimally thereafter is

$$U = h + \beta\pi \int_{b_0}^{b_1} J_0 dF_b^* + \beta(1-\pi)U \tag{21}$$

where π is the steady state probability that an unemployed worker is matched to a firm, F_b^* is the steady state distribution of firm types among open vacancies defined in Appendix B, and

$$J_0 = E_0[\max\{J_1, U\}] \tag{22}$$

is the expected value of employment before the initial signal of match quality is observed.

2.2.2 Vacancies and The Firm's Value of Employment

The firm's value of employing a worker at wage w_τ is today's expected net revenues plus the discounted expected value of employment next period. Thus,

$$\begin{aligned} \Pi_\tau &= E_\tau[q_\tau] - w_\tau + \beta(1-\delta)E_\tau[\max\{\Pi_{\tau+1}, V\}] + \beta\delta V \\ &= \mu + a + b + m_\tau - w_\tau + \beta(1-\delta)E_\tau[\max\{\Pi_{\tau+1}, V\}] + \beta\delta V \end{aligned} \tag{23}$$

so that

$$\Pi_\tau - V = \mu + a + b + m_\tau - w_\tau + \beta(1-\delta)E_\tau[\max\{\Pi_{\tau+1} - V, 0\}] - (1-\beta)V. \tag{24}$$

The production technology (6) implies that the firm's employees produce independently of one another. As a consequence, the firm's decision to open vacancies is static. The number of hires today has no dynamic consequences for future hiring or productivity. When a firm opens υ vacancies, the number l that are filled is a binomial process. The number of vacancies opened by a type b firm in each period therefore solves

$$\max_{\upsilon \in \mathbb{N}} \sum_{l=0}^{\upsilon} \binom{\upsilon}{l} \lambda^l (1-\lambda)^{\upsilon - l} \left[l \int_{a_0}^{a_1} \Pi_0 dF_a^* - \kappa(l) \right] \tag{25}$$

where λ is the steady state probability that a vacancy is filled, F_a^* is the steady state distribution of unemployed worker types defined in Appendix B, and

$$\Pi_0 = E_0[\max\{\Pi_1, V\}] \tag{26}$$

is the expected net revenues from a match before the signal x is observed.

Note that firm size (employment) is indeterminate. However, increasing and convex hiring costs κ guarantee the solution to (25) is well defined, and the firm opens a finite number of vacancies in each period. I derive the average steady state employment of a type b firm in Appendix B.

The equilibrium value of a vacancy satisfies $V = 0$. Since firms are free to open vacancies, they do so until there is no further benefit. Equivalently, since hiring costs are sunk, terminating an employment relationship frees up no resources. Thus the alternative value of a vacancy is zero.

2.2.3 The Equilibrium Wage

With the value functions in hand, it is a simple matter to solve for the equilibrium wage. It takes a remarkably simple form, summarized in Proposition 1.

Proposition 1 (Equilibrium Wage) *At each tenure $\tau > 0$, the equilibrium wage w_τ is linear and additively separable in a person-specific component, a firm-specific component, and the posterior mean of beliefs about match quality; and is independent of s_τ^2.*

Proof. Substituting (20) and (24) into the Nash bargaining condition (18) we obtain

$$\begin{aligned} & (1-\gamma)\left(w_\tau + \beta(1-\delta) E_\tau\left[\max\left\{J_{\tau+1} - U, 0\right\}\right] - (1-\beta) U\right) \\ = \; & \gamma\left(\mu + a + b + m_\tau - w_\tau + \beta(1-\delta) E_\tau\left[\max\left\{\Pi_{\tau+1} - V, 0\right\}\right] - (1-\beta) V\right). \end{aligned} \tag{27}$$

Condition (18) implies

$$(1-\gamma) E_\tau\left[\max\left\{J_{\tau+1} - U, 0\right\}\right] = \gamma E_\tau\left[\max\left\{\Pi_{\tau+1} - V, 0\right\}\right] \tag{28}$$

and thus

$$(1-\gamma)\left(w_\tau - (1-\beta) U\right) = \gamma\left(\mu + a + b + m_\tau - w_\tau - (1-\beta) V\right). \tag{29}$$

Rearranging yields

$$w_\tau = \gamma\mu + \theta + \psi + \gamma m_\tau \tag{30}$$

where the worker specific component is

$$\theta = \gamma a + (1-\gamma)(1-\beta) U \tag{31}$$

and the firm-specific component is

$$\psi = \gamma b - \gamma(1-\beta) V = \gamma b \tag{32}$$

given the equilibrium condition $V = 0$. ■

The equilibrium wage function (30) has the same additively separable structure as the AKM empirical specification (1). In keeping with the empirical literature, I therefore refer to θ and ψ as person and firm effects, respectively. Equations (31) and (32) provide a behavioral interpretation of the AKM person and firm effects. Equation (32) illustrates that the firm effect measures the worker's share γ of the firm's contribution to match surplus. It linearly rescales the firm's productivity b. Rewriting equation (31) as $\theta = \gamma(a - (1-\beta)U) + (1-\beta)U$ demonstrates that the person effect is likewise the worker's share of his contribution to match surplus, plus compensation for forgoing his next-best alternative and adjusted for discounting. It therefore reflects the worker's productivity a, adjusted for discounting and bargaining strength.

The Jovanovic (1979) equilibrium wage is a special case of (30). In that model, workers and firms are identical but matches are heterogeneous, and production occurs according to the continuous time analog of (6) with $a = b = 0$ for every worker and firm. The Jovanovic wage equals expected marginal product, which in that case is also the posterior mean of beliefs about match quality. This fundamental result relies on the assumption that firms earn zero expected profit. Similar to Jovanovic's model, the equilibrium wage (30) is linear and additively separable in expected marginal product, $\mu + a + b + m_\tau$, and in the posterior mean of beliefs about match quality, m_τ. To see that the Jovanovic (1979) equilibrium wage is in fact a special case of (30), notice that when workers capture the entire match surplus, i.e., as $\gamma \to 1$ (so that firms earn zero expected profit), the equilibrium wage approaches $\lim_{\gamma \to 1} w_\tau = \mu + a + b + m_\tau$. That is, the equilibrium wage converges to the expected marginal product of the match. With $a = b = 0$, this is exactly the Jovanovic (1979) equilibrium wage.

The wage function (30) implies rich equilibrium wage dispersion. The person effect, the firm effect, and learning about match quality all contribute. Identical workers earn different wages because of employment at heterogeneous firms and because of heterogeneity of beliefs about match quality. Identical firms pay different wages because they employ heterogeneous workers and because of heterogeneity in beliefs. Even identical workers employed in identical firms earn different wages because of dispersion in beliefs about match quality.

2.2.4 The Separation Decision

In the Nash bargaining framework, the separation decision is made jointly by the worker and firm. The match continues as long as the surplus is non-negative. To characterize the separation decision, it is useful to introduce the Bellman equation for the joint value of employment, W :

$$\begin{aligned} W\left(m_\tau, s_\tau^2\right) &= \max\left\{J_\tau + \Pi_\tau, U + V\right\} \\ &= \max\left\{\mu + a + b + m_\tau + \beta\left(1-\delta\right)E_\tau\left[W\left(m_{\tau+1}, s_{\tau+1}^2\right)\right] + \beta\delta U, U\right\} \end{aligned} \quad (33)$$

given the equilibrium condition $V = 0$. The value function W depends on the complete set of state variables (a, b, m_τ, s_τ^2). However, its dependence on a and b is suppressed in (33) for notational simplicity, and because these quantities do not vary with tenure.

The following Proposition establishes uniqueness of the value function, and its most important properties with regard to the posterior mean of beliefs about match quality. Its proof is in Appendix A.

Proposition 2 (Uniqueness) *There is a unique value function W that satisfies the Bellman equation (33). Furthermore, W is continuous, nondecreasing, and convex in m_τ.*

Workers and firms prefer to continue the employment relationship as long as its value exceeds the value of terminating it. That is, as long as the first argument of the max operator in (33) exceeds the second argument. They prefer to terminate the relationship the first time the inequality is reversed. There are a number of equivalent ways of characterizing this decision in terms of state variables. The most convenient characterization is in terms of beliefs about match quality, since other state variables do not vary over the course of the employment relationship. This characterization corresponds with previous learning models with homogeneous workers and firms. Proposition 3 summarizes the optimal separation policy in terms of beliefs about match quality. Its proof is in Appendix A.

Proposition 3 (Optimal Separation Policy) *At each tenure $\tau > 0$ and for given values of a and b, the optimal separation policy is characterized by a reservation value of beliefs* *about match quality, $\bar{m}_\tau$. Specifically, the optimal policy is to separate if $m_\tau < \bar{m}_\tau$, and continue if $m_\tau \geq \bar{m}_\tau$.*

The reservation level of beliefs about match quality is the value of m_τ at which workers and firms are indifferent between continuing the employment relation and terminating it. Thus $\bar{m}_\tau$ satisfies the Nash continuation condition (16) with equality. Equivalently, it is the value of m_τ that equates the arguments of the max function in the Bellman equation (33). Thus $\bar{m}_\tau$ is implicitly defined by

$$\bar{m}_\tau = (1 - \beta\delta) U - \mu - a - b - \beta (1 - \delta) \bar{E}_\tau \left[W \left(m_{\tau+1}, s_{\tau+1}^2 \right) \right] \tag{34}$$

where $\bar{E}_\tau$ denotes the expectation taken with respect to state variables $(a, b, \bar{m}_\tau, s_\tau^2)$.

It is of considerable interest to characterize how separation behavior evolves with tenure. Proposition 4 establishes a standard result for equilibrium learning models. Its proof is appendicized.

Proposition 4 (Monotonicity) *The reservation value of beliefs about match quality is monotone in tenure, i.e., $\bar{m}_{\tau+1} \geq \bar{m}_\tau$ for all $\tau > 0$.*

The result in Proposition 4 reflects the option value of employment. Early in the match, when beliefs about match quality are imprecise, workers and firms are willing to accept matches of low perceived quality because their point estimate m_τ may be inaccurate. As the worker and firm acquire more information, their beliefs become increasingly precise. As a consequence, the worker and firm become increasingly selective about admissible values of match quality, and the reservation value increases. Asymptotically, $\lim_{\tau\to\infty} \bar{m}_\tau = [1 - \beta(1-\delta)] U - \mu - a - b$.

2.3 Comparative Statics

This section explores how separation behavior varies with worker and firm quality. There are several equivalent characterizations. We begin with a proposition that characterizes how the reservation value of beliefs about match quality varies with worker and firm quality. The Proof is in Appendix A.

Proposition 5 *At each tenure $\tau > 0$, the reservation value of beliefs about match quality is decreasing in worker and firm quality. That is,*

$$\frac{\partial \bar{m}_\tau}{\partial a} < 0, \quad \frac{\partial \bar{m}_\tau}{\partial b} < 0. \tag{35}$$

This result is fairly intuitive. Consider the changing the firm's quality b. This has no effect on the firm's outside option since $V = 0$ in equilibrium, and no effect on the value of the worker's outside option. Thus it only affects the value of remaining in the match, $J_\tau + \Pi_\tau$. Increasing b raises the wage (via ψ), and hence increases J_τ. Likewise, it increases the net value of output $(q_\tau - w_\tau)$ and hence Π_τ. Thus an increase in b raises the value of remaining in the match, and makes the worker and firm less selective about the set of acceptable values of match quality. Having found a "good" employer, the worker is less picky about whether or not it is a "good" match. Since all workers are highly productive at "good" firms, the firm is less picky about whether or not they are "good" matches.

Similar intuition explains why an increase in a reduces the reservation value of beliefs about match quality, with one complication: increasing a raises the worker's productivity not only in the current match, but in all matches. That is, the value of the worker's outside option is increasing in a (see Lemma 6 in Appendix A). Nevertheless, matching frictions ensure that increasing a raises the value of continuing the match more than the value of terminating it.[7] From the firm's perspective, having found a "good" employee, the firm is less picky about whether or not she is a "good" match.

Proposition 5 has obvious implications for the relationship between job duration and worker/firm quality. Expected job duration is decreasing in $\bar{m}_\tau$. Equation (35) therefore implies expected job duration is increasing in worker and firm quality. More productive workers experience longer average job duration than less productive ones, and more productive firms experience less turnover than less productive ones. This is consistent with stylized facts.

To this point, we have characterized the separation decision in terms of $\bar{m}_\tau$. This is the most natural characterization to analyze separation behavior *within* a worker-firm match, since beliefs about match quality are the only quantity that varies over the course of the match. Alternately, we can characterize the separation decision in terms of a reservation level of worker quality, $\bar{a}$, for given firm quality and beliefs about match quality; or symmetrically, in terms of a reservation level of firm quality, $\bar{b}$, for given values of a and m_τ. These quantities are more natural to ask whether matches are assortative. Like $\bar{m}_\tau$, they are defined by equating the two arguments of the max operator in the value function (33). That is, $\bar{a}$ and $\bar{b}$ are implicitly defined by:

$$\bar{a} = (1-\beta\delta)\,U - \mu - b - m_\tau - \beta\,(1-\delta)\,E_\tau\left[W\left(m_{\tau+1}, s^2_{\tau+1}\right)\right] \quad (36)$$
$$\bar{b} = (1-\beta\delta)\,U - \mu - a - m_\tau - \beta\,(1-\delta)\,E_\tau\left[W\left(m_{\tau+1}, s^2_{\tau+1}\right)\right] \quad (37)$$

where the expectation in (36) is taken with respect to $(\bar{a}, b, m_\tau, s^2_\tau)$, and the expectation in (37) is taken with respect to $\left(a, \bar{b}, m_\tau, s^2_\tau\right)$. Differentiating (36) and (37) reveals:[8]

$$\frac{\partial \bar{a}}{\partial b} < 0, \ \frac{\partial \bar{b}}{\partial a} < 0.$$

These show that holding beliefs about match quality constant, matches between workers and firms are negatively assortative: more productive firms are willing to match with less productive workers, and vice versa. This result is consistent with other models where production is additively separable in agents' types, e.g. examples presented in Shimer and Smith (2000) (a search example) and Shimer (2005) (an assignment example). Negative assortative matching is also consistent with recent empirical evidence that finds a negative correlation between empirical person and firm effects.[9] The intuition is straightforward. Using Sattinger's (1975) definition of comparative advantage, the production function (6) implies

[7] That is, using the result of Lemma 6 in Appendix A, it is easy to show that $\partial J_\tau/\partial a > \partial U/\partial a$.

[8] The algebra is omitted but available on request. The method of proof parallels that of Proposition 5.

[9] For instance, ACK find negative correlations in France and Washington State, Goux and Maurin (1999) find a negative correlation in France using different data and methods, Gruetter and Lalive (2004) find a similar result in Austrian data, as do Barth and Dale-Olsen (2003) in Norwegian data.

low-ability workers have a comparative advantage in high-productivity firms.[10] Furthermore, since high-productivity firms have the largest opportunity cost of an unfilled vacancy, they prefer to match with low productivity workers than to leave the vacancy unfilled. This is a boon to low-ability workers. The potential benefit of matching with a high-productivity firm is sufficient that they would rather wait than accept employment at a low-productivity firm. Low-ability workers also have the lowest opportunity cost of unemployment, and hence are most willing to wait for a match with a high-productivity firm. The converse is true for high-ability workers and low-productivity firms.

2.4 Discussion

Before turning to empirics, it is useful to discuss various predictions that stem from the matching model with regards to equilibrium wages, mobility, turnover, and firm size. We will look for the empirical counterparts to these when assessing the empirical specification.

The matching model predicts that wages are additively separable in person- and firm-specific components: the person and firm effects θ and ψ. We know from the empirical work of AKM and others that this additively separable structure is consistent with wage data. In the typical application, it explains over 90 percent of wage variation. Since the person and firm effects are functions of the productivity parameters a and b, productivity differences across workers and firms are one source of equilibrium wage dispersion. These productivity differences are observable by match participants, but not necessarily by the econometrician.

Equilibrium wages are also linear and additively separable in the posterior mean of beliefs about match quality m_τ. Dispersion in beliefs therefore introduces additional equilibrium wage dispersion. Since the person and firm effects do not vary within a worker-firm match, all within-match wage variation is due to the evolution of beliefs about match quality. Furthermore, since beliefs evolve according to Bayes' rule, m_τ is a martingale. Thus the model predicts that within a worker-firm match, wages are a martingale.[11] This is the basis of the test developed in Section 3.4.

The martingale property is common to most Bayesian learning models, e.g. Farber and Gibbons (1996) and Gibbons et al. (2005). It has several economic consequences: it implies wage shocks are permanent and diminish with tenure,[12] and that the variance of wages increases with tenure. The latter follows because the variance of m_τ increases with the number of observed signals: $V_{\tau+1} > V_\tau$. This may seem at odds with the notion that beliefs about match quality become increasingly precise with tenure. However, it is important to

[10] That is, for $a' > a$ abd $b' > b$,

$$\frac{\mu + a' + b' + c + e_\tau}{\mu + a' + b + c + e_\tau} < \frac{\mu + a + b' + c + e_\tau}{\mu + a + b + c + e_\tau}.$$

[11] Due to the selection process that terminates a match if $m_\tau < \bar{m}_\tau$, the wage sequence observed by an econometrician is a submartingale.

[12] Recall the definition of m_τ in equation (9). It implies shocks to beliefs about match quality (z and e_τ) are permanent. Within a match, these are the only shocks to wages. Thus wage shocks are permanent. Recall further that m_τ is a precision-weighted average of $m_{\tau-1}$ and the signal $c + e_\tau$. The precision of the signals (shocks) is constant, but the precision of beliefs increases with tenure. Thus each successive signal (shock) receives smaller weight in the updating process. Asymptotically, new signals receive zero weight.

distinguish between the variance of beliefs, s_τ^2, that declines with tenure, and the variance of the posterior mean of beliefs, V_τ, that increases with tenure.[13] Since tenure and labor marker experience are positively correlated empirically, this is consistent with the well known empirical regularity that the variance of wages increases with experience (e.g., Mincer (1974) and many others).

Section 2.3 presented comparative statics that characterize how separation behavior varies with worker and firm quality, a and b. The empirical exercise that follows focuses on estimating the empirical person and firm effects. We therefore seek predictions regarding the relationship between separation behavior and θ and ψ. Applying Lemma 6 (see Appendix A), it is a simple matter to show $\partial\theta/\partial a > 0$ and $\partial\psi/\partial b > 0$. Combining this with (35), at each $\tau > 0$ we have

$$\frac{\partial \bar{m}_\tau}{\partial \psi} < 0, \quad \frac{\partial \bar{m}_\tau}{\partial \theta} < 0. \tag{38}$$

Because expected job duration is decreasing in $\bar{m}_\tau$, the first inequality in (38) implies that expected job duration is increasing in ψ. As we shall see in Section 5, this is consistent with empirical evidence. A corollary is that turnover is decreasing in ψ. It follows that *ceteris paribus,* firm size (employment) is increasing in ψ. This is consistent with Brown and Medoff (1989) and others who find that conditional on observable worker and firm characteristics, large firms pay higher wages than small firms. A similar prediction arises from the second inequality in (38): on average, workers with large person effects experience longer job duration and change jobs less often than workers with small θ. Lillard (1999) finds a similar result in NLSY data.[14]

The negative assortative matching result of the previous section carries over to the person and firm effects. The model therefore predicts empirical estimates of θ and ψ will be negatively correlated. This is consistent with most empirical studies based on the fixed effect estimator, e.g., AKM, ACK, Barth and Dale-Olsen (2003), Goux and Maurin (1999), and Gruetter and Lalive (2004).

The model also predicts that in a cross-section, workers with longer job tenure earn higher wages on average than their counterparts with lower tenure. This is consistent with stylized facts about labor markets and numerous empirical findings, e.g., Mincer and Jovanovic (1981), Bartel and Borjas (1981), and many others.[15] The argument is as follows. First, larger values of θ and ψ are associated with higher wages and longer expected duration.

[13] The variance of beliefs s_τ^2 declines because agents learn: as they acquire more information about true match quality, their beliefs become increasingly precise. In contrast, the prior variance of the mean of beliefs is zero: all agents have common priors about match quality. As information is acquired, the posterior mean of beliefs converges to the true match quality. It follows that V_τ increases from its prior value (zero) to its asymptotic value (σ_c^2) as tenure increases.

[14] Lillard (1999) estimates simultaneous wage and turnover equations with random person and job effects. His job effect is nested within the person effect and thus is not directly comparable to the firm effects discussed here. He finds a negative correlation between the person effect in the wage equation and the person effect in the job turnover hazard: higher values of the person-wage effect are associated with a reduced turnover hazard.

[15] More recent research has focused on the causal link between job tenure and earnings growth using longitudinal data. Examples include Abraham and Farber (1987), Altonji and Shakotko (1987), and Topel and Ward (1992). Dostie (2005) is a recent study using longitudinal linked data. The matching model implies that conditional on person and firm effects, all returns to tenure are due to accumulated knowledge about match quality. This accumulated knowledge is a form of match-specific human capital. It is not "productive"

Second, conditional on θ and ψ, better matches last longer and are associated with larger values of m_τ on average, and hence with higher wages. Third, because the reservation level of beliefs about match quality is monotone in tenure, the left tail of the wage distribution is increasingly truncated as tenure increases. All three effects operate in concert to induce a positive relationship between tenure and wages.

Finally, we have noted at several points that the additively separable structure of the equilibrium wage function is the same as the AKM empirical specification. This specification has been criticized for relying on an exogenous mobility assumption (discussed further in Section 3). The matching model implies the AKM specification is consistent with endogenous separations due to learning about match quality. However, it is important to note that wages and separations are jointly determined. The posterior mean of beliefs about match quality enters the equilibrium wage, but the optimal separation policy implies the employment relationship only continues as long as $m_\tau \geq \bar{m}_\tau$. Thus an econometrician observes a truncated earnings distribution: earnings are only observed if $w_\tau \geq \gamma\mu + \theta + \psi + \gamma\bar{m}_\tau$. Failing to correct for truncation of the earnings distribution may bias empirical estimates of the person and firm effects. I develop an appropriate correction in what follows.

3 Empirical Specification

At the core of the empirical specification is a linear wage equation like the AKM model. The adopted specification explicitly accounts for truncation of the earnings distribution as implied by the optimal separation policy. I also exploit the matching model's structure to test whether the conditional distribution of earnings is consistent with learning about match quality.

Consider the following empirical counterpart to (30):

$$y_{ijt} = \mu + x_{it}'\beta + \theta_i + \psi_j + \varepsilon_{ijt} \tag{39}$$

$$\varepsilon_{ijt} = \gamma m_{ij\tau} + u_{ijt} \tag{40}$$

where $i = 1, ..., N$ indexes workers and $j = 1, ..., J$ indexes firms; y_{ijt} is the natural logarithm of employment earnings; μ is the grand mean; x_{it} is a vector of observable time-varying covariates;[16] β is a parameter vector; θ_i is the person effect; ψ_j is the firm effect of the firm j at which worker i was employed in t (denoted $j = J(i,t)$); ε_{ijt} is a compound statistical error that consists of the posterior mean of beliefs about match quality $m_{ij\tau}$ times the worker's bargaining strength γ, and classical measurement error u_{ijt}. I assume $m_{ij\tau}$ is independent of u_{ijt}. As in Section 2, τ indexes tenure. Equations (39) and (40) introduce an additional index t of calendar time.[17]

human capital since productivity is constant over the duration of the match except for stochastic variation induced by e_τ. Nevertheless, it has value: it takes time to accumulate, and is lost when the match terminates.

[16] Covariates include year dummies, a quartic in experience (interacted with sex), and dummies for the number of quarters worked in the year (also interacted with sex).

[17] The inclusion of time-varying covariates x_{it} in (39) necessitates the additional calendar time index t. Tenure and calendar time are in general related by a simple function. I therefore frequently suppress one of these indices to economize on notation.

Some comments are in order. First, notice that the dependent variable in (39) is log earnings, while the equilibrium wage function is specified in levels. The semi-log specification is a first-order approximation to a specification on levels.[18] It is also, of course, the standard specification for earnings regressions, and is adopted here for all the usual reasons.[19] Second, notice that (39) includes time-varying covariates x_{it}, whereas the equilibrium wage function (30) does not. The covariates admit variation over time in the theoretical mean of earnings ($\gamma\mu$), and variation in earnings due to labor force experience and attachment.[20] Finally, since beliefs about match quality are not observed by the econometrician, the posterior mean $m_{ij\tau}$ has been subsumed into the error term. Recall that match quality and its signals are drawn independently of the person and firm effects. However, the optimal separation policy of workers and firms introduces a selection process. Since worker-firm matches terminate when $m_{ij\tau} < \bar{m}_{ij\tau}$, the observed distribution of earnings is truncated. Furthermore, since $\partial\bar{m}_{ij\tau}/\partial\theta_i < 0$ and $\partial\bar{m}_{ij\tau}/\partial\psi_j < 0$, the selection process induces a negative correlation between $m_{ij\tau}$ and the person and firm effects. We therefore correct for incidental truncation, as described in Section 3.3 below. The corrected error is uncorrelated with the person and firm effects by construction.

We can further decompose the person effect θ_i into components observed and unobserved by the econometrician as

$$\theta_i = \alpha_i + u_i'\eta \tag{41}$$

where α_i is the unobserved component of the person effect; u_i is a vector of time-invariant person characteristics observed by the econometrician; and η is a parameter vector.

Let N^* denote the total number of observations; q the number of time-varying covariates including the constant term; and p the number of time-invariant person characteristics. Rewriting (39) and (41) in matrix notation, we have

$$y = X\beta + U\eta + D\alpha + F\psi + \varepsilon \tag{42}$$

where y is the $N^* \times 1$ vector of earnings outcomes, X is the $N^* \times q$ matrix of time-varying covariates including the intercept; β is a $q \times 1$ parameter vector; U is the $N^* \times p$ matrix of time-invariant person characteristics; η is a $p \times 1$ parameter vector; D is the $N^* \times N$ design

[18] Write the model on earnings levels as

$$w_{ijt} = \mu\left(1 + x_{it}'\frac{\beta}{\mu} + \frac{\theta_i}{\mu} + \frac{\psi_j}{\mu} + \frac{\varepsilon_{ijt}}{\mu}\right)$$

so that

$$\begin{aligned}\ln w_{ijt} &\approx \ln\mu + x_{it}'\frac{\beta}{\mu} + \frac{\theta_i}{\mu} + \frac{\psi_j}{\mu} + \frac{\varepsilon_{ijt}}{\mu} \\ &= \mu^* + x_{it}'\beta^* + \theta_i^* + \psi_j^* + \varepsilon_{ijt}^*\end{aligned}$$

where the first line uses the first-order Taylor series approximation $\ln(1+x) \approx x$ around $x = 0$, and where $\mu^* = \ln\mu$, $\beta^* = \beta/\mu$, $\theta_i^* = \theta/\mu$, $\psi_j^* = \psi/\mu$, and $\varepsilon_{ijt}^* = \varepsilon_{ijt}/\mu$.

[19] Complete estimation results on earnings levels are available on request. The regressions on levels show some evidence of mis-specification, which was the primary reason for adopting the semi-log alternative.

[20] As noted in footnote 6, a more general model that includes deterministic human capital accumulation and publicly-observable stochastic aggregate productivity shocks yields an equilbrium wage that is additively separable in θ, ψ, m_τ, an experience effect, and time effects, like (39).

matrix of the unobserved component of the person effect; α is the $N \times 1$ vector of person effects; F is the $N^* \times J$ design matrix of the firm effects; ψ is the $J \times 1$ vector of firm effects; and ε is the $N^* \times 1$ vector of errors.

I consider two estimators of the parameters in (42). The first is the direct least squares estimator of ACK, that treats $\beta, \eta, \alpha,$ and ψ as fixed. The second is a mixed effects estimator that treats β and η as fixed, and α and ψ as random. Abowd and Kramarz (1999) briefly discuss a similar mixed effects specification, but do not estimate it.[21] The mixed effects estimator has some compelling properties. First, as shown in Section 3.2.2, the least squares estimator is a special case of the mixed effects estimator. Second, the theoretical person and firm effects are random variables, which argues in favor of treating their empirical counterparts similarly.[22] Third, with a large number of person and firm effects to estimate, the fixed effects estimator is costly in terms of degrees of freedom. The mixed model is less demanding in this regard, since it only requires estimation of a few parameters of the distribution of the random effects. Their realized values, i.e., the Best Linear Unbiased Predictors (BLUPs), are estimated in a subsequent step. Fourth, as Robinson (1991) demonstrates, BLUPs typically have better sampling properties than fixed effect estimates of the unobserved heterogeneity. Fifth, the mixed effects estimator permits simultaneous maximum likelihood estimation of the error covariance within a worker-firm match, which I use to test the learning hypothesis described in Section 3.4. Finally, the mixed effects specification permits out-of-sample prediction of person and firm effects, which I use to validate the specification.

3.1 The Fixed Model

The fixed model is completely specified by (42) and the stochastic assumptions:

$$E\left[\varepsilon_{ijt} | i, j, t, x, u\right] = 0 \quad (43)$$
$$E\left[\varepsilon\varepsilon'\right] = \sigma_\varepsilon^2 I_{N^*} \quad (44)$$

where I_{N^*} is the identity matrix of order N^*. Equation (43) is the exogenous mobility assumption for which the AKM specification has been criticized. The least squares estimator of β, η, α, and ψ solves the normal equations

$$\begin{bmatrix} X'X & X'U & X'D & X'F \\ U'X & U'U & U'D & U'F \\ D'X & D'U & D'D & D'F \\ F'X & F'U & F'D & F'F \end{bmatrix} \begin{bmatrix} \beta \\ \eta \\ \alpha \\ \psi \end{bmatrix} = \begin{bmatrix} X'y \\ U'y \\ D'y \\ F'y \end{bmatrix}. \quad (45)$$

In the data described in Section 4, the cross product matrix on the left hand side of (45) is of sufficiently high dimension to preclude estimation using standard software. I therefore compute least squares solutions $\hat{\beta}$, $\hat{\eta}$, $\hat{\alpha}$, and $\hat{\psi}$ using the iterative conjugate gradient method of ACK to directly minimize the sum of squared residuals. Their algorithm exploits the sparse structure of the cross product matrix after blocking on connected groups of workers

[21] To my knowledge, the application herein is the first to estimate random person and firm effects.

[22] When unobservables are random variables drawn from a homogeneous population distribution, statisticians argue in favor of the mixed effects estimator. See Robinson (1991) or Searle et al. (1992) for such an argument.

and firms.[23] The resulting estimates of α and ψ are not unique, since the design matrices D and F are not full rank. ACK discuss identification of the person and firm effects in detail. I apply their grouping procedure to obtain unique estimates of α and ψ subject to the restriction that their overall and group means are zero. When there are G connected groups of workers and firms, this procedure identifies an overall constant term, and a set of $N + J - G - 1$ person and firm effects measured as deviations from the overall constant and group-specific means.

3.2 The Mixed Model

The mixed model specification treats β and η as fixed, and α and ψ as random. The model is completely specified by (42) and the stochastic assumption

$$\begin{bmatrix} \alpha \\ \psi \\ \varepsilon \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 I_N & 0 & 0 \\ 0 & \sigma_\psi^2 I_J & 0 \\ 0 & 0 & R \end{bmatrix}\right). \tag{46}$$

It is worth noting that (42) and (46) *do not* assume that the design of the random effects (D and F) is orthogonal to the design (X and U) of the fixed effects (β and η). Such an assumption is usually violated in economic data.

I estimate two alternate parameterizations of the error covariance R.[24] The simplest assumes spherical errors, i.e., $R = \sigma_\varepsilon^2 I_{N^*}$. I estimate this specification primarily for comparison with the fixed model. The second parameterization puts no restrictions on the within-match error covariance other than symmetry and positive semi-definiteness. Let M denote the number of worker-firm matches in the data, and let $\bar{\tau}$ denote the maximum observed duration of a worker-firm match. Suppose the data are ordered by t within j within i. In the balanced data case, where there are $\bar{\tau}$ observations on each worker-firm match, we can write the second parameterization as

$$R = I_M \otimes W \tag{47}$$

where W is the $\bar{\tau} \times \bar{\tau}$ within-match error covariance.

The extension to unbalanced data, where each match between worker i and firm j has duration $\tau_{ij} \leq \bar{\tau}$, is straightforward. Define a $\bar{\tau} \times \tau_{ij}$ selection matrix S_{ij} with elements on the principal diagonal equal to 1, and off-diagonal elements equal to zero.[25] S_{ij} selects those rows and columns of W that correspond to observed earnings outcomes in the match between worker i and firm j. We can generalize equation (47) to the unbalanced data case by pre- and post-multiplying W by S_{ij}. That is, R remains block-diagonal, but the diagonal

[23] See Searle (1987) for a statistical discussion of connectedness. In labor market data, firms are connected by common employees; workers are connected by common employers. ACK present a graph-theoretic algorithm to identify connected groups of workers and firms.

[24] Estimates of a variety of ARMA error covariances are also available on request.

[25] For example, if $\bar{\tau} = 3$ and a match between worker i and firm j lasts for 2 periods,

$$S_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

block corresponding to the match between worker i and firm j is now the $\tau_{ij} \times \tau_{ij}$ submatrix $W_{ij} = S'_{ij} W S_{ij}$.

3.2.1 REML Estimation of the Mixed Model

I estimate the variance components $\left(\sigma^2_\alpha, \sigma^2_\psi\right)$ and R by Restricted Maximum Likelihood (REML).[26] REML is frequently described as maximizing that part of likelihood that is invariant to the fixed effects (i.e., β and η). Formally, REML is maximum likelihood on linear combinations of the dependent variable y, chosen so that the linear combinations do not contain any of the fixed effects. The linear combinations $k'y$ are chosen so that

$$k' \left(X\beta + U\eta\right) = 0 \;\; \forall \beta, \eta \tag{48}$$

which implies

$$k' \left[\begin{array}{cc} X & U \end{array}\right] = 0. \tag{49}$$

Thus k' projects onto the space orthogonal to $\left[\begin{array}{cc} X & U \end{array}\right]$ and is of the form

$$k' = c' \left[I_{N^*} - \left[\begin{array}{cc} X & U \end{array}\right] \left(\left[\begin{array}{c} X' \\ U' \end{array}\right] \left[\begin{array}{cc} X & U \end{array}\right] \right)^{-} \left[\begin{array}{c} X' \\ U' \end{array}\right] \right] \equiv c' M_{XU} \tag{50}$$

for arbitrary c', and where A^- denotes the generalized inverse of A. When $\left[\begin{array}{cc} X & U \end{array}\right]$ has rank $r \leq q + p$, there are $N^* - r$ linearly independent vectors k' satisfying (48). Define $K' = CM_{XU}$ with rows k' satisfying (48), and where K' and C have full row rank $N^* - r$. REML estimation is maximum likelihood on $K'y$. For $y \sim N\left(X\beta + U\eta, \mathbf{V}\right)$ it follows that

$$K'y \sim N\left(0, K'\mathbf{V}K\right) \tag{51}$$

where $\mathbf{V} = DD'\sigma^2_\alpha + FF'\sigma^2_\psi + R$ is the covariance implied by (46).

The REML estimator has a number of attractive properties. REML estimates are invariant to the choice of C and are invariant to the value of the fixed effects (i.e., β and η). Furthermore, since REML is based on the maximum likelihood principle, it inherits the consistency, efficiency, asymptotic normality, and invariance properties of ML.

I maximize the REML log-likelihood implied by (51) using the Average Information (AI) algorithm of Gilmour et al. (1995).[27] The AI algorithm is a variant of Fisher scoring that uses a computationally convenient average of the expected and observed information matrices to compute parameter updates during iterative maximization of the REML log-likelihood.

3.2.2 Estimating the Fixed Effects and Realized Random Effects

The REML estimator does not directly estimate the fixed effects β and η. Henderson et al. (1959) derived a system of equations that simultaneously yield the BLUE of fixed effects and

[26] REML is the estimator of choice in applied fields where mixed model estimation is common, e.g., statistical genetics, and among plant and animal breeders.

[27] The AI algorithm is implemented in the ASREML software package.

the BLUP of random effects. These have become known as the mixed model equations or Henderson equations. Define the matrix of variance components

$$G = \begin{bmatrix} \sigma_\alpha^2 I_N & 0 \\ 0 & \sigma_\psi^2 I_J \end{bmatrix}. \tag{52}$$

The mixed model equations are

$$\begin{bmatrix} \begin{bmatrix} X' \\ U' \end{bmatrix} R^{-1} \begin{bmatrix} X & U \end{bmatrix} & \begin{bmatrix} X' \\ U' \end{bmatrix} R^{-1} \begin{bmatrix} D & F \end{bmatrix} \\ \begin{bmatrix} D' \\ F' \end{bmatrix} R^{-1} \begin{bmatrix} X & U \end{bmatrix} & \begin{bmatrix} D' \\ F' \end{bmatrix} R^{-1} \begin{bmatrix} D & F \end{bmatrix} + G^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\beta} \\ \tilde{\eta} \\ \tilde{\alpha} \\ \tilde{\psi} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ U'R^{-1}y \\ D'R^{-1}y \\ F'R^{-1}y \end{bmatrix} \tag{53}$$

where $\tilde{\beta}, \tilde{\eta}, \tilde{\alpha}$, and $\tilde{\psi}$ denote solutions for the various effects. Standard practice, which I apply here, is to solve (53) conditional on REML estimates of the variance components.

The BLUPs $\tilde{\alpha}$ and $\tilde{\psi}$ have the following properties.[28] They are *best* in the sense of minimizing the mean square error of prediction among linear unbiased estimators. They are *linear* in y, and *unbiased* in the sense $E(\tilde{\alpha}) = E(\alpha)$ and $E(\tilde{\psi}) = E(\psi)$.

The mixed model equations make clear the relationship between the fixed and mixed models. In particular, as $G \to \infty$ with $R = \sigma_\varepsilon^2 I_{N^*}$, the mixed model equations (53) converge to the normal equations (45). Thus the mixed model solutions $(\tilde{\beta}, \tilde{\eta}, \tilde{\alpha}, \tilde{\psi})$ converge to the least squares solutions $(\hat{\beta}, \hat{\eta}, \hat{\alpha}, \hat{\psi})$. Thus the least squares estimator is a special case of the mixed model estimator.

3.3 Correcting for Truncation of the Earnings Distribution

In the matching model, a match between worker and firm terminates when the point estimate of match quality m_τ falls below the reservation value $\bar{m}_\tau$. This implies the distribution of earnings is truncated. Earnings are only observed when $m_\tau \geq \bar{m}_\tau$.

If we iterate forward on the definition of $\bar{m}_\tau$ in (34) we obtain

$$\bar{m}_\tau = -[\mu A_\tau + B_\tau] - [a - U(1-\beta)] A_\tau - bA_\tau \tag{54}$$

where

$$A_\tau = 1 + \sum_{s=1}^{\infty} [\beta(1-\delta)]^s \int_{\bar{m}_{\tau+1}}^{\infty} \int_{\bar{m}_{\tau+2}}^{\infty} \cdots \int_{\bar{m}_{\tau+s}}^{\infty} dF_{\tau+s} \cdots dF_{\tau+2} d\bar{F}_{\tau+1} \tag{55}$$

$$B_\tau = \sum_{s=1}^{\infty} [\beta(1-\delta)]^s \int_{\bar{m}_{\tau+1}}^{\infty} \int_{\bar{m}_{\tau+2}}^{\infty} \cdots \int_{\bar{m}_{\tau+s}}^{\infty} m_{\tau+s} dF_{\tau+s} \cdots dF_{\tau+2} d\bar{F}_{\tau+1} \tag{56}$$

where $F_\tau = F\left(m_\tau | m_{\tau-1}, s_{\tau-1}^2\right)$ and $\bar{F}_\tau = F\left(m_\tau | \bar{m}_{\tau-1}, s_{\tau-1}^2\right)$. Reintroducing the person and firm subscripts, I approximate (54) by

$$\bar{m}_{ij\tau} = -\mu_\tau - \zeta_{i\tau} - \xi_{j\tau}. \tag{57}$$

[28] See Robinson (1991) for details.

Recall $m_{ij\tau} \sim N(0, V_\tau)$. Under the approximation (57), the probability of observing earnings at tenure τ is

$$\Pr(m_{ij\tau} \geq \bar{m}_{ij\tau}) = 1 - \Phi\left(\frac{-\mu_\tau - \zeta_{i\tau} - \xi_{j\tau}}{V_\tau^{1/2}}\right) = \Phi\left(\frac{\mu_\tau + \zeta_{i\tau} + \xi_{j\tau}}{V_\tau^{1/2}}\right) \tag{58}$$

where Φ is the standard normal CDF. The conditional expectation of observed earnings is therefore

$$\begin{aligned} E[y_{ijt} | m_{ij\tau} \geq \bar{m}_{ij\tau}] &= \mu + x'_{it}\beta + \theta_i + \psi_j + \gamma V_\tau^{1/2} \frac{\phi\left(\frac{\mu_\tau + \zeta_{i\tau} + \xi_{j\tau}}{V_\tau^{1/2}}\right)}{\Phi\left(\frac{\mu_\tau + \zeta_{i\tau} + \xi_{j\tau}}{V_\tau^{1/2}}\right)} \\ &= \mu + x'_{it}\beta + \theta_i + \psi_j + \gamma V_\tau^{1/2} \lambda_{ij\tau} \end{aligned} \tag{59}$$

where $\lambda_{ij\tau}$ is the familiar Inverse Mills' Ratio.

I perform a two-step truncation correction based on (58) and (59). I estimate a continuation probit at each tenure level with random person- and firm-specific mobility effects.[29] The probits are estimated by Average Information REML applied to the method of Schall (1991).[30] With estimates of the realized random effects $\tilde{\zeta}_{it}$ and $\tilde{\xi}_{j\tau}$ in hand, I construct an estimate $\tilde{\lambda}_{ij\tau}$ of the Inverse Mills' Ratio and include it as an additional covariate in the earnings equation. Note the bargaining strength parameter γ is not identified.

3.4 The Learning Hypothesis

I now turn to a testable hypothesis of the matching model. In the theoretical model, agents update their beliefs about match quality using Bayes' Rule. Bayesian learning implies a specific structure for the within-match error covariance W. Specifically, given the compound error structure $\varepsilon_{ijt} = \gamma m_{ij\tau} + u_{ijt}$, we have $W = \gamma^2 V + \sigma_u^2 I_{\bar{\tau}}$, where V is the $\bar{\tau} \times \bar{\tau}$ within-match covariance of the vector of belief terms $[m_{ij1} \cdots m_{ij\bar{\tau}}]$, and σ_u^2 is the variance of measurement error.[31] Learning about match quality therefore implies

$$W = \begin{bmatrix} \gamma^2 V_1 + \sigma_u^2 & \gamma^2 V_1 & \gamma^2 V_1 & \cdots & \gamma^2 V_1 \\ \gamma^2 V_1 & \gamma^2 V_2 + \sigma_u^2 & \gamma^2 V_2 & \cdots & \gamma^2 V_2 \\ \gamma^2 V_1 & \gamma^2 V_2 & \gamma^2 V_3 + \sigma_u^2 & \cdots & \gamma^2 V_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma^2 V_1 & \gamma^2 V_2 & \gamma^2 V_3 & \cdots & \gamma^2 V_{\bar{\tau}} + \sigma_u^2 \end{bmatrix} \tag{60}$$

where V_τ was defined in (13).

Some aspects of (60) are worthy of note. First, off-diagonal elements within each column of the lower triangle are equal. This is a property of the covariance of any martingale process, and is quite intuitive. Elements of column τ are $Cov(m_{ij\tau}, m_{ij\tau'})$. In the lower triangle, $\tau \leq \tau'$. The common elements in $m_{ij\tau}$ and $m_{ij\tau'}$ are the signals of match quality

[29] Identification of the random effects required pooling of some probit equations across tenure levels.

[30] Schall (1991) extends standard estimation methods for generalized linear models to the random effects case. It is based on REML estimation of a linearized link function (in this case, Φ).

[31] In the unbalanced data case, the vector of belief terms is $[m_{ij1} \cdots m_{ij\tau_{ij}}]$, and $W_{ij} = \gamma^2 S'_{ij} V S_{ij} + \sigma_u^2 I_{\tau_{ij}}$.

received up to tenure τ. Thus the covariance between $m_{ij\tau}$ and $m_{ij\tau'}$ is the variance of the signals received up to tenure τ, i.e., V_τ. A second aspect of note is that the diagonal elements of W are greater than off-diagonal elements in the same column due to measurement error. Finally, as discussed in Section 2.4, $V_{\tau+1} > V_\tau$.[32] Consequently in the lower triangle of (60), elements in each row increase in magnitude from left to right.

Whether the errors have the structure (60) is testable. Furthermore, since the structural parameters σ_c^2, σ_z^2, and σ_e^2 enter into each V_τ, they can be recovered from an estimate of the within-match error covariance (up to the factor of proportionality γ^2). I test the learning hypothesis and recover the structural parameters and σ_u^2 in two steps. The first step is to estimate W. For the mixed model, I use the unrestricted REML estimate of the within-match error covariance.[33] For the fixed model, I use the within-match sample covariance of the residuals.[34] Following Abowd and Card (1989) and Farber and Gibbons (1996), in the second step I fit the martingale covariance $\gamma^2 V + \sigma_u^2 I_{\bar{\tau}}$ to the estimate of W by minimum distance.[35] This yields estimates of the structural parameters up to γ^2. I test the learning hypothesis with the usual χ^2 test of overidentifying restrictions, using the test statistic of Newey (1985).[36]

4 Data

Identifying the person and firm effects requires repeated observations on both workers and firms: longitudinal linked data on employers and employees. I use data from the US Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) program database. The LEHD database spans thirty-two states that represent the majority of American employment. In this paper, I use data from two participating states. Their identity is confidential.

The LEHD data are administrative, constructed from quarterly Unemployment Insurance (UI) system wage reports. Every state, through its Employment Security agency, collects quarterly earnings and employment information to manage its unemployment compensation program. The characteristics of the UI wage data vary slightly from state to state. However 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. Further details regarding UI wage records and the LEHD data can be found in Stevens (2002), Abowd et al. (2000), and LEHD Program (2002). With the UI wage records as its frame, the LEHD data comprise the universe of employers required to

[32] The prior variance of $m_{ij\tau}$ is zero: all agents have common priors about match quality. As agents acquire information, the posterior mean of beliefs converges to true match quality. V_τ increases from its prior value (zero) to its asymptotic value (σ_c^2) as tenure increases.

[33] The relevant block of the REML Average Information matrix estimates the covariance of W.

[34] In this specification, I estimate the covariance of W using the fourth moments of the residuals.

[35] Optimal minimum distance estimation, as discussed in Hansen (1982) and Chamberlain (1984), proved infeasible. For all specifications, the estimated covariance of W was poorly conditioned, and did not invert. I use a diagonal weight matrix instead, with elements equal to the natural logarithm of the number of observations contributing to each element of W. The data are highly unbalanced, with many more observations contributing to elements in the upper-left corner of W than to elements of the lower-right corner. This weight matrix gives greater weight to more precisely estimated elements of W. Weighting by a diagonal matrix of variances yields similar results, as does equally weighted minimum distance.

[36] The Newey (1985) test statistic does not require inverting the variance of the moment conditions.

file UI system wage reports — that is, all employment potentially covered by the UI system in participating states.

Individuals are uniquely identified in the data by a Protected Identity Key (PIK). Employers are identified by an unemployment insurance account number (SEIN). The UI wage records themselves contain only very limited information: PIK, SEIN, and earnings. Earnings are a measure of total compensation that includes 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 LEHD database integrates the UI wage records with internal Census Bureau data to obtain additional demographic and firm characteristics, including sex, race, date of birth, industry, and geography.

Though the underlying data are quarterly, they are aggregated to the annual level for estimation. The full sample consists of over 49 million annualized employment records on full-time workers between 25 and 65 years of age who are employed at nearly 575,000 private-sector non-agricultural firms between 1990 and 1999.

Missing data items are multiply-imputed. Three imputed values are generated for each missing data item. The result is three versions of each of the samples, that I call completed data implicates, each of which contains a different set of imputed values. Further details on sample construction, variable creation, and missing data imputation are given in the Data Appendix.

The direct least squares estimator of ACK is very computationally efficient. It is therefore possible to estimate the fixed model on the full sample. Computational demands of the mixed model, however, necessitate estimating it on a subsample of observations. Sampling from these data is nontrivial because the sample must be sufficiently connected to precisely estimate the person and firm effects.[37] I therefore draw a one percent subsample using the dense sampling algorithm of Woodcock (2005). This sampling procedure ensures that each worker is connected to at least five others by a common employer, but is otherwise equivalent to a simple random sample of individuals employed in a reference year (1997). That is, all individuals employed in 1997 have an equal probability of being sampled. A second (disjoint) one percent dense random subsample is drawn for model validation. In what follows, the subsamples are labeled Dense Sample 1 and Dense Sample 2, respectively. More details on the dense sampling algorithm can be found in the Data Appendix. All mixed model estimates are computed on Dense Sample 1. Tables 1 and 2 report characteristics of the samples. These are discussed further in the Data Appendix.

Estimating the within-match error covariance W requires a measure of job tenure. All employment spells active in the first quarter for which data are available are presumed left-censored. Since job tenure is unknown for left-censored spells, they are excluded from the samples for the purpose of estimating W, and hence for testing the learning hypothesis.[38] Those estimates that do not require a measure of job tenure are based on samples that include the left-censored spells.

[37] Identifying the effects requires multiple observations on employees of most firms and mobility of workers across firms. A small simple random sample of individuals, for example, is usually not sufficiently connected to estimate the person and firm effects precisely.

[38] Alternate results based on multiply-imputed tenure for the left-censored spells are available on request.

5 Results

The fixed and mixed models described in Section 3 are estimated on each of the completed data implicates. Statistics reported in this section combine those obtained on the three implicates using standard formulae in Rubin (1987). Estimates of the fixed covariate effects β and η are available from the author upon request. There is little variation in the estimated covariate effects across specifications. Their values are reasonable.

5.1 Parameter Estimates and Model Fit

Tables 3 and 4 present estimates of the variance components and a summary of model fit for the three specifications of interest: the fixed model, the mixed model with spherical errors ($R = \sigma_\varepsilon^2 I_{N*}$), and the mixed model with unrestricted within-match error covariance W. The estimates reported in Table 3 include left-censored spells,[39] while those in Table 4 exclude them. The two sets of estimates are virtually identical. Both tables report mixed model estimates with and without the truncation correction.[40]

The estimated variance components have a straightforward interpretation. Conditional on all other effects, a one standard deviation increase in the value of the person effect increases real annualized earnings by σ_α log points. Similarly, a one standard deviation increase in the value of the firm effect increases real annualized earnings by σ_ψ log points.

The first thing to note in Tables 3 and 4 is that the truncation correction induces virtually no change in the estimated variance components. This is not overly surprising given the small point estimates of β_λ (between zero and 0.04, depending on specification). It is quite common to find small selection/truncation biases in earnings data.

The second item of note is that in all specifications, the variance of the person effect (σ_α^2) is considerably larger than the variance of the firm effect (σ_ψ^2). That is, individual heterogeneity generates more earnings dispersion than firm heterogeneity. This is consistent with the findings of AKM and others. In Tables 3 and 4, the fixed model yields the largest estimate of σ_α^2 (0.29), but among the smallest estimates of σ_ψ^2 (0.08). These values are comparable to those estimated by ACK for France and the State of Washington, though smaller than Abowd et al. (2003) find in a broader sample of seven LEHD states.[41] The mixed model with spherical errors yields a slightly smaller estimate of σ_α^2 (about 0.23), and an estimate of σ_ψ^2 almost twice that obtained in the fixed model. Relaxing the spherical errors assumption in favor of the unrestricted within-match error covariance reduces the estimate of σ_ψ^2 to a level comparable to the fixed model, and reduces the estimate of σ_α^2 to 0.18. For this specification, a one standard deviation increase in the value of the person effect increases earnings by 0.42 log points, and a one standard deviation increase in the value of the firm

[39] Table 3 estimates of the mixed model with W unrestricted are based on multiply-imputed tenure for left-censored spells.

[40] Complete estimation results for the probit equations are available on request. I have not applied the truncation correction to the fixed model. Doing so is technically possible though computationally demanding: it requires estimating about 25 million probit person and firm effects.

[41] The ACK estimates for Washington State are $\sigma_\theta^2 = 0.21$ and $\sigma_\psi^2 = 0.09$. For France they estimate $\sigma_\theta^2 = 0.23$ and $\sigma_\psi^2 = 0.05$. Abowd et al. (2003) estimate $\sigma_\alpha^2 = 0.64$ and $\sigma_\psi^2 = 0.13$. The discrepancy is at least partly due to different identifying assumptions for state means.

effect increases earnings by 0.28 log points. Clearly, unobserved individual heterogeneity contributes more to earnings variation than unobserved firm heterogeneity. However both are very economically significant.

Tables 3 and 4 also report some measures of model fit. The mixed model with spherical errors obtains the best fit by in-sample measures (AIC, BIC), largely due to its parsimony. To obtain a measure of out-of-sample fit for the mixed models, I solve the mixed model equations (53) on Dense Sample 2, using the variance components $\tilde{G}$ and error covariance $\tilde{R}$ estimated on Dense Sample 1. I compute the prediction error for each observation, and report its variance.[42] By this measure too, the mixed model with spherical errors obtains the best fit.

It is standard to test random effects specifications against the fixed effects alternative using the Hausman (1978) test. Unfortunately, in these data the test proved inconclusive for both mixed effects specifications: the test statistics were negative.

Table 5 presents correlations among the estimated effects. They are qualitatively similar across specifications. Of the estimated effects, the person effect θ_i is most highly correlated with log earnings: between 0.74 and 0.82, depending on specification. In each case, the portion of θ_i corresponding to unobserved heterogeneity (α_i) is much more highly correlated with earnings than the observable component ($u_i\eta$). This is at least partly due to the relatively limited set of observable characteristics in the LEHD data.[43] Correlations between the firm effect and log earnings are lower: between 0.45 and 0.54. Time-varying covariates are even less correlated with earnings, between 0.18 and 0.3 depending on specification.

Recall that the matching model predicts a negative correlation between θ_i and ψ_j. In the fixed model and the mixed model with spherical errors, the correlation between θ_i and ψ_j is essentially zero (0.03). ACK obtain a similarly small, though negative (-0.02), correlation in Washington State data. Abowd et al. (2003), using a sample of seven states from the LEHD data, estimate the correlation to be 0.08. Thus estimates in the first two panels of Table 5 are consistent with other US evidence.[44]

Andrews et al. (2004b) argue that least squares estimates of the correlation between θ_i and ψ_j may be biased downward. I find some corroborating evidence in Table 5. When the mixed model is relaxed to allow an unrestricted within-match error covariance, the correlation between θ_i and ψ_j increases significantly to 0.22. Interestingly, a the same result is obtained with a variety of alternate specifications of the within-match error covariance (not reported here, but available on request), including a variety of ARMA structures or a random "match effect." This suggests that the downward bias noted by Andrews et al. (2004b) may not be a consequence of the fixed effects estimator *per se*, but due to mis-specification of the error distribution. Further research is needed to clarify this issue.

The estimated correlation between θ_i and ψ_j appears to contradict the theoretical model. However Figure 1, which plots the estimated cubic regression of θ_i on ψ_j for each of the specifications, suggests that focusing solely on correlations may be misleading. In each case,

[42] For models with $R = \sigma_\varepsilon^2 I_{N^*}$, the prediction error associated with an observation is the estimated residual. For the model with W unrestricted, the prediction error is the difference between the estimated residual and its conditional expectation given the other within-match residuals under multivariate normality.

[43] Time-invariant covariates include education, race, and a dummy variable to indicate whether initial potential experience is negative, all interacted with sex.

[44] In contrast, European studies typically find a substantial negative correlation.

the relationship between person and firm effects is highly nonlinear. There is a systematic, though weak, positive association between the person and firm effects in the neighborhood of zero. Given the estimated correlations reported in Table 5, it is not surprising that the positive association is strongest in the mixed model with W unrestricted. Among more extreme values of the person and firm effects, however, there is strong "mismatch" of the type predicted by the matching model.[45] This result is common to all three empirical specifications and explains, in part, the near-zero correlations between person and firm effects in Table 5.

The two bottom panels of Table 5 present correlations among the estimated effects after correcting for truncation. The truncation correction has no noticeable effect on the correlations between earnings, the person and firm effects, and covariates. The correction term itself ($\beta_\lambda \lambda_{ij\tau}$) exhibits correlations consistent with the matching model. In the context of the theoretical model, $\beta_\lambda \lambda_{ij\tau}$ estimates the expected value of m_τ given that it exceeds $\bar{m}_\tau$. Since $E[m_\tau | m_\tau > \bar{m}_\tau]$ is increasing in $\bar{m}_\tau$, and recalling the comparative static result (38), the matching model predicts $\partial \beta_\lambda \lambda_{ij\tau}/\partial\theta < 0$ and $\partial \beta_\lambda \lambda_{ij\tau}/\partial\psi < 0$. This prediction is supported by the data, since the estimated correlation between the correction term and the person and firm effects is negative in both mixed model specifications. As for the correlation between $\beta_\lambda \lambda_{ij\tau}$ and earnings, the predicted sign is ambiguous. On the one hand, because $\beta_\lambda \lambda_{ij\tau}$ estimates $E[m_\tau | m_\tau > \bar{m}_\tau]$, larger values of the correction term will be associated with larger values of m_τ, and hence of earnings also. On the other hand, larger values of $\beta_\lambda \lambda_{ij\tau}$ will be associated with smaller values of θ_i and ψ_j, and hence lower earnings. In both mixed model specifications, the estimated correlation between $\beta_\lambda \lambda_{ij\tau}$ and y is negative, implying the latter effect dominates.

5.2 Testing the Learning Hypothesis

Estimates of the within-match error covariance W are presented in Tables 6 and 7. Table 6 presents fixed model estimates based on the full sample (excluding left-censored spells). They exhibit some of the properties of the martingale covariance structure given in (60): in each column, the diagonal elements are larger than the off-diagonal elements, and elements increase in magnitude from left to right within each row. However, the martingale structure also implies that off-diagonal elements within a column should be equal. They are clearly not. Moving from lower-order to higher-order autocovariances, the elements in Table 6 consistently decline in magnitude.

Table 7 presents mixed model estimates of W with and without correcting for truncation. They are virtually identical. Casual inspection indicates the estimates are consistent with the structure implied by the learning process: diagonal elements are larger than off-diagonal elements within a column and elements increase in magnitude from left to right within each row. Compared to the fixed model, there is far less decay in the autocovariances. This is

[45] A scatter plot of the estimated person and firm effects would be more illustrative of mismatch. However, confidentiality concerns preclude presenting one here. Suffice to say, the plotted regression functions are representative of the distribution of the data. The axes are scaled to approximate the observed maximum and minimum values of θ_i and ψ_j. The majority of the data lie within the unit circle, where the association between person and firm effects is weakest. However, a substantial proportion of the data (>10% of observations) lie outside this region. Almost all such observations lie in the northwest and southeast quadrants of Figure 1, which is consistent with mismatch.

more in line with the martingale structure than estimates in Table 6.

Results of the formal test of the learning hypothesis are given in Table 8. I fit (13) and (60) to the fixed and mixed model estimates of W by minimum distance. Table 8 presents the estimated structural parameters and p-values from the chi-squared test of over-identifying restrictions.[46] Recall that the structural parameters σ_c^2, σ_z^2, and σ_e^2 are only identified up to a factor of proportionality: the square of the bargaining strength parameter γ. Estimates in Table 8 are presented on the scale of the data, i.e., for $\gamma = 1$.[47]

The first set of columns in Table 8 present results for the fixed model. The estimated variance of measurement error (σ_u^2) is 0.05, which is reasonable and slightly smaller than the variance of the firm effect. The estimated variance of match quality (σ_c^2) is small (0.01) in comparison to the variance of the person and firm effects. In contrast, the variance of the initial signal of match quality (σ_z^2) is very large (0.97), implying this signal conveys little information. The variance of production outcomes (σ_e^2) is also large (0.07) in comparison to the variance of match quality, which suggests agents learn little about match quality from output realizations. This is confirmed in Figure 2, which plots the variance of beliefs about match quality (s_τ^2) at various tenure levels. It declines very slowly from its prior value ($\sigma_c^2 = 0.01$) to around 0.004 over the course of ten years. Agents learn do not learn much about match quality because there is little to learn. It is therefore no surprise that the over-identifying restrictions implied by the matching model are rejected for the fixed effect specification.

Results for the mixed model are quite different. These are presented in the second set of columns in Table 8. For the most part, results are very similar with and without correcting for truncation. The estimated variance of measurement error is reasonable (0.04) and similar to the fixed model estimate. The estimated variance of match quality is about 0.09, larger than the variance of the firm effect but only about half the variance of the person effect. Unlike the fixed model, this estimate implies substantial wage dispersion due to match quality: a one standard deviation increase in match quality increases earnings by 0.3 log points. The variance of the initial signal is 0.10, and production outcomes have a variance of 0.15. Correcting for truncation reduces these estimates to 0.09 and 0.11, respectively. These estimates imply learning about match quality is quite rapid. This is illustrated in

[46] Rubin (1987) provides formulae for combining chi-square distributed statistics estimated on multiply-imputed data. Let d_m denote the test statistic from the m^{th} implicate, asymptotically distributed χ_k^2. Let M denote the number of implicates, $\bar{d}_m$ the sample mean of the statistics d_m, and s_d^2 their sample variance. Define

$$\hat{r}_m = \frac{\left(1 + M^{-1}\right) s_d^2}{2\bar{d}_m + \left(4\bar{d}_m^2 - 2ks_d^2\right)^{1/2}} \tag{61}$$

and

$$\hat{v} = (M-1)\left(1 + \hat{r}_m^{-1}\right)^2. \tag{62}$$

The quantity $\hat{r}_m$ is a method of moments estimator of the relative increase in variance of the test statistic due to nonresponse. The test statistic

$$\hat{D}_m = \frac{\frac{\bar{d}_m}{k} - \frac{M-1}{M+1}\hat{r}_m}{1 + \hat{r}_m} \tag{63}$$

has an asymptotic F distribution with k and $\left(1 + k^{-1}\right)\hat{v}/2$ degrees of freedom. Reported p-values are based on $\hat{D}_m$.

[47] They can be re-scaled for any other $0 < \gamma < 1$ quite easily: the re-scaled parameter is $\sigma_*^2 = \sigma^2/\gamma^2$.

Figure 2. The initial signal resolves nearly half of the uncertainty about match quality: the variance of beliefs drops from its prior value (0.09) to about 0.047. After observing one production outcome, it falls further to 0.035. It falls to 0.02 by the fifth year, and to 0.01 by the tenth year. Thus most learning about match quality occurs early in the employment relationship. As for the test of over-identifying restrictions, we reject the learning hypothesis at the 5% level of significance but cannot do so at the 1% level. Given the large sample size, it is perhaps surprising that we do not reject the null at *all* conventional significance levels. Being unable to reject the learning hypothesis with a high degree of confidence lends some support to the learning hypothesis, though it is far from overwhelming.

5.3 Additional Predictions From the Matching Model

Figures 3 through 5 address two additional predictions of the matching model: that larger values of θ_i and ψ_j should on average be associated with longer job duration, and that the firm effect should be positively associated with firm size. To address the first of these predictions, I fit a fourth-order polynomial in completed job duration to the estimated person and firm effects.[48] I focus on results for the mixed model with W unrestricted. Results from the other specifications are very similar and available on request. Figures 3 and 4 present the fitted curves. As predicted, larger values of θ_i and ψ_j are associated with longer duration. The profile is much steeper for the person effect than for the firm effect. This is consistent with the much greater variation in θ_i than ψ_j.

To address the second prediction, I fit a fourth-order polynomial in the natural logarithm of the firm's 1997 employment to the estimated firm effect.[49] Again, I focus on the mixed model with W unrestricted. Results obtained on the other specifications and for employment in other years are very similar. Figure 5 presents the fitted curves. As predicted by the matching model, larger values of ψ_j are associated with larger employment. The relationship is nearly linear for small and medium-size firms, and quite convex among the largest firms.

6 Conclusion

The matching model presented in Section 2 predicts rich dispersion in equilibrium wages and employment dynamics. Productivity differences across individuals, technological differences between firms, and learning about match quality all contribute. Given our assumptions about the production technology and wage bargaining, the equilibrium wage function takes the same additively-separable form as the AKM empirical specification. This provides a formal economic interpretation of the AKM person and firm effects. They reflect productivity differences across workers and firms, adjusted for bargaining strength and discounting. Consistent with most previous empirical studies, the model predicts a negative correlation between estimated person and firm effects. The model also makes new predictions about the

[48] Left-censored spells are excluded. Alternate results based on multiply-imputed tenure for left-censored spells are available on request. They are virtually identical.

[49] 1997 was chosen because the Economic Census was conducted in that year, and because the dense samples use 1997 as a reference year.

equilibrium relationship between the person and firm effects and separation behavior, job duration, and firm size.

It is well known by now that person and firm effect models explain (in a statistical sense) the vast majority of wage dispersion. In the LEHD data considered here, our specification of the equilibrium wage function explains about 90 percent of variation in earnings.[50] The empirical results suggests that productivity differences across individuals contribute about twice as much to wage dispersion as do differences across firms. The matching model and our empirical specification attribute the remaining dispersion to learning about match quality and measurement error. Estimates of the structural parameters imply learning about match quality contributes about as much to wage dispersion as firm-level heterogeneity does.

Formal and informal tests of the matching model's empirical predictions yield somewhat mixed results. On the one hand, correlations between the truncation correction term and the person and firm effects have the predicted sign. Likewise, the relationship between the estimated person and firm effects and job duration, and between the firm effect and firm size, are as predicted. We are also unable to confidently reject the learning hypothesis. On the other hand, the truncation correction has surprisingly little effect on estimates. Likewise, the estimated correlation between the person and firm effects has the wrong sign – though most other studies, particularly those on European data, have found correlations in line with the model's predictions. Even in the LEHD data, close inspection of the distribution of person and firm effects reveals some evidence of mismatch between workers and firms, as predicted by the model. Overall, the data provide considerable, though not overwhelming, support for the model.

This paper identifies several fruitful avenues for future research. Investigating the importance of explicit human capital investments, and of the firm's choice of technology, would be considerable contributions. There is also room to consider alternate forms of endogenous mobility, such as on-the-job search. On the empirical front, we found evidence of biases arising from mis-specifying the error distribution. This deserves further investigation.

[50] R^2 in the fixed model is 0.87. There is no exactly comparable measure of variation explained by the mixed models. An approximate measure is given by R^2 in the auxiliary regression of the estimated person, firm, and covariate effects on log earnings. Using this measure, R^2 for the mixed model with spherical errors is 0.91, and $R^2 = 0.88$ for the mixed model with W unrestricted.

A Appendicized Proofs

Proof of Proposition 2. Define the operator T by

$$(TW)(m_\tau) = \max\left\{\mu + a + b + m_\tau + \beta(1-\delta)\int W(m_{\tau+1})\, dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right) + \beta\delta U, U\right\}$$

where $F(m_{\tau+1}|m_\tau, s_\tau^2)$ denotes the normal distribution with mean m_τ and variance $v_{\tau+1}$ defined in equations (14) and (15), and where the second state variable s_τ^2 has been suppressed to simplify the notation. Let S denote the space of bounded, continuous, nondecreasing, convex functions. We first show $T : S \to S$. Let $q(m_\tau) = \mu + a + b + m_\tau$. The boundedness assumption implies $q(m_\tau)$ is bounded, and it is obviously continuous, increasing, and convex. Since U is a constant, it is sufficient to show that the operator M defined by

$$(MW)(m_\tau) = \int W(m_{\tau+1})\, dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right)$$

maps S into itself. So let $W \in S$. Then MW is bounded and continuous. To see that it is nondecreasing, let $m'_\tau > m_\tau$. Then $F(m_{\tau+1}|m'_\tau, \sigma^2) \leq F(m_{\tau+1}|m_\tau, \sigma^2)$ for every $\sigma^2 > 0$. That is, $F(m_{\tau+1}|m'_\tau, \sigma^2)$ first-order stochastically dominates $F(m_{\tau+1}|m_\tau, \sigma^2)$, so that

$$(MW)(m'_\tau) = \int W(m_{\tau+1})\, dF\left(m_{\tau+1}|m'_\tau, \sigma^2\right) \geq \int W(m_{\tau+1})\, dF\left(m_{\tau+1}|m_\tau, \sigma^2\right) = (MW)(m_\tau)$$

since W is nondecreasing by hypothesis. As for convexity, since $m_{\tau+1} \sim N(m_\tau, v_{\tau+1})$ we can write $m_{\tau+1} = m_\tau + \varphi$ where $\varphi \sim N(0, v_{\tau+1})$. Then rewrite the operator M as

$$(MW)(m_\tau) = \int W(m_{\tau+1})\, dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right) = \int W(m_\tau + \varphi)\, dF\left(\varphi|0, s_\tau^2\right)$$

where $F(\varphi|0, s_\tau^2)$ is the normal distribution with mean zero and variance $v_{\tau+1}$. Then for any m_τ, m'_τ, and $\lambda \in [0,1]$,

$$\begin{aligned}
\lambda(MW)(m_\tau) + (1-\lambda)(MW)(m'_\tau) &= \lambda\int W(m_\tau + \varphi)\, dF\left(\varphi|0, s_\tau^2\right) \\
&\quad + (1-\lambda)\int W(m'_\tau + \varphi)\, dF\left(\varphi|0, s_\tau^2\right) \\
&= \int \left[\lambda W(m_\tau + \varphi) + (1-\lambda) W(m'_\tau + \varphi)\right] dF\left(\varphi|0, s_\tau^2\right) \\
&\geq \int W\left[\lambda(m_\tau + \varphi) + (1-\lambda)(m'_\tau + \varphi)\right] dF\left(\varphi|0, s_\tau^2\right) \\
&= \int W\left[\lambda m_\tau + (1-\lambda) m'_\tau + \varphi\right] dF\left(\varphi|0, s_\tau^2\right) \\
&= (MW)\left(\lambda m_\tau + (1-\lambda) m'_\tau\right)
\end{aligned}$$

where the inequality follows because W is convex. Thus $M : S \to S$ and hence $T : S \to S$ also.

To show there is a unique $W \in S$ that satisfies the Bellman equation (33) we need only establish that T is a contraction. Uniqueness then follows immediately from the Contraction Mapping Theorem, since S and the sup norm define a complete metric space. To show that T is a contraction, we verify the Blackwell (1965) sufficient conditions.

Monotonicity: Let $\Psi, W \in S$ and $\Psi(x) \leq W(x)$ for all x. Then

$$\begin{aligned}(T\Psi)(m_\tau) &= \max\left\{q(m_\tau) + \beta(1-\delta)\int \Psi(m_{\tau+1})\,dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right) + \beta\delta U, U\right\} \\ &\leq \max\left\{q(m_\tau) + \beta(1-\delta)\int W(m_{\tau+1})\,dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right) + \beta\delta U, U\right\} \\ &= (TW)(m_\tau)\end{aligned}$$

as required.

Discounting: Let $W \in S$, $y \geq 0$, and $\beta \in (0,1)$. Then

$$\begin{aligned}[T(W+y)](m_\tau) &= \max\left\{q(m_\tau) + \beta(1-\delta)\int [W(m_{\tau+1}) + y]\,dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right) + \beta\delta U, U\right\} \\ &= \max\left\{q(m_\tau) + \beta(1-\delta)\int W(m_{\tau+1})\,dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right) + \beta\delta U + \beta(1-\delta)y, U\right\} \\ &\leq \max\left\{q(m_\tau) + \beta\int W(m_{\tau+1})\,dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right), U\right\} + \beta y \\ &= (TW)(m_\tau) + \beta y\end{aligned}$$

as required. ■

Proof of Proposition 3. The proof of Proposition 2 showed that when W is continuous and nondecreasing in m_τ, so is $\int W\left(m_{\tau+1}, s_{\tau+1}^2\right) dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right)$. Therefore the first argument of the max operator in the Bellman equation (33) is continuous and increasing in m_τ. The second argument is a constant, and the result follows immediately. ■

Proof of Proposition 4. We first establish an intermediate result. Namely, that

$$\varrho\left(m_\tau, s_\tau^2\right) = \bar{E}_\tau\left[W\left(m_{\tau+1}, s_{\tau+1}^2\right)\right] = \int W\left(m_{\tau+1}, s_{\tau+1}^2\right) dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right)$$

is increasing in s_τ^2. We already know from the proof of Proposition 2 that ϱ is nondecreasing in m_τ. Recall that $F\left(m_{\tau+1}|m_\tau, s_\tau^2\right)$ is the normal distribution with mean m_τ and variance $v_{\tau+1}$ given by equation 14. Notice that

$$\frac{\partial v_{\tau+1}}{\partial s_\tau^2} = s_\tau^2\left(1 + \frac{\sigma_e^2}{s_\tau^2 + \sigma_e^2}\right) > 0$$

so an increase in s_τ^2 constitutes a mean-preserving spread on $m_{\tau+1}$. Since W is convex in its first argument, for any $\tilde{s}_\tau^2 > s_\tau^2$ we have

$$\begin{aligned}\varrho\left(m_\tau, \tilde{s}_\tau^2\right) &= \int W\left(m_{\tau+1}, s_{\tau+1}^2\right) dF\left(m_{\tau+1}|m_\tau, \tilde{s}_\tau^2\right) \\ &\geq \int W\left(m_{\tau+1}, s_{\tau+1}^2\right) dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right) = \varrho\left(m_\tau, s_\tau^2\right)\end{aligned}$$

as required.

As for the main result, suppose to the contrary that $\bar{m}_{\tau+1} < \bar{m}_\tau$. Then from (34),

$$\begin{aligned}\bar{m}_{\tau+1} - \bar{m}_\tau &= \beta(1-\delta)\bar{E}_\tau\left[W\left(m_{\tau+1}, s^2_{\tau+1}\right)\right] - \beta(1-\delta)\bar{E}_{\tau+1}\left[W\left(m_{\tau+2}, s^2_{\tau+1}\right)\right] \\ &= \beta(1-\delta)\left(\varrho\left(\bar{m}_\tau, s^2_\tau\right) - \varrho\left(\bar{m}_{\tau+1}, s^2_{\tau+1}\right)\right).\end{aligned}$$

The right-hand side is negative because ϱ is nondecreasing in its first argument and $\bar{m}_{\tau+1} < \bar{m}_\tau$ by hypothesis, and because ϱ is increasing in its second argument and $s^2_{\tau+1} < s^2_\tau$ for all $\tau > 0$. But the left-hand side is negative, a contradiction. ■

The following lemmata are useful for the proof of Proposition 5.

Lemma 6 $\partial U/\partial b = 0, \partial U/\partial a \in \left(0, \frac{1}{1-\beta}\right)$.

Proof. Write the value of the worker's outside option as:

$$U = h + \beta\pi\int_{b_0}^{b_1} J_0 dF^*_b + \beta(1-\pi)U = \frac{h + \beta\pi\int_{b_0}^{b_1} J_0 dF^*_b}{1-\beta(1-\pi)} \tag{64}$$

where

$$J_0 = E_0\left[\max\{J_1, U\}\right] = U + \int_{\bar{m}_1}^{\infty}(J_1 - U)\,dF\left(m_1|0, \sigma^2_c + \sigma^2_z\right) \tag{65}$$

is the prior expected value of employment defined in (22) and F^*_b is defined Appendix B. That $\partial U/\partial b = 0$ is obvious, since U doesn't depend on b.

From (64) we have

$$\frac{\partial U}{\partial a} = \frac{\beta\pi}{1-\beta(1-\pi)}\int_{b_0}^{b_1}\frac{\partial J_0}{\partial a}dF^*_b. \tag{66}$$

Differentiating (65) using Leibniz's Rule,

$$\begin{aligned}\frac{\partial J_0}{\partial a} &= \frac{\partial U}{\partial a} - \frac{\partial \bar{m}_1}{\partial a}\left(\bar{J}_1 - U\right)f\left(\bar{m}_1|0, \sigma^2_c + \sigma^2_z\right) + \int_{\bar{m}_1}^{\infty}\frac{\partial(J_1 - U)}{\partial a}dF_1 \\ &= \frac{\partial U}{\partial a} + \int_{\bar{m}_1}^{\infty}\frac{\partial(J_1 - U)}{\partial a}dF_1\end{aligned} \tag{67}$$

where $F_{\tau+1}$ is shorthand for $F\left(m_{\tau+1}|m_\tau, s^2_\tau\right)$ and $\bar{J}_\tau$ is shorthand for the value of J_τ when $m_\tau = \bar{m}_\tau$. Note $\bar{J}_\tau = U$ by definition of $\bar{m}_\tau$ and the individual rationality property of the Nash Bargain. Differentiating (20) gives the recursion

$$\frac{\partial(J_\tau - U)}{\partial a} = \gamma - \gamma(1-\beta)\frac{\partial U}{\partial a} + \beta(1-\delta)\int_{\bar{m}_{\tau+1}}^{\infty}\frac{\partial(J_{\tau+1} - U)}{\partial a}dF_{\tau+1} \tag{68}$$

for all $\tau > 0$. Substituting (68) into (67) repeatedly gives

$$\frac{\partial J_0}{\partial a} = \frac{\partial U}{\partial a}\left(1 - \gamma(1-\beta)Z\right) + \gamma Z \tag{69}$$

where

$$Z=\sum_{\tau=1}^{\infty}\left[\beta\left(1-\delta\right)\right]^{\tau-1}\int_{\bar{m}_1}^{\infty}\int_{\bar{m}_2}^{\infty}\cdots\int_{\bar{m}_\tau}^{\infty}dF_\tau\cdots dF_2dF_1>0. \tag{70}$$

Substituting (69) into (66) and simplifying gives

$$\frac{\partial U}{\partial a}=\frac{1}{1-\beta}\left[\frac{\beta\pi\gamma\int_{b_0}^{b_1}ZdF_b^*}{1+\beta\pi\gamma\int_{b_0}^{b_1}ZdF_b^*}\right]\in\left(0,\frac{1}{1-\beta}\right) \tag{71}$$

because $Z>0$ implies the term in square brackets is between zero and one, and $\beta\in(0,1)$. ■

Lemma 7 *The joint value of continuing the employment relationship, $J_\tau+\Pi_\tau$, is strictly increasing in a and b.*

Proof. Since $J_\tau+\Pi_\tau=\mu+a+b+m_\tau+\beta\left(1-\delta\right)\int W\left(m_{\tau+1},s_{\tau+1}^2\right)dF\left(m_{\tau+1}|m_\tau,s_\tau^2\right)+\beta\delta U$, and given lemma 6, it is sufficient to show the value function W is nondecreasing in a and b. Recall the space of functions S from the proof of Proposition 2, and define the space of functions $S'\subseteq S$ that are also nondecreasing in a and b. Recall also the operators T and M defined in the proof of Proposition 2. Write the value function as $W\left(m_\tau;a,b\right)$. If $W\in S'$, then for any $a'>a$, $b'>b$, and any m_τ, we have $W\left(m_\tau;a',b'\right)\geq W\left(m_\tau;a,b\right)$ so that

$$\begin{aligned}\left(MW\right)\left(m_\tau;a',b'\right)&=\int W\left(m_{\tau+1};a',b'\right)dF\left(m_{\tau+1}|m_\tau,s_\tau^2\right)\\&\geq\int W\left(m_{\tau+1};a,b\right)dF\left(m_{\tau+1}|m_\tau,s_\tau^2\right)\\&=\left(MW\right)\left(m_\tau;a,b\right)\end{aligned}$$

which shows that $M:S'\rightarrow S'$. Applying lemma 6 again, it follows that $\left(TW\right)\left(m_\tau;a',b'\right)\geq\left(TW\right)\left(m_\tau;a,b\right)$ also. Hence $T:S'\rightarrow S'$. Since S' is a closed subset of S, the unique fixed point of T is $W\in S'$. ■

Proof of Proposition 5. Rewrite the threshold value of beliefs as

$$\bar{m}_\tau=\left(1-\beta\right)U-\mu-a-b-\beta\left(1-\delta\right)\int_{\bar{m}_{\tau+1}}^{\infty}\left(J_{\tau+1}+\Pi_{\tau+1}-U\right)d\bar{F}_{\tau+1} \tag{72}$$

where $\bar{F}_{\tau+1}=F\left(m_{\tau+1}|\bar{m}_\tau,s_\tau^2\right)$. Let $x\in\{a,b\}$. Differentiating (72) using Leibniz's Rule,

$$\begin{aligned}\frac{\partial\bar{m}_\tau}{\partial x}&=\left(1-\beta\right)\frac{\partial U}{\partial x}-1+\beta\left(1-\delta\right)\frac{\partial\bar{m}_{\tau+1}}{\partial x}\left(\bar{J}_{\tau+1}+\bar{\Pi}_{\tau+1}-U\right)f\left(\bar{m}_{\tau+1}|\bar{m}_\tau,s_\tau^2\right)\\&\quad-\beta\left(1-\delta\right)\int_{\bar{m}_{\tau+1}}^{\infty}\left[\frac{\partial\left(J_{\tau+1}+\Pi_{\tau+1}-U\right)}{\partial x}+\frac{\partial\bar{m}_\tau}{\partial x}\left(J_{\tau+1}+\Pi_{\tau+1}-U\right)\left(\frac{m_{\tau+1}-\bar{m}_\tau}{s_\tau^2}\right)\right]d\bar{F}_{\tau+1}\\&=\frac{\left(1-\beta\right)\frac{\partial U}{\partial x}-1-\beta\left(1-\delta\right)\int_{\bar{m}_{\tau+1}}^{\infty}\frac{\partial\left(J_{\tau+1}+\Pi_{\tau+1}-U\right)}{\partial x}d\bar{F}_{\tau+1}}{1+\beta\left(1-\delta\right)\int_{\bar{m}_{\tau+1}}^{\infty}\left(J_{\tau+1}+\Pi_{\tau+1}-U\right)\left(\frac{m_{\tau+1}-\bar{m}_\tau}{s_\tau^2}\right)d\bar{F}_{\tau+1}}\end{aligned} \tag{73}$$

where $\bar{J}_{\tau+1}$ is shorthand for the value of $J_{\tau+1}$ when $m_{\tau+1} = \bar{m}_{\tau+1}$, $\bar{\Pi}_{\tau+1}$ is defined analogously, and $\bar{J}_{\tau+1} + \bar{\Pi}_{\tau+1} = U$ by definition of $\bar{m}_{\tau+1}$.

Applying the first result from Lemma 6,

$$\frac{\partial \bar{m}_\tau}{\partial b} = \frac{-1 - \beta(1-\delta)\int_{\bar{m}_{\tau+1}}^{\infty} \frac{\partial (J_{\tau+1}+\Pi_{\tau+1})}{\partial b} d\bar{F}_{\tau+1}}{1 + \beta(1-\delta)\int_{\bar{m}_{\tau+1}}^{\infty} (J_{\tau+1} + \Pi_{\tau+1} - U)\left(\frac{m_{\tau+1}-\bar{m}_\tau}{s_\tau^2}\right) d\bar{F}_{\tau+1}}. \tag{74}$$

Since $\partial(J_{\tau+1} + \Pi_{\tau+1})/\partial b > 0$ by Lemma 7, the numerator is negative. The denominator is positive because $J_{\tau+1}+\Pi_{\tau+1} \geq U$ for $m_{\tau+1} \geq \bar{m}_{\tau+1}$ (with equality only when $m_{\tau+1} = \bar{m}_{\tau+1}$); and $m_{\tau+1} \geq \bar{m}_{\tau+1} \geq \bar{m}_\tau$ by Proposition 4. Thus $\partial \bar{m}_\tau/\partial b < 0$.

Letting $x = a$ in (73) gives

$$\frac{\partial \bar{m}_\tau}{\partial a} = \frac{(1-\beta)\frac{\partial U}{\partial a} - 1 - \beta(1-\delta)\int_{\bar{m}_{\tau+1}}^{\infty} \frac{\partial (J_{\tau+1}+\Pi_{\tau+1}-U)}{\partial a} d\bar{F}_{\tau+1}}{1 + \beta(1-\delta)\int_{\bar{m}_{\tau+1}}^{\infty} (J_{\tau+1} + \Pi_{\tau+1} - U)\left(\frac{m_{\tau+1}-\bar{m}_\tau}{s_\tau^2}\right) d\bar{F}_{\tau+1}}. \tag{75}$$

As in (74), the denominator is positive. To sign the numerator note that for all $s \geq 1$,

$$\begin{aligned} J_{\tau+s} + \Pi_{\tau+s} - U &= \mu + a + b + m_{\tau+s} - (1-\beta)U \\ &\quad + \beta(1-\delta)\int_{\bar{m}_{\tau+s+1}}^{\infty} (J_{\tau+s+1} + \Pi_{\tau+s+1} - U)\, dF_{\tau+s+1} \end{aligned} \tag{76}$$

and differentiating gives the recursion

$$\frac{\partial (J_{\tau+s} + \Pi_{\tau+s} - U)}{\partial a} = 1 - (1-\beta)\frac{\partial U}{\partial a} + \beta(1-\delta)\int_{\bar{m}_{\tau+s+1}}^{\infty} \frac{\partial (U - J_{\tau+s+1} - \Pi_{\tau+s+1})}{\partial a} dF_{\tau+s+1}. \tag{77}$$

Repeated substitution of (77) into the numerator of (75) gives

$$\frac{\partial \bar{m}_\tau}{\partial a} = \frac{\left[(1-\beta)\frac{\partial U}{\partial a} - 1\right]\bar{Z}_\tau}{1 + \beta(1-\delta)\int_{\bar{m}_{\tau+1}}^{\infty} (J_{\tau+1} + \Pi_{\tau+1} - U)\left(\frac{m_{\tau+1}-\bar{m}_\tau}{s_\tau^2}\right) d\bar{F}_{\tau+1}} \tag{78}$$

where

$$\bar{Z}_\tau = 1 + \sum_{s=1}^{\infty} [\beta(1-\delta)]^s \int_{\bar{m}_{\tau+1}}^{\infty}\int_{\bar{m}_{\tau+2}}^{\infty} \cdots \int_{\bar{m}_{\tau+s}}^{\infty} dF_{\tau+s} \cdots dF_{\tau+2} d\bar{F}_{\tau+1} > 0. \tag{79}$$

The numerator of (78) is negative, since $\partial U/\partial a < \frac{1}{1-\beta}$ by lemma 6. ■

B Appendix: The Steady State

B.1 Flows Into Unemployment

Let $l(a, b, \tau)$ denote the density of type a workers employed at type b firms with tenure τ. The number of such workers entering unemployment in a given period is

$$\begin{aligned} & (1-u)\, l(a, b, \tau) \left[\Pr(m_\tau < \bar{m}_\tau) + \delta\right] \\ = \; & (1-u)\, l(a, b, \tau) \left[\Phi\left(\frac{\bar{m}_\tau}{V_\tau^{1/2}}\right) + \delta\right] \end{aligned} \tag{80}$$

where Φ denotes the standard normal CDF. The flow into unemployment of all type a workers from type b firms is

$$(1-u) \sum_{\tau=1}^{\infty} l(a, b, \tau) \left[\Phi\left(\frac{\bar{m}_\tau}{V_\tau^{1/2}}\right) + \delta\right] \tag{81}$$

and the aggregate flow into unemployment is

$$(1-u) \int_{a_0}^{a_1} \int_{b_0}^{b_1} \sum_{\tau=1}^{\infty} l(a, b, \tau) \left[\Phi\left(\frac{\bar{m}_\tau}{V_\tau^{1/2}}\right) + \delta\right] dbda. \tag{82}$$

B.2 Vacancies

In the steady state, the probability λ that a vacancy is filled is constant. Thus the equilibrium number of vacancies opened by each firm, i.e., the solution to (25), is also constant. Let υ_b^* denote the steady state number of vacancies opened by a type b firm, i.e.,

$$\upsilon_b^* = \arg\max_{\upsilon \in \mathbb{N}} \sum_{l=0}^{\upsilon} \binom{\upsilon}{l} \lambda^l (1-\lambda)^{\upsilon - l} \left[l \int_{a_0}^{a_1} \Pi_0 dF_a^* - \kappa(l) \right] \tag{83}$$

when λ takes its steady state value and where F_a^* is the steady state distribution of unemployed worker types defined below. Let f_b denote the density function associated with the distribution F_b of firm types. The steady state number of vacancies opened by all type b firms is $\phi \upsilon_b^* f_b(b)$, and the steady state stock of vacancies in the economy is

$$v = \phi \int_{b_0}^{b_1} \upsilon_b^* f_b(b)\, db. \tag{84}$$

B.3 Steady State Type Distributions

Each open vacancy is associated with a firm type b. Let $f_b^*(b)$ denote the steady state distribution of firm types among open vacancies. This is

$$f_b^*(b) = \phi \frac{\upsilon_b^*}{v} f_b(b) \tag{85}$$

with corresponding CDF F_b^*. Workers use F_b^* to compute the expected value of employment in new matches before the identity of the matching firm is known.

Similarly, we can define the distribution F_a^* of unemployed worker types. Firms use F_a^* to compute the expected value of employment in new matches before the identity of the matching worker is known. Define the density of employed type a workers:

$$l(a) = \int_{b_0}^{b_1} \sum_{\tau=1}^{\infty} l(a, b, \tau)\, db. \tag{86}$$

Then the density function f_a^* associated with F_a^* is $f_a^*(a) = u^{-1}[f_a(a) - (1-u)\, l(a)]$, where f_a is the density function associated with the distribution F_a of worker types.

B.4 Flows Out of Unemployment

The flow of type a workers out of unemployment and into type b firms is

$$m(u, v)\, f_a^*(a)\, f_b^*(b). \tag{87}$$

Thus the aggregate flow out of unemployment is

$$m(u, v) \int_{a_0}^{a_1} \int_{b_0}^{b_1} f_a^*(a)\, f_b^*(b)\, db da = m(u, v). \tag{88}$$

The steady state flow-balance condition is the equality of (81) and (87) for all worker types a and all firm types b. This implies the aggregate steady state flow-balance (82) = (88). The steady state level of unemployment is implicitly defined by this equality when v takes its steady state value.

B.5 Firm size

Let

$$l(b) = \int_{a_0}^{a_1} \sum_{\tau=1}^{\infty} l(a, b, \tau)\, da \tag{89}$$

be the density of employment at type b firms. Then the average size of type b firms is

$$\frac{(1-u)\, l(b)}{\phi f_b(b)}. \tag{90}$$

C Data Appendix

C.1 Sample Construction

The sample is restricted to full-time private sector employees at their dominant employer,[51] between 25 and 65 years of age, who had no more than 44 employers in the sample period,[52] with real annualized earnings between \$1,000 and \$1,000,000 (1990 dollars), employed in non-agricultural jobs that included at least one full quarter of employment.[53] The resulting sample consists of 174 million quarterly earnings observations on 9.3 million individuals employed at approximately 575,000 firms, for a total of over 15 million unique worker-firm matches. The quarterly records are annualized for estimation, for a sample of 49.3 million annual records.

C.1.1 The Dense Samples

The dense sampling algorithm of Woodcock (2005) ensures that individuals are connected to a specified minimum number of other workers by means of a common employer. In brief, this result is achieved by sampling firms first, with probabilities proportional to employment in a reference period. Workers are then sampled within firms, with probabilities inversely proportional to firm employment. A minimum of n employees are sampled from each firm. The resulting sample is equivalent to a simple random sample of workers employed in the reference period (that is, each worker has an equal probability of being sampled), but guarantees that each worker is connected to at least n others by a common employer.

I draw two disjoint one percent dense random samples of workers employed in 1997 using this algorithm. Each worker is connected to at least $n = 5$ others.[54] For comparison, I also draw a one percent simple random sample of workers employed in 1997. Table 1 presents connectedness properties of the full sample, Dense Sample 1, and the simple random sample.[55] The full sample is highly connected: the largest connected group contains 99.06 percent of jobs. The dense sample remains quite highly connected: about 92 percent of jobs are contained in the two largest connected groups. This is in contrast to the simple random sample: though about 80 percent of jobs are contained in the two largest groups, only 84 percent are in groups containing at least 5 worker-firm matches. By construction, all jobs in the dense samples are contained in groups of at least 5 matches. In the simple random sample, fully 5.5 percent of jobs are connected to no other.

[51] A dominant employer is identified for each individual in each year. An individual's dominant employer in year t is the employer at which her reported UI earnings were largest in t. About 87 percent of the UI system wage records correspond to employment at a dominant employer.

[52] There is some concern that observing an extreme number of employment spells may be due to measurement error in the person and firm identifiers. Around 0.5 percent of quarterly wage observations corresponded to individuals employed at more than 44 employers over the sample period.

[53] An individual employed at SEIN j in quarter q is considered to have worked a full quarter if she was employed at j in quarters $q-1$ and $q+1$.

[54] The other parameters used to draw the dense samples, defined in Woodcock (2005), are $m = 0.5$ and $p = 0.004$.

[55] Characteristics of the two dense samples are virtually identical.

C.2 Variable Creation and Missing Data Imputation

Time-varying covariates X include a quartic in labor force experience (interacted with sex), four dummy variables to indicate the number of full quarters the individual worked in the year (interacted with sex), and year effects. Time-invariant person characteristics U are education (five categories, interacted with sex), race (3 categories, interacted with sex), and a dummy variable to indicate if the initial experience measure was negative (interacted with sex).[56]

Missing data items include full-time status, education, tenure (for left-censored job spells), initial experience, and (in some cases discussed below) the earnings measure. Missing data items are multiply-imputed using the Sequential Regression Multivariate Imputation (SRMI) method. See Rubin (1987) for a general treatment of multiple-imputation; the SRMI technique is due to Raghunathan et al. (2001); Abowd and Woodcock (2001) generalize SRMI to the case of longitudinal linked data. SRMI imputes missing data in a sequential and iterative fashion on a variable-by-variable basis. Each missing data item is multiply-imputed with draws from the posterior predictive distribution of an appropriate generalized linear model under a diffuse prior. Full estimation results of each of the imputation regressions are available on request. I generate three imputed values of each missing data item. The result is three versions of the analysis sample, each containing different imputed values for the missing data items. In keeping with the statistical literature on multiple imputation, I refer to these as completed data implicates.

C.2.1 Real Annualized Earnings

The dependent variable for the earnings regressions is log real annualized earnings. The annualized measure is constructed from real full-quarter earnings. Full quarter earnings are defined as follows. For individuals who worked a full quarter at firm j in t, full-quarter earnings equal UI reported earnings (about 80 percent of the analysis sample). For individuals who did not work a full quarter in t, one of two earnings measures is used. If the individual worked at least one full quarter in the four previous or subsequent quarters, and if reported earnings in quarter t were at least 80 percent of average real earnings in the full quarters, the individual is presumed to have worked a full quarter.[57] That is, reported earnings are treated as full-quarter earnings (12.5 percent of the analysis sample). If on the other hand reported earnings are less than 80 percent of average real average earnings in the full quarters, earnings are imputed to the full-quarter level (7.5 percent of the analysis sample). The imputation model is a linear regression on log real full quarter earnings. Conditioning variables include up to four leads and four lags of full quarter earnings (where available), year and quarter dummies, race, education (5 categories), labor market experience (linear through quartic terms), and SIC division. Separate imputation models were estimated for men and for women. For each quarter in which earnings are imputed to the full-quarter level, three imputed values are drawn from the posterior predictive distribution under a diffuse prior. After constructing the real full-quarter earnings measure, these are annualized to obtain real

[56] As described in Section C.2.3, initial experience is set to zero in this case.

[57] The 80 percent cutoff was chosen as follows. For individuals that worked a full quarter in q, the median ratio or quarter q earnings to average full-quarter earnigns in quarters $q-4$ to $q+4$ was 0.8.

annualized earnings.

C.2.2 Education

Education is multiply-imputed from the 1990 Decennial Census long form. The imputation model is an ordered logit. There are 13 outcome categories, corresponding to 0 through 20 years of education. Conditioning variables include age (10 categories), vintiles of real annual earnings at the dominant employer in 1990 or the year the individual first appeared in the sample, and SIC division. Separate imputation models were estimated for men and for women. For each person, three imputed values are drawn from the normal approximation (at the mode) to the posterior predictive distribution under a diffuse prior. The education measure is subsequently collapsed to five categories: Less than high school, High school graduate, Some college or vocational training, Undergraduate degree, and Graduate or professional degree.

C.2.3 Labor Market Experience

In the first quarter that an individual appears in the sample, I calculate potential labor market experience (in years) as age at the beginning of the quarter minus years of education minus 6. In cases where this measure is negative, potential experience is set to zero. In each subsequent quarter, labor market experience is accumulated using the individual's realized labor market history. Note that since initial experience depends on the multiply-imputed education measure, calculated labor market experience varies across the three completed data implicates.

C.2.4 Tenure

Jobs fall into two categories with respect to the calculation of job tenure: spells that are left-censored and spells that are not. In one state the data series begins in the first quarter of 1990; in the other state, the data series begins earlier. All jobs with positive earnings in the first quarter of available data for that state are presumed left censored. Such spells comprise 33 percent of jobs in the full sample.

For non-left-censored spells, tenure is set to 1 in the first quarter that there is a UI wage record, and is subsequently accumulated using the individual's employment history. For left-censored spells, tenure as of the first quarter of 1990 is imputed using data from the 1996 and 1998 CPS February supplements. The imputation model is a linear regression on the natural logarithm of tenure. Conditioning variables include age (10 categories), vintiles of real annual earnings at the dominant employer in 1990, education (5 categories), and SIC division. For each left-censored job, three imputed values of tenure in 1990 quarter 1 were drawn from the posterior predictive distribution under a diffuse prior. In subsequent quarters, tenure is accumulated using the individual's employment history. Note that left-censored spells are excluded from the reported estimates of W.

C.2.5 Full-Time Status

Full-time status is multiply-imputed using the 1982-1999 CPS March supplements. The imputation model is a binary logit. Conditioning variables include a quadratic in age, SIC division, year dummies, and vintiles of reported annual earnings at the dominant employer. Separate imputation models were estimated for men and for women. For each worker-firm match in each year, three imputed values were drawn from the normal approximation (at the mode) to the posterior predictive distribution under a diffuse prior.

C.3 Characteristics of the Samples

Table 2 presents basic summary statistics for the full analysis sample, the two dense samples, and a simple random sample. The dense samples exhibit properties virtually identical to those of the simple random sample, confirming the analytic proof of equivalence in Woodcock (2005). Since these are point-in-time samples, their properties differ slightly from those of the full sample. In particular, they exhibit properties consistent with a sample of individuals with a somewhat stronger-than-average labor force attachment: individuals in the point-in-time samples are somewhat more likely to be male, are slightly more educated, have somewhat longer average job tenure, earn slightly more, and are somewhat more likely to work a full calendar year. These are all small differences, however.

References

Abowd, J. M. and D. Card (1989). On the covariance structure of earnings and hours changes. *Econometrica 57*(2), 411–445.

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. and F. Kramarz (1999). Econometric analysis of linked employer-employee data. *Labour Economics 6*(1), 53–74.

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, P. Lengermann, and S. Roux (2005). Persistent inter-industry wage differences: Rent sharing and oppotunity costs. 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., J. Lane, and R. Prevost (2000). Design and conceptual issues in realizing analytical enhancements through data linkages of employer and employee data. In *Proceedings of the Federal Committee on Statistical Methodology.*

Abowd, J. M., P. Lengermann, and K. McKinney (2003). The measurement of human capital in the U.S. economy. Mimeo.

Abowd, J. M. and S. D. Woodcock (2001). Disclosure limitation in longitudinal linked data. In P. Doyle, J. I. Lane, J. M. Theeuwes, and L. M. Zayatz (Eds.), *Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies*, Chapter 10, pp. 215–277. Elsevier Science.

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

Albrecht, J. and S. Vroman (2002). A matching model with endogenous skill requirements. *International Economic Review 43*(1), 283–305.

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

Andersson, F., H. J. Holzer, and J. I. Lane (2005). *Moving Up or Moving On: Who Advances in the Low-Wage Labour Market.* Russell Sage Foundation.

Andrews, M., T. Schank, and R. Upward (2004a). Practical estimation methods for linked employer-employee data. IAB Discussion Paper No. 3/2004.

Andrews, M. J., T. Schank, and R. Upward (2004b). High wage workers and low wage firms: Negative assortative matching or statistical artefact? Mimeo.

Bartel, A. P. and G. J. Borjas (1981). Wage growth and job turnover: An empirical analysis. In S. Rosen (Ed.), *Studies in Labor Markets*, Chapter 2, pp. 65–90. New York: NBER.

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.

Blackwell, D. (1965). Discounted dynamic programming. *Annals of Mathematical Statistics 36*, 226–235.

Brown, C. and J. L. Medoff (1989). The employer-size wage effect. *Journal of Political Economy 97*, 1027–1059.

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

Chamberlain, G. (1984). Panel data. In Z. Grilliches and M. Intriligator (Eds.), *Handbook of Econometrics, Volume II*, Chapter 22, pp. 1248–1318. New York: Elsevier.

Dostie, B. (2005). Job turnover and the returns to seniority. *Journal of Business and Economic Statistics 23*(2), 192–199.

Farber, H. S. and R. Gibbons (1996). Learning and wage dynamics. *The Quarterly Journal of Economics 111*(4), 1007–1047.

Felli, L. and C. Harris (1996). Learning, wage dynamics, and firm-specific human capital. *Journal of Political Economy 104*(4), 838–868.

Flinn, C. J. (1986). Wages and the job mobility of young workers. *Journal of Political Economy 94*(3), S88–S110.

Gautier, P. A. (2002). Unemployment and search externalities in a model with heterogeneous jobs and workers. *Economica 69*(273), 21–40.

Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2005). Comparative advantage, learning, and sectoral wage discrimination. *Journal of Labor Economics 23*(4), 681–724.

Gilmour, A. R., R. Thompson, and B. R. Cullis (1995). Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics 51*, 1440–1450.

Goux, D. and E. Maurin (1999). Persistence of interindustry wage differentials: A reexamination using matched worker-firm panel data. *Journal of Labor Economics 17*(3), 492–533.

Gruetter, M. (2005). Identification of firm effects when mobility is endogenous: A simulation study. Mimeo.

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

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*(4), 1029–1054.

Harris, M. and B. Holmstrom (1982). A theory of wage dynamics. *Review of Economic Studies XLIX*, 315–333.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica 46*(6), 1251–1271.

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.

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

Kohns, S. (2000). Different skill levels and firing costs in a matching model with uncertainty - an extension of Mortensen and Pissarides (1994). IZA Discussion Paper No. 104.

LEHD Program (2002). The Longitudinal Employer-Household Dynamics Program: Employment Dynamics Estimates Project versions 2.2 and 2.3. LEHD Technical Paper 2002-05.

Lillard, L. A. (1999). Job turnover heterogeneity and person-job-specific time-series wages. *Annales D'Économie et de Statistique 55-56*, 183–210.

McCulloch, C. E. and S. R. Searle (2001). *Generalized, Linear, and Mixed Models*. New York: John Wiley and Sons.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: NBER.

Mincer, J. and B. Jovanovic (1981). Labor mobility and wages. In S. Rosen (Ed.), *Studies in Labor Markets*, Chapter 1, pp. 21–64. New York: NBER.

Moscarini, G. (2003). Skill and luck in the theory of turnover. Yale University Working Paper.

Newey, W. K. (1985). Generalized method of moments specification tests. *Journal of Econometrics 29*, 229–256.

Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica 70*(4), 2295–2350.

Raghunathan, T. E., J. M. Lepkowski, J. V. Hoewyk, and P. Solenberger (2001). A multivariate technique for multiply-imputing missing values using a sequence of regression models. *Survey Methodology 27*(1), 85–95.

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

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Sattinger, M. (1975). Comparative advantage and the distributions of earnings and abilities. *Econometrica 43*(3), 455–468.

Sattinger, M. (1995). Search and the efficient assignment of workers to jobs. *International Economic Review 36*(2), 283–302.

Schall, R. (1991). Estimation in generalized linear models with random effects. *Biometrika 78*(4), 719–727.

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

Searle, S. R., G. Casella, and C. E. McCulloch (1992). *Variance Components*. 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.

Shimer, R. and L. Smith (2001). Matching, search, and heterogeneity. *B.E. Journals in Macroeconomics 1*(1), Article 5.

Stern, S. (1990). The effects of firm optimizing behaviour in matching models. *Review of Economic Studies 57*, 647–660.

Stevens, D. W. (2002). State UI wage records: Description, access and use. Mimeo.

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. (2005). Sampling connected histories from longitudinal linked data. Mimeo.

TABLE 1
PROPERTIES OF CONNECTED GROUPS OF WORKERS AND FIRMS

	Full Sample[a]	Dense Sample 1[b]	Simple Random Sample[c]
Number of Groups	84,708	1,140	9,457
Number of Workers	9,271,766	49,425	49,200
Number of Firms	573,237	27,421	40,064
Number of Worker-Firm Matches	15,305,508	92,539	93,182
Number of Matches in Smallest Group	1	5	1
Percentage of Matches in:			
Largest Group	99.06	67.25	59.37
Second Largest Group	0.0006	24.70	20.30
Third Largest Group	0.0003	0.04	0.06
Groups containing 5 or more matches	99.21	100	84.44
Groups containing only 1 match	0.35	0	5.50

[a] Results combined across three completed data implicates.
[b] One percent dense random samples of workers employed in 1997. Results are combined across three completed data
[c] One percent simple random sample of workers employed in 1997. Results are for one completed data implicate.

TABLE 2
SUMMARY STATISTICS
Estimates Combined Across 3 Completed Data Implicates

	Full Sample		Dense Sample 1		Simple Random Sample	
Variable	Mean[a]	Std. Dev[b]	Mean[a]	Std. Dev[b]	Mean[a]	Std. Dev[b]
Demographic Characteristics						
Male (Proportion)	0.56	0.50	0.56	0.50	0.57	0.50
Age (Years)	40.6	10.2	40.4	9.5	40.3	9.5
Men						
Nonwhite (Proportion)	0.21	0.57	0.21	0.57	0.21	0.57
Race Missing (Proportion)	0.04	0.25	0.03	0.24	0.03	0.24
Less than high school (Proportion)	0.12	0.45	0.11	0.43	0.11	0.42
High school (Proportion)	0.30	0.67	0.29	0.66	0.30	0.66
Some college (Proportion)	0.23	0.60	0.23	0.60	0.23	0.59
Associate or Bachelor Degree (Proportion)	0.25	0.62	0.26	0.62	0.26	0.62
Graduate or Professional Degree (Proportion)	0.10	0.42	0.11	0.43	0.11	0.42
Women						
Nonwhite (Proportion)	0.24	0.69	0.24	0.70	0.24	0.60
Race Missing (Proportion)	0.02	0.22	0.02	0.22	0.02	-0.01
Less than high school (Proportion)	0.09	0.45	0.09	0.43	0.09	0.28
High school (Proportion)	0.31	0.78	0.30	0.77	0.30	0.39
Some college (Proportion)	0.25	0.71	0.25	0.71	0.25	0.31
Associate or Bachelor Degree (Proportion)	0.26	0.72	0.28	0.75	0.27	0.42
Graduate or Professional Degree (Proportion)	0.08	0.42	0.09	0.44	0.09	0.31
Work History Characteristics						
Tenure (Years)	4.5	3.5	4.9	3.6	4.8	3.6
Job is Left Censored (Proportion)	0.33	0.47	0.36	0.48	0.35	0.48
Real Annualized Earnings (1990 Dollars)	53755	50804	57209	51196	56483	50074
Men						
Labor Market Experience (Years)	11.8	13.1	11.7	12.6	11.7	12.6
Initial Experience <0 (Proportion)	0.02	0.20	0.02	0.20	0.02	0.20
Worked 0 Full Quarters in Calendar Year (Proportion)	0.08	0.36	0.06	0.32	0.06	0.32
Worker 1 Full Quarter in Calendar Year (Proportion)	0.15	0.49	0.11	0.44	0.12	0.44
Worker 2 Full Quarters in Calendar Year (Proportion)	0.13	0.47	0.12	0.45	0.12	0.45
Worker 3 Full Quarters in Calendar Year (Proportion)	0.14	0.48	0.14	0.47	0.13	0.47
Worker 4 Full Quarters in Calendar Year (Proportion)	0.50	0.91	0.57	0.96	0.56	0.97
Women						
Labor Market Experience (Years)	9.5	13.0	9.3	12.6	9.2	12.5
Initial Experience <0 (Proportion)	0.02	0.23	0.02	0.23	0.02	0.22
Worked 0 Full Quarters in Calendar Year (Proportion)	0.07	0.39	0.05	0.35	0.06	0.36
Worker 1 Full Quarter in Calendar Year (Proportion)	0.14	0.54	0.11	0.49	0.11	0.50
Worker 2 Full Quarters in Calendar Year (Proportion)	0.13	0.53	0.12	0.50	0.12	0.51
Worker 3 Full Quarters in Calendar Year (Proportion)	0.14	0.55	0.13	0.53	0.13	0.53
Worker 4 Full Quarters in Calendar Year (Proportion)	0.52	0.96	0.59	1.00	0.59	1.01
Number of Observations	49,281,533		357,725		357,009	
Number of Workers	9,271,766		49,425		49,200	
Number of Firms	573,237		27,421		40,064	
Number of Worker-Firm Matches	15,305,508		92,539		93,182	

[a] Means are computed on each completed data implicate for each sample. The reported value is the arithmetic mean of the means computed on each implicate.

[b] The variance of each variable is computed on each completed data implicate for each sample. The reported value is the square root of the arithmetic mean of the variances computed on each implicate.

TABLE 3
ESTIMATED VARIANCE COMPONENTS AND SUMMARY OF MODEL FIT
Combined Results From 3 Completed Data Implicates

	Fixed Model		Mixed Model With Spherical Error		Mixed Model With Unrestricted Within-Match Error Covariance	
	Estimate[a,e]	Standard Error[b]	Estimate[a]	Standard Error[b]	Estimate[a,f]	Standard Error[b]
No Correction for Truncation						
Variance of person effect (σ^2_α)	0.29	(0.002)	0.23	(0.005)	0.18	(0.002)
Variance of firm effect (σ^2_ψ)	0.08	(0.000)	0.15	(0.002)	0.08	(0.007)
Residual variance (σ^2_ε)	0.06	(0.000)	0.04	(0.001)	n/a	n/a
AIC[c]	-2.4	(0.000)	-3.1	(0.016)	-2.9	(0.009)
BIC[c]	0.7	(0.001)	-3.1	(0.016)	-2.9	(0.009)
Var(out-of-sample prediction error)[c,d]			0.036	(0.000)	0.038	(0.000)
Corrected for Truncation						
Variance of person effect (σ^2_α)			0.23	(0.005)	0.18	(0.002)
Variance of firm effect (σ^2_ψ)			0.15	(0.002)	0.08	(0.006)
Residual variance (σ^2_ε)			0.04	(0.001)	n/a	n/a
Truncation Correction (β_λ)			0.04	(0.002)	0.02	(0.001)
AIC[c]			-3.1	(0.016)	-2.9	(0.006)
BIC[c]			-3.1	(0.016)	-2.9	(0.006)
Var(out-of-sample prediction error)[c,d]			0.036	(0.000)	0.038	(0.000)
Number of Observations	49,281,533	(9103)	357,725	(2363)	357,725	(2363)
Number of Workers	9,271,766	(710)	49,425	(150)	49,425	(150)
Number of Firms	573,237	(118)	27,421	(13)	27,421	(13)
Number of Worker-Firm Matches	15,305,508	(3196)	92,539	(470)	92,539	(470)

[a] Arithmetic mean of parameter estimate across three completed data implicates.
[b] Square root of estimated total variance of parameter estimate across three completed data implicates, as defined in Rubin (1987).
[c] Value in column labeled "Standard Error" is the square root of the estimated between-implicate variance.
[d] Computed on Dense Sample 2.
[e] Sample variance of estimated person and firm effects, averaged across three completed data implicates.
[f] Estimates based on multiply-imputed tenure for left-censored spells.

TABLE 4
ESTIMATED VARIANCE COMPONENTS AND SUMMARY OF MODEL FIT
Combined Results From 3 Completed Data Implicates, Left-Censored Spells Excluded

	Fixed Model		Mixed Model With Spherical Error		Mixed Model With Unrestricted Within-Match Error Covariance	
	Estimate[a,e]	Standard Error[b]	Estimate[a]	Standard Error[b]	Estimate[a]	Standard Error[b]
No Correction for Truncation						
Variance of person effect (σ^2_α)	0.29	(0.002)	0.24	(0.006)	0.18	(0.003)
Variance of firm effect (σ^2_ψ)	0.08	(0.000)	0.15	(0.003)	0.08	(0.001)
Residual variance (σ^2_ε)	0.07	(0.000)	0.05	(0.001)	n/a	n/a
AIC[c]	-2.0	(0.000)	-3.0	(0.018)	-2.8	(0.025)
BIC[c]	4.0	(0.001)	-3.0	(0.018)	-2.8	(0.025)
Var(out-of-sample prediction error)[c,d]			0.038	(0.000)	0.045	(0.000)
Corrected for Truncation						
Variance of person effect (σ^2_α)			0.24	(0.005)	0.18	(0.003)
Variance of firm effect (σ^2_ψ)			0.15	(0.003)	0.08	(0.001)
Residual variance (σ^2_ε)			0.05	(0.001)	n/a	n/a
Truncation Correction (β_λ)			0.00	(0.003)	0.02	(0.003)
AIC[c]			-3.0	(0.018)	-2.8	(0.005)
BIC[c]			-3.0	(0.018)	-2.8	(0.005)
Var(out-of-sample prediction error)[c,d]			0.038	(0.000)	0.045	(0.000)
Number of Observations	32,800,936	(7217)	228,386	(2018)	228,386	(2018)
Number of Workers	12,289,989	(2855)	39,816	(168)	39,816	(168)
Number of Firms	544,254	(177)	24,624	(22)	24,624	(22)
Number of Worker-Firm Matches	7,577,051	(694)	73,307	(475)	73,307	(475)

[a] Arithmetic mean of parameter estimate across three completed data implicates.
[b] Square root of estimated total variance of parameter estimate across three completed data implicates, as defined in Rubin (1987).
[c] Value in column labeled "Standard Error" is the square root of the estimated between-implicate variance.
[d] Computed on Dense Sample 2.
[e] Sample variance of estimated person and firm effects, averaged across three completed data implicates.

TABLE 5
CORRELATIONS AMONG ESTIMATED EFFECTS
Combined Results From 3 Completed Data Implicates

No Correction for Truncation

Fixed Model	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1	0.74	0.66	0.34	0.45	0.18
Total Person Effect (θ)	0.74	1	0.91	0.41	0.03	-0.30
Unobserved Component (α)	0.66	0.91	1	0.00	0.00	-0.27
Observed Component (Uη)	0.34	0.41	0.00	1	0.09	-0.12
Total Firm Effect (ψ)	0.45	0.03	0.00	0.09	1	0.05
Time-Varying Covariates (Xβ)	0.18	-0.30	-0.27	-0.12	0.05	1

Mixed Model With Spherical Error	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1	0.80	0.71	0.38	0.47	0.29
Total Person Effect (θ)	0.80	1	0.91	0.41	0.03	0.02
Unobserved Component (α)	0.71	0.91	1	-0.01	-0.01	-0.03
Observed Component (Uη)	0.38	0.41	-0.01	1	0.08	0.11
Total Firm Effect (ψ)	0.47	0.03	-0.01	0.08	1	0.04
Time-Varying Covariates (Xβ)	0.29	0.02	-0.03	0.11	0.04	1

Mixed Model With Unrestricted Error Covariance

	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1	0.82	0.73	0.36	0.54	0.30
Total Person Effect (θ)	0.82	1	0.87	0.49	0.22	0.02
Unobserved Component (α)	0.73	0.87	1	-0.01	0.20	-0.03
Observed Component (Uη)	0.36	0.49	-0.01	1	0.09	0.09
Total Firm Effect (ψ)	0.54	0.22	0.20	0.09	1	0.041
Time-Varying Covariates (Xβ)	0.30	0.02	-0.03	0.09	0.04	1

Corrected For Truncation

Mixed Model With Spherical Error	y	θ	α	Uη	ψ	Xβ	$\beta_\lambda\lambda$
Log Earnings (y)	1	0.80	0.71	0.38	0.48	0.29	-0.06
Total Person Effect (θ)	0.80	1	0.91	0.41	0.03	0.02	-0.01
Unobserved Component (α)	0.71	0.91	1	-0.01	0.00	-0.03	-0.01
Observed Component (Uη)	0.38	0.41	-0.01	1	0.08	0.12	-0.02
Total Firm Effect (ψ)	0.48	0.03	0.00	0.08	1	0.05	-0.18
Time-Varying Covariates (Xβ)	0.29	0.02	-0.03	0.12	0.05	1	0.01
Truncation Correction ($\beta_\lambda\lambda$)	-0.06	-0.01	-0.01	-0.02	-0.18	0.01	1

Mixed Model With Unrestricted Error Covariance

	y	θ	α	Uη	ψ	Xβ	$\beta_\lambda\lambda$
Log Earnings (y)	1	0.82	0.73	0.36	0.54	0.30	-0.06
Total Person Effect (θ)	0.82	1	0.87	0.49	0.22	0.02	0.00
Unobserved Component (α)	0.73	0.87	1	-0.01	0.20	-0.03	0.00
Observed Component (Uη)	0.36	0.49	-0.01	1	0.09	0.09	-0.01
Total Firm Effect (ψ)	0.54	0.22	0.20	0.09	1	0.04	-0.25
Time-Varying Covariates (Xβ)	0.30	0.02	-0.03	0.09	0.04	1	0.00
Truncation Correction ($\beta_\lambda\lambda$)	-0.06	0.00	0.00	-0.01	-0.25	0.00	1

TABLE 6
FIXED MODEL ESTIMATES OF WITHIN-MATCH ERROR COVARIANCE
Combined Results From 3 Completed Data Implicates, Left-Censored Spells Excluded

Tenure	1	2	3	4	5	6	7	8	9	10
1	0.088									
2	0.022	0.062								
3	0.004	0.019	0.052							
4	-0.004	0.005	0.017	0.049						
5	-0.010	-0.002	0.005	0.016	0.047					
6	-0.015	-0.008	-0.002	0.005	0.018	0.052				
7	-0.015	-0.010	-0.004	0.001	0.008	0.018	0.047			
8	-0.016	-0.011	-0.007	-0.003	0.002	0.007	0.019	0.050		
9	-0.017	-0.013	-0.009	-0.005	-0.002	0.001	0.010	0.020	0.052	
10	-0.017	-0.013	-0.011	-0.007	-0.005	-0.002	0.005	0.012	0.022	0.060

TABLE 7
MIXED MODEL ESTIMATES OF WITHIN-MATCH ERROR COVARIANCE
Combined Results From 3 Completed Data Implicates, Left-Censored Spells Excluded

No Correction for Truncation

Tenure	1	2	3	4	5	6	7	8	9	10
1	0.124									
2	0.066	0.094								
3	0.052	0.066	0.094							
4	0.047	0.059	0.070	0.099						
5	0.041	0.053	0.063	0.075	0.098					
6	0.036	0.049	0.058	0.068	0.078	0.104				
7	0.031	0.045	0.054	0.064	0.072	0.083	0.107			
8	0.030	0.043	0.052	0.063	0.071	0.078	0.086	0.115		
9	0.032	0.043	0.051	0.062	0.068	0.075	0.081	0.091	0.118	
10	0.030	0.041	0.051	0.061	0.066	0.072	0.077	0.083	0.092	0.119

Corrected for Truncation

Tenure	1	2	3	4	5	6	7	8	9	10
1	0.124									
2	0.066	0.094								
3	0.052	0.066	0.094							
4	0.047	0.059	0.071	0.099						
5	0.041	0.053	0.063	0.076	0.099					
6	0.036	0.049	0.059	0.069	0.079	0.105				
7	0.032	0.045	0.055	0.065	0.073	0.084	0.107			
8	0.031	0.043	0.052	0.064	0.071	0.079	0.087	0.115		
9	0.032	0.043	0.051	0.062	0.069	0.075	0.081	0.091	0.118	
10	0.029	0.041	0.050	0.061	0.066	0.072	0.077	0.083	0.091	0.118

TABLE 8
MINIMUM DISTANCE ESTIMATES OF STRUCTURAL PARAMETERS, SCALE PARAMETER γ=1
Combined Results From 3 Completed Data Implicates, Left-Censored Spells Excluded

	Fixed Model		Mixed Model	
	Parameter Estimate[b]	Standard Error[c]	Parameter Estimate[b]	Standard Error[c]
No Correction for Truncation				
Variance of Measurement Error (σ^2_u)	0.05	(0.000)	0.04	(0.001)
Variance of Match Quality (σ^2_c)[a]	0.01	(0.000)	0.09	(0.004)
Variance of Initial Signal (σ^2_z)[a]	0.97	(0.024)	0.10	(0.011)
Variance of Production Outcomes (σ^2_e)[a]	0.07	(0.002)	0.15	(0.052)
p-value from Test of Overidentifying Restrictions	< 0.0001		0.048	
Corrected for Truncation				
Variance of Measurement Error (σ^2_u)			0.04	(0.002)
Variance of Match Quality (σ^2_c)[a]			0.09	(0.005)
Variance of Initial Signal (σ^2_z)[a]			0.09	(0.010)
Variance of Production Outcomes (σ^2_e)[a]			0.11	(0.018)
p-value from Test of Overidentifying Restrictions			0.039	

[a] Estimates can be rescaled for alternate values $0<\gamma<1$ by dividing the pararameter estimate by γ^2.
[b] Average of parameter estimates across three completed data implicates.
[c] Square root of total variance of parameter estimate across three completed data implicates, as defined in Rubin (1987).

Figure 1
Estimated Cubic Regression of
Person Effect (Theta) on Firm Effect (Psi)

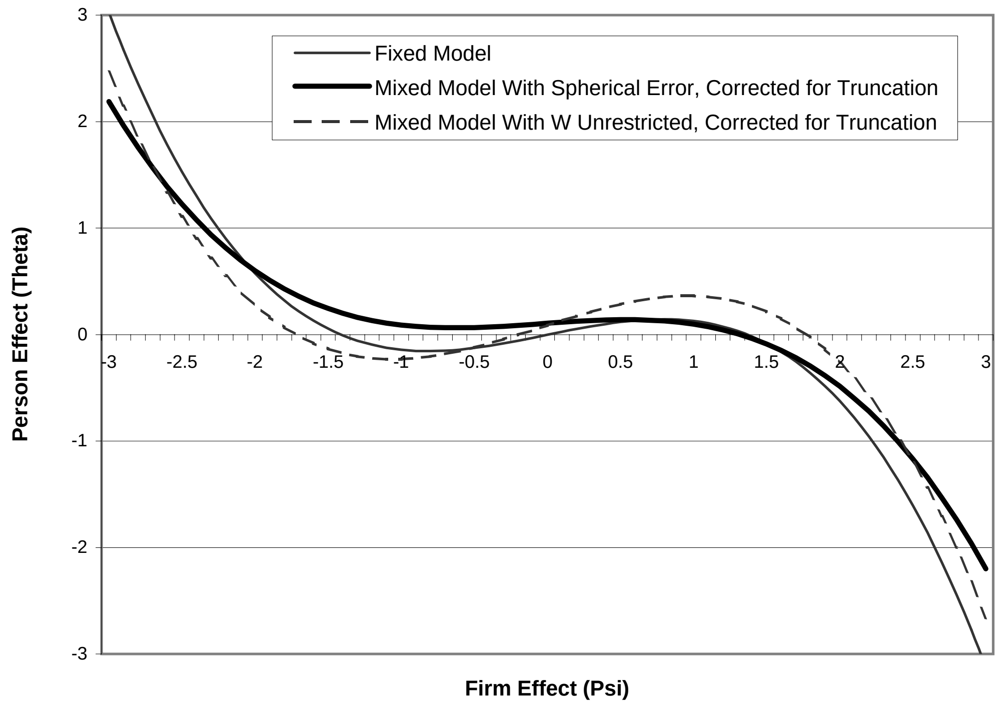

Figure 2
Estimated Sequence of Belief Variances

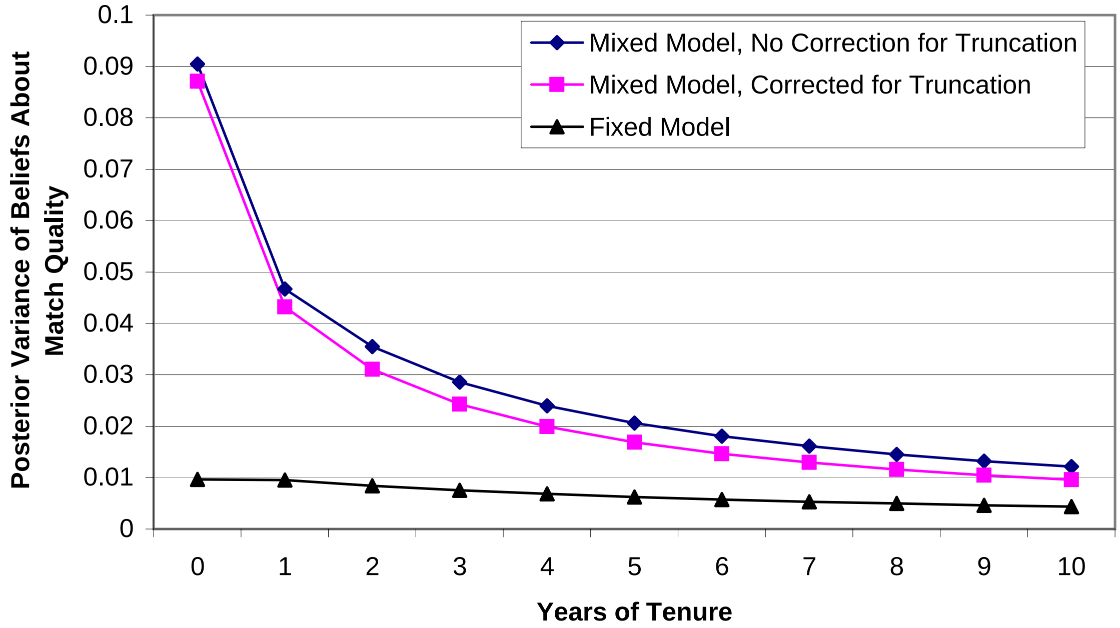

Figure 3

Estimated Relationship Between Person Effect and Completed Job Duration
Mixed Model With Unrestricted Within-Match Error Covariance
Left-Censored Spells Excluded (N=44,062 Completed Jobs)

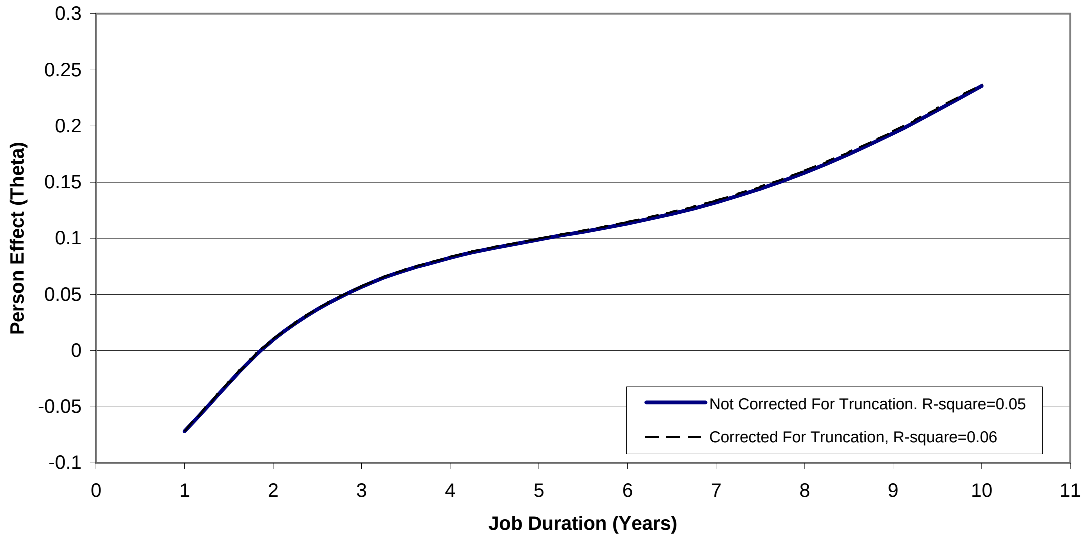

Figure 4

Estimated Relationship Between Firm Effect and Completed Job Duration
Mixed Model With Unrestricted Within-Match Error Covariance
Left-Censored Spells Excluded (N=44,062 Completed Jobs)

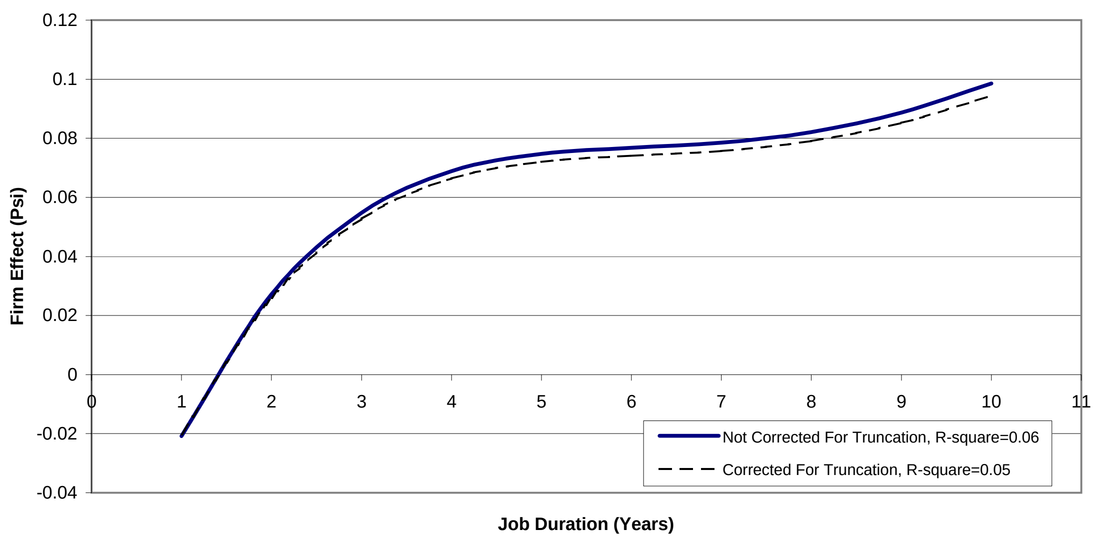

Figure 5

Estimated Relationship Between Firm Effect and Log(1997 Employment)
Mixed Model With Unrestricted Within-Match Error Covariance
(N=27,421 Firms)