Heterogeneity and Learning in Labor Markets*

Simon D. Woodcock†
Cornell University
sdw9@cornell.edu

July 3, 2003

*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, 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 is preparing to support external researchers' use of these data; please contact U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233, USA.

†I would like to thank John M. Abowd, David Easley, George Jakubson, Robert Shimer, Martha Stinson, members of the LEHD program staff, and seminar participants at the 2002 CEA meetings, Cornell University, the University of Western Ontario, Simon Fraser University, the University of Washington, RAND, the Federal Reserve Bank of Atlanta, the University of Wisconsin-Madison, Queen's University, McGill University, and the University of Montreal for helpful comments on earlier versions of this paper.

Abstract

I develop a matching model with heterogeneous workers, firms, and worker-firm matches, and apply it to longitudinal linked data on employers and employees. Workers vary in their marginal product when employed. Firms vary in their marginal product and cost of maintaining a vacancy. The marginal product of a worker-firm match also depends on a match-specific interaction between worker and firm that I call match quality. Agents have complete information about worker and firm heterogeneity, and symmetric but incomplete information about match quality. They learn its value slowly by observing production outcomes. There are two key results. First, under a Nash bargain, the equilibrium wage is linear in a person-specific component, a firm-specific component, and the posterior mean of beliefs about match quality. Second, in each period the optimal separation policy is characterized by a reservation level of beliefs about match quality. The reservation value varies across workers and firms, and is monotone in job tenure. These results have several implications for empirical work. The first implies that residuals within a worker-firm match are a martingale; the second implies the distribution of earnings is truncated.

I test predictions from the matching model using data from the Longitudinal Employer-Household Dynamics (LEHD) Program at the US Census Bureau. I present both fixed and mixed model specifications of the equilibrium wage function, taking account of structural aspects implied by the learning process. In the most general specification, earnings residuals have a completely unstructured covariance within a worker-firm match. I estimate structural parameters of the matching model, and test the martingale structure implied by the learning process. I find considerable support for the matching model in these data.

1 Introduction

It has long been recognized that observationally indistinguishable workers employed in seemingly identical firms earn different wages and have vastly different employment histories. Typically, observable worker and firm characteristics explain only about 30 percent of wage variation. Numerous authors have addressed this issue from a wide variety of perspectives. One branch of early empirical work focused on the role of unobserved heterogeneity on the part of workers as a determinant of employment outcomes. Another considered the importance of unobserved heterogeneity on the part of firms. Recent advances in the creation and analysis of longitudinal linked data on employers and employees have brought together these diverse literatures, and spawned a new one that examines the relative importance of unobserved worker and firm heterogeneity as determinants of employment outcomes, e.g., Abowd et al. (1999), and Abowd et al. (2002). This work has shown that most of the wage dispersion not explained by observable characteristics can be attributed to unmeasured characteristics of workers and firms. Does this reflect productivity differences, rent-sharing, or something else?

The purpose of this paper is twofold. The first is to provide a theoretical context in which to conceptualize the source of worker and firm differences, and their role in determining employment outcomes. To this end, I present a matching model with heterogeneous workers, heterogeneous firms, and heterogeneous worker-firm interactions. Workers and firms are imperfectly informed about the location of worker, firm, and match types. This precludes the optimal assignment of workers to firms. I endogenize employment mobility via a learning process. Workers and firms learn about the quality of a match by observing production outcomes. I show that the Nash-bargained equilibrium wage is linear in a person-specific component, a firm-specific component, and the posterior mean of beliefs about match quality. The optimal separation policy is of the reservation-wage type: workers and firms share a common reservation value of beliefs about match quality. The employment relation persists only so long as beliefs lie above this reservation value. Like most learning models, this reservation value reflects the option value of employment and increases with tenure. Unlike models with homogeneous agents, its value varies across employment matches.

The second goal of this paper is to extend the empirical literature on heterogeneity and labor markets. I apply the matching model to longitudinal linked data on employers and employees. I present both fixed and mixed model specifications of the equilibrium wage function predicted by the matching model, taking account of structural aspects implied by the learning process. Specifically, the learning process implies that the distribution of observed earnings is truncated, and that earnings residuals within a worker-firm match are a martingale. The latter implies a specific covariance structure for earnings residuals. In the most general empirical specification, I allow wage residuals to have a completely unstructured covariance within-match. I then fit the martingale structure to an estimate of the within-match residual covariance, and test the learning hypothesis. I find considerable support for these and other predictions of the matching model in the data.

The matching model is related to several established literatures. The first is the literature on search and matching with heterogeneous agents. A recent survey is Burdett and Coles (1999). In general, work in this area has focused on economies with heterogeneous workers and heterogeneous worker-firm matches.[1] Typically firms employ only a single worker. Thus there is no need to separately model heterogeneity at the firm and match level. In contrast, I model an economy in which firms employ many workers, and introduce an exogenous firm-specific technology that affects the marginal product of all its employees. A similar approach is taken by Postel-Vinay and Robin (2002), who present a dynamic search model with heterogeneous workers and firms that employ many workers. Unlike the model presented here, their workers are equally productive in every firm. Their work is exceptional, however, in its empirical application of the search model to longitudinal linked data.

A second related literature concerns learning in labor markets. Work in this area has provided new interpretations of important characteristics of labor market data, such as the returns to tenure and the relationship between earnings dispersion and labor market experience. The seminal Jovanovic (1979) matching model considered the case where identical workers and firms learn about the quality of a match. Flinn (1986) extends the Jovanovic (1979) model to the case of observably heterogeneous workers in discrete time. Moscarini (2002) develops a related model in continuous time. 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. (2002) 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.

The empirical portion of the paper draws heavily on recent work by Abowd et al. (1999), Abowd and Kramarz (1999), and Abowd et al. (2002), and the extensive statistical literature on mixed models. Abowd et al. (1999) and Abowd et al. (2002) develop and estimate linear wage models with fixed person and firm effects. Abowd and Kramarz (1999) describe but do not estimate the mixed model specification where person and firm effects are treated as random. Excellent references on mixed model theory are Searle et al. (1992) and McCulloch and Searle (2001).

The remainder of the paper is structured as follows. I present the matching model in Section 2. In Section 3, I develop the econometric specification. I give a detailed description of the data in Section 4. I present the results in Section 5 and conclude in Section 6.

2 A Matching Model with Heterogeneous Workers, Firms, and Worker-Firm Matches

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 $a \in [\underline{a}, \overline{a}]$. I refer to a as worker quality. Assume

$$a \sim F_a \text{ iid across workers} \tag{1}$$

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

where F_a is a probability distribution with zero mean, known to all agents. Assume 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, and there is no human capital accumulation over the life cycle. Unemployed workers receive income $h \in \mathbb{R}$ from home production.[2] 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. The price of output is normalized to 1. Firms can only produce output when matched with workers. Firms seek to maximize the expected net revenues of a match: the expected value of output minus a wage payment to the worker.

Firms are heterogeneous along two dimensions. They vary in their technology, which determines the marginal productivity of all their employees, denoted $b \in \left[\overline{b}, \underline{b}\right]$; and their cost of opening a vacancy, denoted $k > 0$. Assume

$$b, k \sim F_{b,k} \text{ iid across firms} \tag{2}$$

where $F_{b,k}$ is a probability distribution known to all agents. Without loss of generality, assume $E\left[b\right] = 0$. I refer to b as firm quality. Assume that firms know their own values of b and k, and that these are observed by the worker when the worker and firm meet. Both b and k are exogenous. Firms incur cost $\kappa\left(l\right)$ to hire l workers in the current period. Assume κ is continuous, increasing, and convex.

The marginal productivity of worker a when employed at firm b 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. Assume

$$c \sim N\left(0, \sigma_c^2\right) \text{ iid across matches.} \tag{3}$$

Let F_c denote the normal distribution function in (3). The normality assumption follows Jovanovic (1979) and others and is required to obtain a closed form solution for beliefs.

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

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

Let F_z denote the normal distribution function in (4). The worker and firm form beliefs about the value of c on the basis of 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, b, and k, and have common beliefs about c at every point in time.

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

$$q_\tau = \mu + a + b + c + e_\tau \tag{5}$$

[2] Assume for simplicity that h includes all search costs, the value of leisure, etc. Allowing h to vary across individuals does not change any of the 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 static excercise in Section 2.3.

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-specific idiosyncratic shock. Assume

$$e_\tau \sim N\left(0, \sigma_e^2\right) \text{ iid across matches and tenure.} \tag{6}$$

The linear production technology (5) generalizes that of Jovanovic (1979) to the case of heterogeneous workers and firms in a discrete time setting. Note that there are no aggregate shocks to productivity, and no human capital accumulation over the life cycle.[3] 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 (3), (4), and (6) are approximate. Of course the approximation can be made arbitrarily precise by appropriate choice of support.

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.

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,k) 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, k 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, k and current beliefs about c. The current period wage w_τ is simultaneously determined by a Nash bargain.

[3] Introducing a publicly-observable stochastic aggregate shock to productivity is a relatively straightforward generalization. Likewise, introducing a deterministic trend to individual productivity (i.e., an experience effect) presents no serious complication provided that it is observable by the 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 (5) can be viewed as a measure of output net of aggregate shocks and deterministically acquired human capital. The same is true of the equlibrium wage w_τ in (31) and the net value of output $q_\tau - w_\tau$. That is, in the more general model, the equilibrium wage (see Proposition 1) remains linear in person- and firm- specific components and in the posterior mean of beliefs, and is linear in the productivity shock and the experience effect.

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.

4. Firms open new vacancies υ at per-vacancy cost k.

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, k, c, z, and e_τ are governed by equations (1), (2), (3), (4), and (6). Recall that the worker's type a and the firm's type (b, k) are observed by both parties when the match forms. 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

$$m_1 = x\left(\frac{\sigma_c^2}{\sigma_c^2+\sigma_z^2}\right) \tag{7}$$

$$s_1^2 = \frac{\sigma_c^2\sigma_z^2}{\sigma_c^2+\sigma_z^2}. \tag{8}$$

In each subsequent period that the match persists, 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

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

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

Clearly the evolution of s_τ^2 is deterministic and does not depend on the value of the signals received. At each $\tau > 0$, $s_\tau^2 > s_{\tau+1}^2$. Equation (9) says that the posterior mean of beliefs m_τ is a precision-weighted average of the prior mean $m_{\tau-1}$ and the 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, it follows that each new signal is given successively smaller weight in the updating process. Asymptotically, beliefs converge to point mass at true match quality. That is,

$$\lim_{\tau\to\infty} m_\tau = c \tag{11}$$

$$\lim_{\tau\to\infty} s_\tau^2 = 0 \tag{12}$$

which is a standard result for Bayesian learnings with "correct" priors (see e.g., Blume and Easley (1998)).

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

$$m_\tau \sim N(0, V_\tau) \tag{13}$$

$$V_\tau = s_\tau^2 \sigma_c^2 \left(\frac{1}{\sigma_z^2} + \frac{\tau - 1}{\sigma_e^2} \right). \tag{14}$$

With a little algebra, one can 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

$$m_p | m_\tau \sim N(m_\tau, v_p) \tag{15}$$

$$v_p = \frac{s_\tau^4 (p - \tau)}{s_\tau^2 (p - \tau) + \sigma_e^2} \tag{16}$$

for any $p > \tau$. Note that (15) implies the posterior mean of beliefs is a martingale. Conditional on current information, expectations about future realizations of the random variable m_τ are equal to 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. Since the Nash bargain is efficient, in each period the match continues only if the expected joint surplus of the match is nonnegative. Expectations are taken with respect to tenure τ beliefs about match quality, conditional on the worker's type a and the firm's type (b, k). It follows that the equilibrium wage maps tenure τ information about the match $(a, b, k, m_\tau, s_\tau^2)$ into a payment from the firm to the worker.

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

$$J_\tau + \Pi_\tau \geq U + V. \tag{17}$$

When (17) is satisfied, the equilibrium wage w_τ solves the Nash bargaining condition

$$J_\tau - U = \delta [J_\tau + \Pi_\tau - U - V] \tag{18}$$

or equivalently,

$$(1 - \delta)(J_\tau - U) = \delta (\Pi_\tau - V) \tag{19}$$

where δ is the worker's exogenous share of the joint surplus.

2.2.1 The Worker's Value of Employment and Unemployment

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

$$J_\tau = w_\tau + \beta E\left[\max\left\{J_{\tau+1}, U\right\} | m_\tau, s_\tau^2\right]. \tag{20}$$

For notational simplicity, I omit the implicit conditioning on a, b, k. It is convenient to rewrite this value net of the value of unemployment, i.e.,

$$J_\tau - U = w_\tau - (1-\beta) U + \beta E\left[\max\left\{J_{\tau+1} - U, 0\right\} | m_\tau, s_\tau^2\right]. \tag{21}$$

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

$$U = h + \beta\pi \int_{\underline{b}}^{\overline{b}} J_0 dF_b^* + \beta (1-\pi) U \tag{22}$$

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

$$J_0 = E\left[\max\left\{J_1, U\right\} | 0, \sigma_c^2 + \sigma_z^2\right] \tag{23}$$

is the expected value of employment prior to observing the initial signal of match quality.

2.2.2 Vacancies and The Firm's Value of Employment

I now turn to 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,

$$\begin{aligned} \Pi_\tau &= E\left[q_\tau | m_\tau, s_\tau^2\right] - w_\tau + \beta E\left[\max\left\{\Pi_{\tau+1}, V\right\} | m_\tau, s_\tau^2\right] \\ &= \mu + a + b + m_\tau - w_\tau + \beta E\left[\max\left\{\Pi_{\tau+1}, V\right\} | m_\tau, s_\tau^2\right]. \end{aligned} \tag{24}$$

It follows that

$$\Pi_\tau - V = \mu + a + b + m_\tau - w_\tau - (1-\beta) V + \beta E\left[\max\left\{\Pi_{\tau+1} - V, 0\right\} | m_\tau, s_\tau^2\right]. \tag{25}$$

The production technology (5) implies that the firm's employees produce independently of one another. As a consequence, in each period the firm's decision to open vacancies is a static one. The number of hires today has no dynamic consequences for future hiring or productivity. When a firm opens υ vacancies, we can model the number l that are filled as a binomial process. The number of vacancies opened by a type (b, k) firm in each period solves[4]

$$\max_{\upsilon \in \mathbb{N}} \sum_{l=0}^{\upsilon} \binom{\upsilon}{l} \lambda^l (1-\lambda)^{\upsilon - l} \left[l \int_{\underline{a}}^{\overline{a}} \Pi_0 dF_a^* - \kappa(l) \right] - k\upsilon \tag{26}$$

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

$$\Pi_0 = E\left[\max\left\{\Pi_1, V\right\} | 0, \sigma_c^2 + \sigma_z^2\right] \tag{27}$$

[4] This approach to modeling vacancies follows Nagypal (2000).

is the expected present value of net revenues from a match before observing the initial signal of match quality.

Note that firm size (employment) is indeterminate. However, increasing and convex hiring costs κ guarantee the solution to (26) is well defined and the number of vacancies opened in any period by the firm is finite. I derive the average steady state employment of a type (b, k) firm in Appendix B.

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

2.2.3 The Equilibrium Wage

With expressions for the value functions in hand it is a simple matter to solve for the equilibrium wage. The equilibrium wage 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 in person- and firm-specific components, linear in m_τ, and independent of s_τ^2.*

Proof. Substituting (21) and (25) into the Nash bargaining condition (19) we obtain

$$\begin{aligned} & (1-\delta)\left\{w_\tau - (1-\beta)U + \beta E\left[\max\left\{J_{\tau+1} - U, 0\right\} | m_\tau, s_\tau^2\right]\right\} \\ = \; & \delta\left\{\mu + a + b + m_\tau - w_\tau - (1-\beta)V + \beta E\left[\max\left\{\Pi_{\tau+1} - V, 0\right\} | m_\tau, s_\tau^2\right]\right\}. \end{aligned} \tag{28}$$

It follows immediately from (19) that

$$(1-\delta) E\left[\max\left\{J_{\tau+1} - U, 0\right\} | m_\tau, s_\tau^2\right] = \delta E\left[\max\left\{\Pi_{\tau+1} - V, 0\right\} | m_\tau, s_\tau^2\right] \tag{29}$$

and thus

$$(1-\delta)\left\{w_\tau - (1-\beta)U\right\} = \delta\left\{\mu + a + b + m_\tau - w_\tau - (1-\beta)V\right\}. \tag{30}$$

Rearranging,

$$w_\tau = \delta\mu + \theta + \psi + \delta m_\tau \tag{31}$$

where

$$\begin{aligned} \theta &= \delta a + (1-\delta)(1-\beta)U && (32) \\ \psi &= \delta b - \delta(1-\beta)V && (33) \\ &= \delta b \end{aligned}$$

upon imposing the equilibrium condition $V = 0$. ■

The linear wage structure in (31) is at the core of the empirical strategy developed in Section 3. In related work, Abowd et al. (1999) present an empirical specification for earnings with fixed worker and firm effects. I thus refer to θ and ψ as empirical person and firm effects, respectively. The model thus far provides a behavioral interpretation to such empirical constructs. Equation (33) illustrates that the firm effect is simply the worker's share δ of the firm's contribution to match surplus. It reflects the firm's productivity parameter b. Rewriting equation (32) as $\theta = \delta\left(a - (1-\beta)U\right) + (1-\beta)U$ demonstrates that the person effect is the worker's share of his contribution to the joint surplus, plus compensation for forgoing his next-best alternative. It reflects the worker's productivity parameter a.

Equation (31) also demonstrates that the equilibrium wage is linear in the posterior mean of beliefs about match quality and independent of the posterior variance of beliefs. It is worthwhile relating this result to the Jovanovic (1979) equilibrium wage. In his model, workers and firms are *ex-ante* identical but matches are heterogeneous, and production occurs according to the continuous time analog of (5) with $a = b = 0$. The Jovanovic (1979) equilibrium wage is equal to expected marginal product, which in his case is also the posterior mean of beliefs about match quality. His result relies on the assumption that firms earn zero expected profit. Similar to Jovanovic's model, the equilibrium wage (31) is linear in expected marginal product, $\mu + a + b + m_\tau$ and in the posterior mean of beliefs about match quality, m_τ. A stronger result is that when workers capture all the quasi–rents associated with the match, i.e., as $\delta \to 1$ (so that firms earn zero expected profit), the equilibrium wage converges to $w'_\tau = \mu + a + b + m_\tau$. That is, the equilibrium wage converges to the expected marginal product of the match. In this sense, the Jovanovic (1979) equilibrium wage is a special case of (31).

2.2.4 The Separation Decision

Under the Nash bargaining framework, the separation decision is made jointly by the worker and firm. The match persists as long as the joint net surplus accruing to employment is non-negative. To characterize the separation decision, it is useful to introduce the Bellman equation that characterizes the joint value of employment, W_τ.

$$\begin{aligned} W_\tau &= \max\left\{J_\tau + \Pi_\tau, U + V\right\} \\ &= \max\left\{\mu + a + b + m_\tau + \beta E\left[W_{\tau+1} | m_\tau, s_\tau^2\right], U\right\} \end{aligned} \tag{34}$$

given the equilibrium condition $V = 0$. Lemma 2 establishes a fundamental property of W_τ, namely that it is increasing in the posterior mean of beliefs about match quality. The proof is rather lengthy, and relegated to Appendix A.

Lemma 2 *At each tenure $\tau > 0$, the joint value of employment, W_τ, is increasing in m_τ.*

The proof of Lemma 2 shows that the first argument of the max operator in (34) is strictly increasing in m_τ. Since the second argument is constant, it follows that the optimal separation policy is of the reservation-wage type.

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

Proof. Follows immediately from Lemma 2. ∎

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 (17) with equality. Equivalently, it is the value of m_τ that equates the arguments of the max function in (34). Thus $\bar{m}_\tau$ is implicitly defined by

$$\begin{aligned} \bar{m}_\tau &= U - \mu - a - b - \beta E\left[W_{\tau+1} | \bar{m}_\tau, s_\tau^2\right] \\ &= U - \mu - a - b - \beta \int W_{\tau+1} dF\left(m_{\tau+1} | \bar{m}_\tau, s_{\tau+1}^2\right) \end{aligned} \tag{35}$$

where $F\left(m_{\tau+1} | \bar{m}_\tau, s_{\tau+1}^2\right)$ is the normal distribution (15) with mean $\bar{m}_\tau$.

It is of considerable interest to characterize how $\bar{m}_\tau$ varies with tenure. The proof of Lemma 4 is in Appendix A.

Lemma 4 *The expected joint value of employment at $\tau+1$ given tenure τ information, $E\left[W_{\tau+1} | m_\tau, s_\tau^2\right]$, is increasing in s_τ^2 for all $\tau > 0$.*

Proposition 5 (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 $\tau > 0$.*

Proof. Suppose not, so that $\bar{m}_{\tau+1} < \bar{m}_\tau$. Consider

$$\bar{m}_{\tau+1} - \bar{m}_\tau = \beta \int W_{\tau+1} dF\left(m_{\tau+1} | \bar{m}_\tau, s_\tau^2\right) - \beta \int W_{\tau+2} dF\left(m_{\tau+2} | \bar{m}_{\tau+1}, s_{\tau+1}^2\right). \tag{36}$$

We know from the proof of Lemma 2 that $\int W_{\tau+1} dF\left(m_{\tau+1} | \bar{m}_\tau, s_\tau^2\right)$ is increasing in $\bar{m}_\tau$ (see equation (75) in Appendix A). Lemma 4 established that $\int W_{\tau+1} dF\left(m_{\tau+1} | \bar{m}_\tau, s_\tau^2\right)$ is increasing in s_τ^2. Since $s_{\tau+1}^2 < s_\tau^2$ for all $\tau > 0$ and $\bar{m}_{\tau+1} < \bar{m}_\tau$ by hypothesis, the right hand side of (36) is nonnegative. But the left hand side of (36) is negative, a contradiction. ■

The result in Proposition 5 is standard in equilibrium learning models, e.g. Jovanovic (1979) and Flinn (1986). It reflects the option value of employment. Early in the match, when beliefs about match quality are imprecise, workers and firms are willing to enter into employment relationships of low perceived quality because their point estimate m_τ may be inaccurate. Later, 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, $\lim_{\tau\to\infty} \bar{m}_\tau = (1-\beta) U - \mu - a - b$.

2.3 Comparative Statics

In this section, I explore how separation behavior varies across workers and firms. This is completely characterized by variation in the reservation value of beliefs about match quality. The key result is summarized in Proposition 6. The Proof is in Appendix A.

Proposition 6 *Increasing worker or firm quality reduces the reservation value of beliefs about match quality at each tenure. That is for each $\tau > 0$,*

$$\frac{\partial \bar{m}_\tau}{\partial a} < 0, \quad \frac{\partial \bar{m}_\tau}{\partial b} < 0. \tag{37}$$

This result is fairly intuitive. Consider the varying the productivity parameter b common to all employees of a firm. Variation in b has no effect on the firm's outside option V, since $V = 0$ in equilibrium. Thus variation in b only affects the value of remaining in the match, $J_\tau + \Pi_\tau$. Increasing b increases the wage w_τ (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 increases the joint value of remaining in the match, and at each tenure makes the worker and firm less selective about the set of acceptable values of match quality. Having found a "good" firm, 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 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 increases 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 8 in Appendix A). Nevertheless, search frictions ensure that increasing a increases the value of continuing the match more than the value of terminating it.[5] Furthermore, having found a "good" employee, the firm is less picky about whether or not she is a "good" match.

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 theoretical predictions when assessing the empirical specification.

First and foremost, the model predicts that wages are linear in person- and firm-specific components. In keeping with the empirical literature (e.g., Abowd et al. (1999)), I have called these empirical person and firm effects, and denoted them θ and ψ. They are functions of the random variables a, b, and k, and thus are random variables themselves.

Equilibrium wages are also linear in the posterior mean beliefs about match quality m_τ. This has a number of implications for the equilibrium distribution of wages and their evolution within a worker-firm match. First, since m_τ is a normally distributed random variable, conditional on the person and firm effects, equilibrium wages are as well. Second, 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. Since beliefs evolve according to Bayes' rule, m_τ is a martingale. Thus the model predicts that within a worker-firm match, wages are also a martingale.[6] The martingale property is common to most learning models, see e.g. Farber and Gibbons (1996) and Gibbons et al. (2002).[7] Its

[5]That is, using the result of Lemma 8 in Appendix A, it is easy to show that

$$\frac{\partial J_\tau}{\partial a} > \frac{\partial U}{\partial a}. \tag{38}$$

[6]Due to the selection process that terminates a match if $m_\tau < \bar{m}_\tau$, *observed* wages within a worker-firm match are a submartingale.

[7]In these two papers, firms and matches are homogeneous. Workers vary in their ability, which is unknown

econometric implications are discussed in Section 3.4. However, the martingale structure also has a number of economic consequences. First, recalling the definition of m_τ in equation (9), shocks to beliefs about match quality (z and e_τ) are permanent. Within a worker-firm match, these are the only shocks to wages. Thus wage shocks are permanent. Second, on average, wage shocks diminish with tenure. To see this, recall 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 as agents learn about match quality, each successive signal (shock) receives smaller weight in the updating process. Asymptotically, new signals receive zero weight. Third, within a worker-firm match the variance of earnings increases with tenure. This arises because the variance of m_τ increases with the number of observed signals. This may seem at odds with the notion that beliefs about match quality become increasingly precise with tenure. However, it is important to 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.[8] That the variance of earnings increases with tenure is broadly consistent with the empirical observation that the variance of earnings increases with labor market experience (see e.g., Mincer (1974)).

Section 2.3 presented comparative statics that characterize how separation behavior behavior varies with worker and firm quality, a and b. The empirical exercise that follows focuses on estimation of the empirical person and firm effects θ and ψ. Thus it is of interest to obtain predictions regarding the relationship between separation behavior and θ and ψ. It is a simple matter to show $\frac{\partial a}{\partial \theta} > 0$ and $\frac{\partial b}{\partial \psi} > 0$. Combining this with (37), at each $\tau > 0$ we have

$$\frac{\partial \bar{m}_\tau}{\partial \psi} < 0 \tag{39}$$

$$\frac{\partial \bar{m}_\tau}{\partial \theta} < 0. \tag{40}$$

Note that lower values of $\bar{m}_\tau$ are on average associated with longer job duration. Thus equation (39) implies that on average, jobs last longer at firms with larger firm effects ψ. A corollary is that firms with larger values of ψ experience less turnover than firms with smaller values. Since firms with large firm effects are better able to retain workers, *ceteris paribus* these firms have larger employment at any point in time than firms with smaller values. This is consistent with Brown and Medoff (1989) and others who find that conditional on observable worker and firm characteristics, larger firms pay higher wages than smaller firms. Abowd et al. (1999) find that estimated firm-size wage effects are well explained by firm-size category average person and firm effects.[9]

---

to either worker or firm. All agents in the economy observe signals of the worker's ability, and update their beliefs using Bayes' rule. Thus, individual earnings are a martingale, both within and between worker-firm matches. Farber and Gibbons (1996) test this hypothesis using data from the NLSY, with mixed results.

[8] Intuitively, the variance of beliefs s_τ^2 declines with tenure 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.

[9] However, they note that person effects are much more important in explaining firm-size wage effects than are firm effects.

Similar predictions arise from equation (40). On average, workers with large person effects enjoy longer job duration and change jobs less often. Lillard (1999) finds a comparable result in NLSY data.[10] Combining (39) and (40), we should expect to find a positive duration-weighted correlation between estimates of θ and ψ. Abowd et al. (2003) find evidence consistent with this prediction in LEHD data. Abowd et al. (2002) find the reverse in France and in the State of Washington.

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.[11] The result stems from several observations. First, larger values of θ and ψ are associated with higher wages and longer expected duration. Second, conditional on θ and ψ, better matches last longer and are associated with larger values of m_τ and hence higher wages. This effect is reinforced by monotonicity in tenure of the reservation level of beliefs about match quality.

Finally, for the empirical specification that follows, it is important to note that wages and employment duration are simultaneously determined. The posterior mean of beliefs about match quality m_τ enters the equilibrium wage, but the employment relationship only continues as long as $m_\tau \geq \bar{m}_\tau$. Thus the observed distribution of earnings is truncated: earnings outcomes are only observed if $w_\tau \geq \mu + \theta + \psi + \bar{m}_\tau$.

3 Empirical Specification

The empirical specification is based primarily on the equilibrium wage function (31). It takes account of the wage and mobility dynamics implied by the learning process. For clarity, I develop the empirical specification in stages.

Abowd et al. (1999), Abowd et al. (2002), and others have estimated earnings models with fixed person and firm effects. Since the equilibrium wage (31) is linear in person- and firm- specific components I adopt a similar approach. However, I depart from earlier work in two important respects.

The first departure is to focus primarily on a mixed model specification, where the person and firm effects are treated as random. There are a number of compelling reasons to do so. First, as shown in Section 3.2.2, least squares estimates of the fixed effects are a special

[10] Lillard (1999) estimates simultaneous wage and job turnover hazard 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 equation: higher values of the person-wage effect are associated with a reduced turnover hazard. Using a similar specification, Stinson (2002) obtains the same result in 1990 SIPP data, but finds the reverse in 1996 SIPP data.

[11] 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). Recent studies using longitudinal linked data include Dostie (2002) and Stinson (2002).

In the context of this debate, the learning 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" human capital since productivity is constant over the duration of the match. Nevertheless, it has value: it takes time to accumulate, and is lost when the match terminates.

case of mixed model estimates. Second, the theoretical person and firm effects θ and ψ are random variables. This suggests treating their empirical counterparts similarly. Third, with a large number of person and firm effects to estimate, it is their distribution that is of primary interest. Their realizations for specific workers and firms are of secondary importance. A mixed model specification estimates parameters of the (assumed) distribution of the random effects. One can also recover estimates of the realized values of the random effects, called Best Linear Unbiased Predictors (BLUPs). Fourth, the data can be considered a random sample of workers and firms from a larger population, and the person and firm effects a random sample from a larger population of values. When making inferences about a population of effects from which those in the data are considered a random sample, Searle et al. (1992) argue in favor of treating the effects as random. Finally, a mixed model specification permits out-of-sample prediction of person and firm effects.

The second important departure from earlier empirical work is to explicitly account for the structure implied by the learning process on earnings residuals. Recall the equilibrium wage function (31), into which θ, ψ, and δm_τ enter linearly. If we ignore for a moment the selection process that terminates an employment match whenever $m_\tau < \bar{m}_\tau$, it is a simple matter to show $E(m_\tau \theta) = E(m_\tau \psi) = 0$ for all $\tau > 0$. That is, signals of match quality are drawn independently of the person and firm effects. Furthermore, m_τ is a normally distributed random variable with zero expected value in the population, i.i.d across matches, and satisfying

$$E\left[m_\tau m_{\tau'}\right] = V_\tau \text{ for } \tau \leq \tau' \tag{41}$$

within matches, where V_τ is the unconditional variance of m_τ defined in equation (14).[12] Taken together, these facts suggest treating the term δm_τ in (31) as a normally distributed statistical residual with a within-match covariance structure defined by 41.

Of course the learning model of the previous Section predicts a selection process that complicates matters slightly. Since worker-firm matches terminate when $m_\tau < \bar{m}_\tau$, the observed distribution of earnings residuals δm_τ (and hence earnings) is truncated. Furthermore, since $\frac{\partial \bar{m}_\tau}{\partial \theta} < 0$ and $\frac{\partial \bar{m}_\tau}{\partial \psi} < 0$, the selection process induces a negative correlation between m_τ and the person and firm effects. I correct for truncation in the distribution of observed earnings residuals using standard methods. The corrected residual is uncorrelated with the person and firm effects by construction. The correction yields consistent estimates of the effects of interest.

Let $i = 1, ..., N$ index workers and $j = 1, ..., J$ index firms. The empirical specification for earnings is

$$w_{ijt} = \mu + x_{it}'\beta + \theta_i + \psi_j + \varepsilon_{ijt} \tag{42}$$

where w_{ijt} is a measure of earnings; μ is the grand mean of earnings; x_{it} is a vector of observable time-varying individual characteristics;[13] β is a parameter vector; θ_i is the pure

[12] The fact that $Cov(m_\tau, m_{\tau'}) = Var(m_\tau)$ for $\tau \leq \tau'$ is a standard property of Bayesian learning. The only information common to m_τ and $m_{\tau'}$ is the set of signals observed through tenure τ. Hence their covariance is the variance of the common signals. This is discussed further in Section 3.4.

[13] There are no explicitly time-varying covariates in the theoretical model of the previous Section. However, as noted in footnote 3, if we generalize the model to include deterministic human capital accumulation and publicly-observable stochastic aggregate productivity shocks, the equilbrium wage is linear in θ, ψ, m_τ, an additive experience effect, and time effects. The vector x_{it} includes a polynomial in experience (interacted with sex) and time effects. There is no conceptual leap required to include these in (42).

person effect; ψ_j is the pure firm effect of the firm j at which worker i was employed in t (denoted $j = J(i,t)$); and ε_{ijt} is a statistical residual. Note that in Section 2, τ was used to index tenure; here t indexes calendar time.[14] We can further decompose the pure person effect θ_i into components observed and unobserved by the econometrician as

$$\theta_i = \alpha_i + u_i'\eta \tag{43}$$

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.[15].

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 (42) and (43) in matrix notation, we have

$$w = X\beta + U\eta + D\alpha + F\psi + \varepsilon \tag{44}$$

where w 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 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 pure firm effects; and ε is the $N^* \times 1$ vector of residuals.

Abowd and Kramarz (1999) discuss fixed and mixed model specifications of equations like (44). A fixed model specification treats all the effects β, η, α, and ψ as fixed. A mixed model specification treats some of the effects as random. I consider the case where β and η are fixed, and α and ψ are random. I estimate both fixed and mixed model specifications in what follows, but focus primarily on the mixed model. For completeness, I discuss estimation of both fixed and mixed model specifications.

3.1 The Fixed Model

The fixed model is completely specified by (44) and the following assumptions on ε_{ijt} :

$$\begin{aligned} E\left[\varepsilon_{ijt}|i,j,t,x,u\right] &= 0 && (45)\\ E\left[\varepsilon\varepsilon'\right] &= \sigma_\varepsilon^2 I_{N^*} && (46)\end{aligned}$$

where I_{N^*} is the identity matrix of order N^*. The least squares estimator of β, η, α, and ψ solves the normal equations

$$\begin{bmatrix} X'X & X'U & X'D & X'F \\ U'X & U'U & U'D & U'F \\ D'X & D'U & D'D & D'F \\ F'X & F'U & F'D & F'F \end{bmatrix} \begin{bmatrix} \beta \\ \eta \\ \alpha \\ \psi \end{bmatrix} = \begin{bmatrix} X'w \\ U'w \\ D'w \\ F'w \end{bmatrix}. \tag{47}$$

[14] The inclusion of time-varying covariates x_{it} in (42) necessitates the additional calendar time index t. Since tenure and calendar time are in general related by a simple function, the tenure index τ is suppressed to avoid undue notational clutter.

[15] A similar decomposition could be done on the pure firm effect ψ_j. This is left for future research.

In the data described in Section 4, the cross product matrix on the left hand side of (47) is of sufficiently high dimension to preclude estimation using standard software packages. Instead, I obtain least squares solutions $\hat{\beta}$, $\hat{\eta}$, $\hat{\alpha}$, and $\hat{\psi}$ using the iterative conjugate gradient method of Abowd et al. (2002). Their algorithm exploits the sparse structure of the cross product matrix after blocking on connected groups of workers and firms.[16] The resulting estimates of α and ψ are not unique, since the design matrices D and F are not full rank. Abowd et al. (2002) discuss identification of the person and firm effects in detail. I apply their 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

In the remainder of this Section, I focus on a mixed model specification based on (44). The particular specification that I consider treats β and η as fixed, and the unobserved components of the person and firm effects α and ψ as random. The model is completely specified by (44) and the assumption

$$\begin{bmatrix} \alpha \\ \psi \\ \varepsilon \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 I_N & 0 & 0 \\ 0 & \sigma_\psi^2 I_J & 0 \\ 0 & 0 & R \end{bmatrix} \right). \tag{48}$$

It is worth noting that unlike the usual random effects specification considered in the econometric literature, (44) and (48) *do not* assume that the random effects are orthogonal to the design (X and U) of the fixed effects (β and η). Such an assumption is almost always violated in economic data. More general specifications than (48) are technically feasible though computationally demanding, e.g., allowing for a nonzero correlation between the person and firm effects. These are left for future research.

I present results for two different parameterizations of the residual covariance R.[17] The simplest of these is $R = \sigma_\varepsilon^2 I_{N^*}$. I estimate this specification primarily for comparison with the fixed model. The second parameterization of R allows for a completely unstructured residual covariance within a worker-firm match. 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

$$R = I_M \otimes W \tag{49}$$

where W is the $\bar{\tau} \times \bar{\tau}$ positive-semidefinite matrix of within-match residual covariances.[18]

[16]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. Abowd et al. (2002) develop a graph-theoretic algorithm for finding connected groups of workers and firms in longitudinal linked employer-employee data.

[17]Estimates of a variety of alternate residual covariance structures are available from the author on request. These include a number of common one and two parameter specifications for the within-match residual covariance, e.g., AR(1), AR(2), MA(1), MA(2), and ARMA(1,1).

[18]Symmetry and postive-semidefiniteness are the only restrictions imposed on W during estimation.

The extension to unbalanced data, where each match between worker i and firm j has duration $\tau_{ij} \leq \bar{\tau}$, is fairly 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.[19] S_{ij} selects those rows and columns of W that correspond to observed earnings outcomes in the match between worker i and firm j. In the unbalanced data case, the second parameterization of R is

$$R = I_M \otimes S_{ij}' W S_{ij}. \tag{50}$$

3.2.1 REML Estimation of the Mixed Model

Mixed model estimation is discussed at length in Searle et al. (1992) and McCulloch and Searle (2001). There are three principal methods that can be applied to estimate the variance components $\left(\sigma_\alpha^2, \sigma_\psi^2\right)$ and R : ANOVA, Maximum Likelihood (ML), and Restricted Maximum Likelihood (REML). ANOVA and ML methods are familiar to most economists; REML less so.[20] Since I apply the REML method in this application, it is worth giving it a brief treatment.

REML is frequently described as maximizing that part of likelihood that is invariant to the fixed effects (e.g., β and η). More precisely, REML is maximum likelihood on linear combinations of the dependent variable w, chosen so that the linear combinations do not contain any of the fixed effects. As Searle et al. (1992, pp. 250-251) show, these linear combinations are equivalent to residuals obtained after fitting the fixed effects via OLS. The linear combinations $k'w$ are chosen so that

$$k' \left(X\beta + U\eta\right) = 0 \;\; \forall \beta, \eta \tag{51}$$

which implies

$$k' \left[\begin{array}{cc} X & U \end{array} \right] = 0. \tag{52}$$

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

$$\begin{aligned} k' &= c' \left[I_{N^*} - \left[\begin{array}{cc} X & U \end{array} \right] \left(\left[\begin{array}{c} X' \\ U' \end{array} \right] \left[\begin{array}{cc} X & U \end{array} \right] \right)^{-} \left[\begin{array}{c} X' \\ U' \end{array} \right] \right] && (53) \\ &\equiv c' M_{XU} && (54) \end{aligned}$$

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 only $N^* - r$ linearly independent vectors k' satisfying (51).

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

$$S_{ij} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{array} \right].$$

[20] REML estimation of mixed models is commonplace in statistical genetics and in the plant and animal breeding literature. In recent years, REML has in fact become the mixed model estimation method of choice in these fields, superceding ML and ANOVA.

Define $K' = TM_{XU}$ with rows k' satisfying (51), and where K' and T have full row rank $N^* - r$. REML estimation is maximum likelihood on $K'w$. For $w \sim N(X\beta + U\eta, \mathbf{V})$ it follows that

$$K'w \sim N(0, K'\mathbf{V}K) \tag{55}$$

where $\mathbf{V} = DD'\sigma_\alpha^2 + FF'\sigma_\psi^2 + R$ is the covariance of earnings implied by (48). The REML log-likelihood (i.e., the log-likelihood of $K'w$) is therefore

$$\log L_{REML} = -\frac{1}{2}(N^* - r)\log 2\pi - \frac{1}{2}\log|K'\mathbf{V}K| - \frac{1}{2}w'K(K'\mathbf{V}K)^{-1}K'w. \tag{56}$$

REML estimates of the variance components and residual covariance have a number of attractive properties. First, REML estimates are invariant to the choice of K'.[21] Second, REML estimates are invariant to the value of the fixed effects (i.e., β and η). Third, in the balanced data case, REML is equivalent to ANOVA.[22] Under normality, it thus inherits the minimum variance unbiased property of the ANOVA estimator.[23] Finally, since REML is based on the maximum likelihood principle, it inherits the consistency, efficiency, asymptotic normality, and invariance properties of ML.

I estimate the variance components and residual covariance using the ASREML software package. ASREML implements the Average Information (AI) algorithm of Gilmour et al. (1995) to maximize the REML log-likelihood (56). The AI algorithm is a variant of Fisher scoring. AI uses a computationally convenient average of the expected and observed information matrices to compute parameter updates during iterative maximization of (56).[24]

Inference based on REML estimates of the variance components and parameters of the residual covariance is straightforward. Since REML estimation is just maximum likelihood on (56), REML likelihood ratio tests (REMLRTs) can be used. In most cases, REMLRTs are equivalent to standard likelihood ratio tests. The exception is testing for the presence of some random effect γ.[25] The null is $\sigma_\gamma^2 = 0$. Denote the restricted REML log-likelihood by $\log L_{REML}^*$. The REMLRT statistic is $\Lambda = -2(\log L_{REML}^* - \log L_{REML})$. Since the null puts σ_γ^2 on the boundary of the parameter space under the alternative hypothesis, Λ has a non-standard distribution. Stram and Lee (1994) show the asymptotic distribution of Λ is a 50:50 mixture of a χ_0^2 and χ_1^2. The approximate p-value of the test is thus $0.5(1 - \Pr(\chi_1^2 \leq \Lambda))$.

3.2.2 Estimating the Fixed Effects and Realized Random Effects

A disadvantage of REML estimation is that it provides no means for estimating the fixed effects β and η. Henderson, in Henderson et al. (1959) derived a system of equations that

[21] Subject to rows k' satisfying (51).

[22] The usual statistical definition of balanced data can be found in Searle (1987). Under this definitions, longitudinal linked data on employers and employees are balanced if we observe each worker employed at every firm, and all job spells have the same duration. Clearly, this is not the usual case.

[23] In contrast, ML estimators of variance components are biased since they do not take into account degrees of freedom used for estimating the fixed effects.

[24] The expected information matrix is the inverse of the negative expected Hessian of the REML log-likelihood (56). The observed information matrix is the inverse of the negative Hessian. The familiar Newton-Raphson method of maximizing the log-likelihood uses the observed information matrix to compute parameter updates. The Fisher scoring method uses the expected information for this purpose.

[25] For example, I test for the presence of an interaction effect between workers, i.e., a "match effect" γ_{ij}.

simultaneously yield the BLUE of the fixed effects and BLUP of the random effects for known values of the variance components and R. These equations have become known as the mixed model equations or Henderson equations. Define the matrix of variance components

$$G=\left[\begin{array}{cc}\sigma_{\alpha}^{2}I_{N} & 0\\ 0 & \sigma_{\psi}^{2}I_{J}\end{array}\right]. \tag{57}$$

The mixed model equations are

$$\left[\begin{array}{cc}\left[\begin{array}{c}X'\\ U'\end{array}\right]R^{-1}\left[\begin{array}{cc}X & U\end{array}\right] & \left[\begin{array}{c}X'\\ U'\end{array}\right]R^{-1}\left[\begin{array}{cc}D & F\end{array}\right]\\ \left[\begin{array}{c}D'\\ F'\end{array}\right]R^{-1}\left[\begin{array}{cc}X & U\end{array}\right] & \left[\begin{array}{c}D'\\ F'\end{array}\right]R^{-1}\left[\begin{array}{cc}D & F\end{array}\right]+G^{-1}\end{array}\right]\left[\begin{array}{c}\tilde{\beta}\\ \tilde{\eta}\\ \tilde{\alpha}\\ \tilde{\psi}\end{array}\right]=\left[\begin{array}{c}\left[\begin{array}{c}X'\\ U'\end{array}\right]R^{-1}w\\ \left[\begin{array}{c}D'\\ F'\end{array}\right]R^{-1}w\end{array}\right] \tag{58}$$

where $\tilde{\beta}$ and $\tilde{\eta}$ denote solutions for the fixed effects, and $\tilde{\alpha}$ and $\tilde{\psi}$ denote solutions for the random effects. In practice, of course, solving (58) requires estimates of R and G. Common practice, which I apply here, is to use the REML estimates $\tilde{G}$ and $\tilde{R}$ for this purpose.

The BLUPs $\tilde{\alpha}$ and $\tilde{\psi}$ have the following properties. They are *best* in the sense of minimizing the mean square error of prediction

$$E\left(\left[\begin{array}{c}\tilde{\alpha}\\ \tilde{\psi}\end{array}\right]-\left[\begin{array}{c}\alpha\\ \psi\end{array}\right]\right)'A\left(\left[\begin{array}{c}\tilde{\alpha}\\ \tilde{\psi}\end{array}\right]-\left[\begin{array}{c}\alpha\\ \psi\end{array}\right]\right) \tag{59}$$

where A is any positive definite symmetric matrix. They are *linear* in w, 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 (58) converge to the normal equations (47). Thus the mixed model solutions $(\tilde{\beta},\tilde{\eta},\tilde{\alpha},\tilde{\psi})$ converge to the least squares solutions $(\hat{\beta},\hat{\eta},\hat{\alpha},\hat{\psi})$. In this sense the least squares solutions are a special case of the mixed model solutions.

Equation (58) also makes clear the relationship between the mixed model and the usual "random effects" specification of Nerlove (1971) and others. In such models, the design of the random effects is assumed orthogonal to the design of the fixed effects. That is, $X'D=X'F=U'D=U'F=0$. Thus the off-diagonal blocks of (58) are zero, and we can solve for $\tilde{\beta}$ and $\tilde{\eta}$ independently of $\tilde{\alpha}$ and $\tilde{\psi}$. Furthermore, due to orthogonality the covariance of the random effects G is subsumed into R. The usual GLS estimator for $\tilde{\beta}$ and $\tilde{\eta}$ results.

3.3 Correcting for Residual Truncation

Under 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 residuals is truncated. Only earnings observations such that $m_{\tau}\geq\bar{m}_{\tau}$ are observed.

If we iterate forward on the definition of $\bar{m}_{\tau}$ in (35) we obtain

$$\bar{m}_{\tau}=-\mu A_{\tau}-\left[a-U\left(1-\beta\right)\right]A_{\tau}-bA_{\tau}-B_{\tau} \tag{60}$$

where

$$A_\tau = 1+\sum_{s=1}^{\infty}\beta^s\int_{\bar{m}_{\tau+1}}^{\infty}\int_{\bar{m}_{\tau+2}}^{\infty}\cdots\int_{\bar{m}_{\tau+s}}^{\infty}dF_{\tau+s}\cdots dF_{\tau+2}d\bar{F}_{\tau+1} \tag{61}$$

$$B_\tau = \sum_{s=1}^{\infty}\beta^s\int_{\bar{m}_{\tau+1}}^{\infty}\int_{\bar{m}_{\tau+2}}^{\infty}\cdots\int_{\bar{m}_{\tau+s}}^{\infty}m_{\tau+s}dF_{\tau+s}\cdots dF_{\tau+2}d\bar{F}_{\tau+1} \tag{62}$$

where $F_\tau = F\left(m_\tau|m_{\tau-1}, s^2_{\tau-1}\right)$ and $\bar{F}_\tau = F\left(m_\tau|\bar{m}_{\tau-1}, s^2_{\tau-1}\right)$. In keeping with the empirical discussion, I now add i and j subscripts to the posterior mean of beliefs about match quality and its reservation value. To correct earnings residuals for truncation, I approximate (60) by

$$\bar{m}_{ij\tau} \approx -\mu_\tau - \zeta_{i\tau} - \xi_{j\tau}. \tag{63}$$

Since $m_{ij\tau} \sim N(0, V_\tau)$, under the approximation (63) the marginal probability of observing the earnings outcome $w_{ij\tau}$ is

$$\begin{aligned}\Pr\left(m_{ij\tau} \geq \bar{m}_{ij\tau}\right) &= 1-\Phi\left(\frac{-\mu_\tau - \zeta_{i\tau} - \xi_{j\tau}}{V_\tau^{1/2}}\right)\\ &= \Phi\left(\frac{\mu_\tau + \zeta_{i\tau} + \xi_{j\tau}}{V_\tau^{1/2}}\right)\end{aligned} \tag{64}$$

where Φ is the standard normal CDF. Then we have

$$\begin{aligned}E\left[w_{ijt}|m_{ij\tau} \geq \bar{m}_{ij\tau}\right] &= \mu + x'_{it}\beta + \theta_i + \psi_j + V_\tau^{1/2}\frac{\phi\left(\frac{\mu_\tau+\zeta_{i\tau}+\xi_{j\tau}}{V_\tau^{1/2}}\right)}{\Phi\left(\frac{\mu_\tau+\zeta_{i\tau}+\xi_{j\tau}}{V_\tau^{1/2}}\right)}\\ &= \mu + x'_{it}\beta + \theta_i + \psi_j + V_\tau^{1/2}\lambda_{ij\tau}\end{aligned} \tag{65}$$

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

I perform a simple truncation correction based on (64) and (65). I estimate a continuation probit at each tenure level with random person- and firm-specific mobility effects.[26] The probits are estimated by Average Information REML applied to the method of Schall (1991).[27] 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 term for each observation. To correct for truncation in the distribution of earnings, I include $\tilde{\lambda}_{ij\tau}$ as an additional time-varying covariate in the earnings equation (42).

3.4 The Learning Hypothesis

Having discussed fixed and mixed model estimation in some detail, I now turn to a testable hypothesis of the matching model. In the matching model, agents update their beliefs about

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

[27] The Schall (1991) method extends standard methods for estimating generalized linear models to the random effects case. The basic idea is to perform REML on a linearization of the link function Φ. The process requires an iterative reweighting of the design matrices of fixed and random effects in the linearized system, see Schall (1991) for details.

match quality using Bayes' Rule. Bayesian learning implies a specific structure for the within-match residual covariance W. Since the empirical residual is $\varepsilon_{ijt} = \delta m_{ij\tau}$, we can write $W = \delta^2 V$, where V is the $\bar{\tau} \times \bar{\tau}$ within-match covariance of the vector of belief terms $[m_{ij1} \ \cdots \ m_{ij\bar{\tau}}]$.[28] Overlaying V with classical measurement error as in Farber and Gibbons (1996), Bayesian learning implies

$$V = \begin{bmatrix} V_1 + \sigma_u^2 & V_1 & V_1 & \cdots & V_1 \\ V_1 & V_2 + \sigma_u^2 & V_2 & \cdots & V_2 \\ V_1 & V_2 & V_3 + \sigma_u^2 & \cdots & V_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ V_1 & V_2 & V_3 & \cdots & V_{\bar{\tau}} + \sigma_u^2 \end{bmatrix} \tag{66}$$

where σ_u^2 is the variance of measurement error, and where the elements V_τ of V are the unconditional variance of $m_{ij\tau}$ in (14).

Some aspects of V are worthy of note. First, concentrating on the lower triangle of (66), off-diagonal elements within each column are equal. This is a property of the covariance of any martingale process, and is quite intuitive. Elements of column τ are $Cov\left(m_{ij\tau}, m_{ij\tau'}\right)$. In the lower triangle, $\tau \leq \tau'$. The common elements in $m_{ij\tau}$ and $m_{ij\tau'}$ are the signals of match quality 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: $Var\left(m_{ij\tau}\right) \equiv V_\tau$. A second aspect of note is that the diagonal elements of V are greater than off-diagonal elements in the same column due to measurement error. Finally, it can be shown that $V_{\tau+1} > V_\tau$. As mentioned in Section 2.4, 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. Thus V_τ increases from its prior value (zero) to its asymptotic value (σ_c^2) as tenure increases. Consequently in the lower triangle of (66), elements within a row increase in magnitude from left to right.

Whether or not earnings residuals have the structure implied by (66) is a testable hypothesis. 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 residual covariance. I test the learning hypothesis and recover the structural parameters and σ_u^2 using a two-step procedure. The first step is to obtain an estimate of the within-match residual covariance W. I use the unstructured estimate of the residual covariance $\tilde{W}$ obtained under the mixed model specification. An estimate of the covariance of elements of $\tilde{W}$ is provided by the relevant block of the REML Average Information matrix. Following Abowd and Card (1989) and Farber and Gibbons (1996), the second step is to fit the martingale covariance $\delta^2 V$ to $\tilde{W}$ by minimum distance.[29] This yields estimates of the structural parameters up to the factor

[28] In the unbalanced data case where a match between worker i and firm j lasts $\tau_{ij} \leq \bar{\tau}$ periods, the vector of belief terms is $\left[m_{ij1} \ \cdots \ m_{ij\tau_{ij}}\right]$. The residual covariance is $R = I_M \otimes \delta^2 S_{ij}' V S_{ij}$, where S_{ij} is the selection matrix defined earlier.

[29] Optimal minimum distance estimation, as discussed in Hansen (1982) and Chamberlain (1984), proved infeasible. The covariance of elements of $\tilde{W}$ was poorly conditioned, and did not invert. Instead I use a diagonal weight matrix, with elements equal to the natural logarithm of the number of observations contributing to each element of $\tilde{W}$. The data are highly unbalanced, with many more observations contributing to the estimation of elements in the upper-left corner of W than to elements of the lower-right corner. This weight

of proportionality δ^2. I test the learning hypothesis with the usual χ^2 test of overidentifying restrictions, using the test statistic of Newey (1985).[30]

3.5 Logs or Levels?

The matching model of Section 2 was developed in earnings *levels.* In keeping with this, the empirical specification has been developed thus far in levels also. However, it is customary to model earnings in logs, partly because the distribution of earnings is typically found to be approximately lognormal, but also to alleviate heteroskedasticity. Should the empirical specification proceed in earnings logs or levels? I opt to do both. I model earnings in levels to test the learning hypothesis and other predictions of the matching model, and model earnings in logs to facilitate comparison with earlier work.[31]

4 Data

Identifying the person and firm effects requires longitudinal linked data on employers and employees: data with repeated observations on both workers and firms. I use data from the Longitudinal Employer-Household Dynamics (LEHD) program database, under development at the U.S. Census Bureau. At present, the LEHD database spans eighteen states. Together, these states represent well over 50 percent of U.S. employment. In this paper, I use data from two of the eighteen participating states. The identity of the two states cannot be revealed for confidentiality reasons.

The LEHD data are administrative, constructed from quarterly Unemployment Insurance (UI) system wage reports. Every state in the U.S., 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 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

matrix gives greater weight to more precisely estimated elements of W. Weighting by a diagonal matrix of the variances of each element of $\tilde{W}$ yields similar results, as does equally weighted minimum distance.

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

[31]The log-linear specification can be interpreted as an approximation to the levels specification as follows. Rewrite (42) 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

$$\begin{aligned} \ln w_{ijt} &\approx \ln\mu + x'_{it}\frac{\beta}{\mu} + \frac{\theta_i}{\mu} + \frac{\psi_j}{\mu} + \frac{\varepsilon_{ijt}}{\mu} \\ &= \mu^* + x'_{it}\beta^* + \theta_i^* + \psi_j^* + \varepsilon_{ijt}^* \end{aligned} \tag{67}$$

where the first line of (67) uses $\ln(1+x) \approx x$, and where $\mu^* = \ln\mu$, $\beta^* = \beta/\mu$, $\theta_i^* = \theta/\mu$, $\psi_j^* = \psi/\mu$, and $\varepsilon_{ijt}^* = \varepsilon_{ijt}/\mu$.

employers required to file UI system wage reports — that is, all employment covered by the UI system in the eighteen participating states. For the two states used in this analysis, the data span the first quarter 1990 through the fourth quarter 1999.[32]

Individuals are uniquely identified in the data by a Protected Identity Key (PIK). Employers are identified by a state unemployment insurance account number (SEIN). The UI wage records themselves contain only very limited information: PIK, SEIN, and quarterly earnings. Quarterly earnings are a measure of total compensation, including 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)). In the LEHD database, the UI wage records are integrated with internal Census Bureau data to obtain additional demographic characteristics. Such characteristics include sex, race, and date of birth.

Though the underlying data are quarterly, they are aggregated to the annual level for estimation. All preliminary data processing is done on the quarterly records.

4.1 Sample Construction

Before discussing the estimation sample, variables, and the imputation of missing data, it is necessary to develop several concepts. The first concept is that of a *dominant employer.* A dominant employer is identified for each individual in each year. Individual i's dominant employer in year t is the employer at which i's actual UI earnings were largest in t. About 87 percent of the UI system wage records correspond to employment at a dominant employer. The second concept is *full quarter employment.* Individual i 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 $q+1$.

The analysis sample is restricted to full-time private sector employees at their dominant employer, between 25 and 65 years of age, who had no more than 44 employers in the sample period,[33] 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, at firms with at least five employees in 1997. The resulting analysis 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 an analysis sample of 49.3 million annual records.

Using the method of Abowd et al. (2002), it is feasible to estimate the fixed model on the entire analysis sample. Unfortunately, estimating the mixed model on the full sample remains computationally intractable. Drawing an appropriate random sample of observations for estimating the mixed model is not trivial. Obtaining precise estimates of the variance components and BLUPs requires a highly connected sample of workers and firms. In a small simple random sample of individuals, there may not be sufficient connectivity to be confident that the person and firm effects are well identified. For this reason I develop a dense sampling algorithm, described in detail in Appendix C. The basic idea behind the algorithm

[32]For one of the two states, the data series begins earlier. All estimation is done on the pooled states for the years 1990-1999.

[33]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.

is to sample 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. In Appendix C, I show that the resulting sample has all the properties of a simple random sample of workers employed in the reference period (i.e., 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 1 percent dense random samples of workers employed in 1997 using this algorithm. Each worker is connected to at least $n = 5$ others.[34] I label the two samples Dense Sample 1 and Dense Sample 2. All mixed model estimation is performed on Dense Sample 1. Dense Sample 2 is used for model validation. For comparison, I also draw a 1 percent simple random sample of workers employed in 1997. Table 1 presents connectedness properties of the full analysis sample, the two dense samples, and the simple random sample. The full analysis sample is highly connected: the largest connected group contains 99.06 percent of jobs. The dense samples remain 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 full analysis sample and 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.

4.2 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).[35]

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. (1998); 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 from the author 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.

[34] The other parameters used to draw the dense samples, defined in Appendix C, are $m = 0.5$ and $p = 0.004$.

[35] As described in Section 4.2.3, initial experience is set to zero in this case.

4.2.1 Real Annualized Earnings

The dependent variable for the earnings regressions is 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, the full-quarter earnings measure is reported UI system 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 real 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.[36] 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 annualized earnings.

4.2.2 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. For the earnings model I collapse the education measure to five categories: Less than high school, High school graduate, Some college or vocational training, Undergraduate degree, and Graduate or professional degree.

4.2.3 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

[36]The 80 percent cutoff rule was chosen to reduce error in the construction of the full quarter earnings measure. To determine the cutoff, for each quarter I computed real average full quarter earnings in the four previous and subsequent quarters (a nine quarter moving window). For full quarter employees, the median ratio of real earnings in quarter t to real average full quarter earnings in the nine quarter window around t was 0.8.

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.

4.2.4 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 analysis sample.

For non-left-censored spells, tenure is set to 1 in the first quarter that there is a UI system 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.

4.2.5 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.

4.3 Characteristics of the Samples

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

Figures 1 and 2 confirm these properties of the samples. Figure 1 plots the yearly time series of average real annualized earnings in each of the samples. The same trend is apparent in all four samples. The dense and simple random samples are virtually indistinguishable. However, average real annualized earnings are greater in each year in the point-in-time samples than in the full analysis sample, as expected. Figure 2 plots the yearly time series of employment in each of the samples. By construction, employment in the point-in-time samples is greatest in 1997.[37] As a consequence, the dense and simple random samples are indistinguishable, but their employment series differ somewhat from that of the full analysis sample.

5 Results

The econometric specification of Section 3 is estimated on each of the completed data implicates. Statistics reported in this section combine those obtained on the three implicates using 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 Estimated Variance Components and Model Fit

5.1.1 The Earnings Model

Table 3 presents estimates of the variance components and a summary of model fit for the four earnings models, estimated on the natural logarithm of annualized earnings. The four models are: 1) the fixed model; 2) a mixed model with random person and firm effects and a spherical error ($R = \sigma_\varepsilon^2 I_{N*}$); 3) a mixed model with random person, firm, and match effects and a spherical error;[38] and 4) a mixed model with random person and firm effects and an unstructured within-match residual covariance W. Mixed model estimates are presented both with and without the truncation correction.[39] In the case of the fixed model, the reported "variance components" are the sample variance of the estimated person and firm effects. The estimated variance components have a fairly straightforward interpretation. Conditional on all other effects, a one standard deviation increase in the value of the person effect α_i 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 Table 3 is that applying the truncation correction induces virtually no change in the estimated variance components. This is not overly surprising given the small selection/truncation bias typically found in earnings data. The second item of note is that in each of the models, the variance of the person effect (σ_α^2) is considerably larger than the variance of the firm effect (σ_ψ^2). That is, in the log earnings dimension, individuals are more heterogeneous than firms. This is consistent with Abowd et al. (1999) and others, who find unobserved individual heterogeneity to be a more important determinant of earnings

[37]Recall these are random samples of individuals employed in 1997. There are no individuals in the dense and simple random samples who were not employed in that year.

[38]This model is included to test for the presence of an interaction effect between person and firm effects.

[39]I have not applied the truncation correction to the fixed model. Doing so would require computing realized random effects $\tilde{\zeta}_{i\tau}$ and $\tilde{\xi}_{j\tau}$ for each worker and firm in the full analysis sample (about 25 million effects total). Although this is technically feasible, it is extraordinarily demanding from a computational standpoint. I leave this exercise for future research.

than unobserved firm heterogeneity. In Table 3, the fixed model yields the largest estimate of σ^2_α (0.290), but one of the smallest estimates of σ^2_ψ (0.077). These values are slightly larger than those estimated by Abowd et al. (2002) for France and the State of Washington. The mixed model with random person and firm effects and a spherical error yields a slightly smaller estimate of σ^2_α (0.23), and an estimate of σ^2_ψ twice that obtained under the fixed model. Relaxing the mixed model specification to allow for a match effect or an unstructured within-match residual covariance reduces the estimated variance of the firm effect to levels comparable to the fixed model, and reduces the estimated variance of the person effect to around 0.175. Under the most general specification (W unrestricted), 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 effect increases earnings by 0.28 log points. These are very close in magnitude to Abowd et al. (2002).

Table 3 also reports some measures of model fit. Not surprisingly, the mixed model with the unrestricted within-match residual covariance obtains the best fit by all in-sample measures (REML log-likelihood, AIC, BIC). To obtain a measure of out-of-sample fit, I solve the mixed model equations (58) on Dense Sample 2, using the variance components $\tilde{G}$ and residual covariance $\tilde{R}$ estimated on Dense Sample 1. To facilitate comparison with models estimated on earnings levels, the dependent variable is first scaled to have unit variance (and parameters are re-scaled accordingly). I compute the prediction error for each observation, and report its variance in Table 3.[40] By this measure, the mixed model specification with a match effect has the smallest out-of-sample prediction error. The mixed model with W unrestricted has the largest, though only slightly larger than the simple mixed model.

I perform a REMLRT for the presence of a match effect in the mixed model with a spherical error. The p-value for this test is extremely small ($< 10^{-8}$), so we reject the null of no match effect. I do not test the match effect specification against the specification with W unrestricted, since these hypotheses are not nested. The AIC and BIC statistics indicate the model with W unrestricted fits the data better than the match effect specification.

Table 4 reproduces that in Table 3 for models estimated on earnings levels. To put the parameter estimates on a recognizable scale, the dependent variable is scaled to have unit variance. Parameter estimates exhibit the same stylized facts as those obtained on earnings logs: the truncation correction has almost no influence on the estimated variance components; the estimates of σ^2_α are considerably larger than estimates of σ^2_ψ; the estimate of σ^2_α is largest under the fixed model; and relaxing the mixed model specification to allow for a match effect or unstructured within-match residual covariance reduces the magnitude of the estimated variance components. Under the mixed model with W unrestricted, a one standard deviation increase in the value of the person effect α_i increases real annualized earnings by 0.63 standard deviations (about $32,500 1990 Dollars). Employment at a firm with a value of ψ_j one standard deviation above the mean increases real annualized earnings by 0.21 standard deviations (about $10,600 1990 Dollars). Clearly, unobserved individual heterogeneity is a much more important determinant of earnings variation than unobserved firm heterogeneity. Nevertheless, the effect of unobserved firm heterogeneity is very economically significant.

[40]For models with $R = \sigma^2_\varepsilon I_{N^*}$, the prediction error associated with an observation is the estimated residual. For the model with the unstructured residual covariance, prediction error is the difference between the estimated residual and its expected value given the other within-match residuals. The conditional expectation is calculated using standard formulae for the multivariate normal distribution.

As in Table 3, the mixed model with unstructured W obtains the best fit to earnings levels using in-sample measures. Using the out-of-sample measure, its fit is the same as the simple mixed model with spherical error. By this measure, the mixed model with the match effect obtains the best fit. Prediction errors are nevertheless considerably more variable than under the log specification. It is no surprise that the linear model fits the logarithm of earnings better than it does earnings levels.

Tables 5-8 present correlations among the estimated effects for the various specifications. Table 5 presents correlations for models estimated on log earnings without the truncation correction. There is only slight variation across specifications. Of the estimated effects, the pure person effect θ_i is most highly correlated with log earnings: between 0.74 and 0.83, depending on the specification. The portion of θ_i corresponding to unobserved heterogeneity (α_i) is much more highly correlated with earnings than the observable component ($u_i\eta$). Correlations between the firm effect and log earnings are considerably lower: between 0.45 and 0.54 depending on the specification. The match effect is highly correlated with log earnings (0.62) in the specification that includes one.

Recall that the matching model predicts a positive correlation between the pure person effect θ_i and the pure firm effect ψ_j. In the fixed model and the mixed model with no match effect and spherical error, there is a small positive correlation between θ_i and ψ_j (about 0.03). When the mixed model is relaxed to allow for a match effect or an unstructured within-match residual covariance, the correlation between θ_i and ψ_j increases markedly to around 0.22. This is consistent with the matching model's prediction that larger values of θ_i and ψ_j are associated with longer job duration. It is also consistent with conventional wisdom that "good" workers sort themselves into "good" firms.

Table 6 presents correlations among estimated effects for the mixed model specifications on log earnings after correcting for truncation. Once again, correcting for truncation has little effect on the results. The truncation correction term $\beta_\lambda\lambda_{ij\tau}$ has a small positive correlation with earnings (0.056 in each specification), exhibits a small positive correlation with θ_i (about 0.11), and a small negative correlation with ψ_j (between -0.06 and -0.13 depending on specification).

Tables 7 and 8 reproduce the information in Tables 5 and 6 for models estimated on earnings levels. Again, there is little change in the estimates when correcting for truncation. The person effect exhibits a slightly higher correlation with earnings levels than with logs; the reverse is true for the firm effect. In the fixed model, the correlation between θ_i and ψ_j remains small and positive. In all the mixed model specifications, including the model with no match effect and a spherical error, the correlation between θ_i and ψ_j is larger: between 0.17 and 0.27 depending on the specification.

5.1.2 The Continuation Probit

Table 9 presents parameter estimates for the probit models used to correct for truncation. The longest observed tenure in Dense Sample 1 was 21 years.[41] I estimate eight continuation

[41] Only ten years of earnings data are used in the estimation. In one state, the data series is longer. Consequently, in that state there are individuals with (true) tenure values in excess of 10 years (maximum value 14 years). Other cases where tenure exceeds 10 years are due to the tenure imputation of left-censored job spells.

probits: one each for 1-2 years of tenure, 3-4 years of tenure, 5-6 years, 7-8 years, 9-10 years, 11-12 years, 13-14 years, and 15+ years. In each of the these the residual variance is scaled to unity. At all tenure levels the variance of the firm effect far exceeds the variance of the person effect. This is in contrast to the earnings models, and suggests greater heterogeneity in the separation policies of firms than of workers. The variance of the person effect increases from a very small value (0.001) at 1-2 years of tenure, to its maximum 0.31 at 5-6 years of tenure, and then declines with further increases in tenure. Likewise, the variance of the firm effect increases from 0.09 at 1-2 years of tenure, increases to its maximum of 1.34 at 11-12 years, and then declines.

5.2 Testing the Learning Hypothesis

Estimates of the within-match residual covariance W are presented in Tables 10-13. Table 10 presents the estimate of W obtained on log earnings without correcting for truncation. Estimates in Table 10 exhibit a number of the properties of the martingale covariance structure overlaid with classical measurement error given in (66): in each column, the diagonal elements are larger than the off-diagonal elements, and elements increase in magnitude from left to right within a 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 10 consistently decline.

Table 11 presents the estimate of W obtained on log earnings after correcting for truncation. Once again, there is little difference between the estimates obtained with and without the truncation correction. A casual comparison of Tables 10 and 11 suggests that there is slightly less decay in the autocovariances corrected for truncation than without the correction.

In Tables 12 and 13 I present estimates of W obtained on earnings levels, with and without correcting for truncation. They are virtually identical. Once again, casual inspection indicates the estimates are consistent with the structure implied by the learning process. The diagonal elements are larger than off-diagonal elements within a column. Elements increase in magnitude from left to right within each row. Unlike the estimates obtained on earnings logs, there is little decline moving from lower-order to higher-order autocovariances, which is consistent with the martingale structure implied by the learning model. This is particularly true for the first 10 years of tenure.

I formally test whether the within-match residual covariance has the martingale structure predicted by the learning model. I fit (14) and (66) to estimates of W by minimum distance. Table 14 presents the estimates of structural parameters and p-values from the chi-squared test of over-identifying restrictions.[42] Recall that the structural parameters σ_u^2, σ_c^2, σ_z^2, and

[42] Rubin (1987) provides formulae for combining statistics with chi-squared distributions obtained on multiply-imputed data. Let d_m denote the test statistic from the m^{th} implicate, with an asymptotic χ_k^2 distribution. Let M denote the number of implicates, $\bar{d}_m$ the mean of the statistics d_m, and s_d^2 their variance. Define

$$\hat{r}_m = \frac{\left(1 + M^{-1}\right) s_d^2}{2\bar{d}_m + \left(4\bar{d}_m^2 - 2ks_d^2\right)^{1/2}} \tag{68}$$

and

$$\hat{v} = (M-1)\left(1 + \hat{r}_m^{-1}\right)^2. \tag{69}$$

σ_e^2 are only identified up to a factor of proportionality: the square of the bargaining strength parameter δ. The estimates in Table 14 are presented on the scale of the data, i.e., for $\delta = 1$.[43]

The first column of Table 14 presents the results obtained on log earnings. Given that estimates of W obtained with and without the truncation correction are virtually identical, it is no surprise that the minimum distance estimates are also. The estimated variance of measurement error (σ_u^2) and variance of match quality (σ_c^2) are both approximately 0.05. This is of similar magnitude to the variance of the firm effect in this model. The variance of the initial signal of match quality (σ_z^2) and of production outcomes (σ_e^2) are considerably smaller: 0.02 and 0.004, respectively. This implies learning about match quality is very rapid. Figure 3 plots the estimated posterior variance of beliefs about match quality (s_τ^2) at each tenure. Upon receipt of the initial signal x, the posterior variance drops from the prior level ($\sigma_c^2 = 0.05$) to less than 0.02; after observing one production outcome it drops well below 0.01. However, these results should be viewed cautiously, since the p-value of the χ^2 test of overidentifying restrictions is 0.001. Though the log specification is only an approximation to the matching model, this is nevertheless evidence against the learning hypothesis.

Column (2) of Table 14 presents results estimated on earnings levels. The estimated variance of match quality is quite large: about 0.38, which is more than a third of the variance of annual earnings. The estimate of σ_z^2 (8.66) indicates the initial signal of match quality conveys virtually no information. However, learning is quite rapid once production outcomes are observed: the posterior variance of beliefs drops below 0.2 after observing one production outcome, and is about 0.1 after two (see Figure 4). On the basis of the p-value from the χ^2 test (0.002), we are inclined to reject the learning hypothesis. However, recall from Tables 12 and 13 that for the first 10 years of tenure, casual inspection indicated $\tilde{W}$ was highly consistent with the martingale covariance structure. In column (3) of Table 14 I present results estimated on only the first 10 rows and columns of $\tilde{W}$. These yield a considerably larger variance of match quality (0.58) and indicate that learning about match quality is quite slow (see Figure 4). Both with and without the truncation correction, the p-value of the χ^2 test is 0.026. Thus we reject the learning hypothesis at the 5 percent level, but fail to do so at the 1 percent level. Though far from conclusive evidence in favor of the matching model, these results certainly provide some support to the learning hypothesis. However, they also raise the possibility that imputing initial tenure for left-censored job spells may be the source of some inconsistencies between the data and the matching model. All tenure observations in excess of 14 years, and most in excess of 10 years, are the result of imputing tenure for left-censored job spells (see footnote 41).

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

$$\hat{D}_m = \frac{\frac{\bar{d}_m}{k} - \frac{M-1}{M+1}\hat{r}_m}{1 + \hat{r}_m} \tag{70}$$

has an asymptotic F distribution with k and $\left(1 + k^{-1}\right)\hat{v}/2$ degrees of freedom.

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

5.3 Additional Predictions From the Matching Model

Thus far we have seen several predictions of the matching model confirmed in the LEHD data: an earnings specification linear in person and firm effects fits the data very well;[44] we observe a positive correlation between the estimated person and firm effects; and limited evidence in favor of the learning hypothesis. Though the matching model predicted the distribution of earnings residuals is truncated, this appears to have little influence on parameter estimates. I now address two other predictions: that larger values of θ_i and ψ_j should on average be associated with longer job duration, and that firms with larger estimated firm effects should have greater employment on average.

To address the first of these predictions, I fit a fourth-order polynomial in job duration to the estimated person and firm effects. Right-censored spells are excluded from the regression. I focus on the effects estimated in the mixed model with the unstructured within-match residual covariance. Results from the other specifications are very similar and available on request. Figures 5 and 6 present the fitted curves. As the matching model predicted, larger values of θ_i and ψ_j are associated with longer duration. This is true on both logs and levels, with and without the truncation correction. 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 1997 employment to the estimated firm effects. Again, I focus on effects estimated from the mixed model with the unstructured within-match residual covariance. Results obtained on the other specifications are qualitatively similar. Figure 7 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

I presented a matching model with heterogeneous workers, firms, and worker-firm matches. The model generalizes the seminal Jovanovic (1979) model to the case of heterogeneous workers and firms. I showed that the equilibrium wage is linear in a person-specific component, a firm-specific component, and the posterior mean of beliefs about match quality. The matching model has numerous predictions for empirical person and firm effects, and earnings residuals. I then developed a mixed model specification for the equilibrium wage function that takes account of structural aspects of the learning process. I found considerable support for the various predictions of the matching model in UI wage data. A formal test of the learning hypothesis proved inconclusive.

Conventional wisdom suggests that “good” workers sort themselves into “good” firms. The matching model contains no explicit sorting mechanism. Nevertheless, it predicts that given time, good workers sort themselves into good firms. This arises simply because matches between bad workers and bad firms dissolve. Since agents need to enter a match to learn

[44] In a simple auxiliary regression of log earnings on covariates and the estimated person and firm effects from the simple mixed model, $R^2 = 0.92$; on earnings levels, $R^2 = 0.88$. Excluding covariates, the R^2 is apprximately 0.7 under both specifications.

one another's type, only by sampling a number of employment relationships do they eventually find a match worth pursuing. Empirically, the result is a positive duration-weighted correlation between person and firm effects. The data support this prediction.

The empirical analysis demonstrates that unobserved worker and firm heterogeneity are extremely important determinants of earnings. The estimated person and firm effects are highly correlated with earnings; much more so than observable characteristics such as education and labor market experience. Together, they explain about 70 percent of earnings variation. An important contribution of the matching model is that it yields an economic interpretation of these effects. They reflect worker and firm productivity, adjusted for the worker's bargaining strength and the value of each agent's outside option.

The paper suggests several fruitful areas for future work. One is to consider the case where a worker invests in productive human capital. On the empirical front, it would be of considerable value to refine the imputation of initial tenure for left-censored jobs.

A Appendix: Omitted Proofs

Proof of Lemma 2. Following Flinn (1986), consider a finite-period model with terminal date T. Abusing notation slightly, the joint value of employment in the terminal period is

$$W_T = \max\{\mu + a + b + m_T, U_T\} \tag{71}$$

where $U_T = h$ is the value of the worker's outside option in the terminal period. W_T is clearly increasing in m_T. The proof proceeds by induction. Take as the induction hypothesis that W_{t+1} is increasing in m_{t+1}, and I establish that W_t is increasing in m_t, $t < T$.

Consider the joint value of employment in period t :

$$\begin{aligned} W_t &= \max\left\{\mu + a + b + m_t + \beta E\left[W_{t+1}|m_t, s_t^2\right], U_t\right\} \\ &= \max\left\{\mu + a + b + m_t + \beta \int W_{t+1} dF\left(m_{t+1}|m_t, s_\tau^2\right), U_t\right\} \end{aligned} \tag{72}$$

where $F\left(m_{t+1}|m_t, s_\tau^2\right)$ is the normal distribution given in (15),

$$U_t = h + \beta\pi_{t+1} \int E\left[J_{t+1}|0, \sigma_c^2 + \sigma_z^2\right] dF_{b,t+1}^* + \beta\left(1 - \pi_{t+1}\right) U_{t+1} \tag{73}$$

is the value of the worker's outside option in t, π_{t+1} is the probability of drawing a new match in $t+1$, and $F_{b,t+1}^*$ is the period $t+1$ distribution of firm types among open vacancies.

We need to show W_t is increasing in m_t. Note U_t is independent of m_t, and the current period return to employment $\mu + a + b + m_t$ is strictly increasing in $m_{t,}$ so it is sufficient to establish that

$$\beta \int W_{t+1} dF\left(m_{t+1}|m_t, s_\tau^2\right) \tag{74}$$

is increasing in m_t. Period t beliefs affect (74) in two ways: via W_{t+1}, and via the transition distribution. The effect via W_{t+1} is straightforward: W_{t+1} is increasing in m_{t+1} (by hypothesis), and m_{t+1} is increasing in m_t from (9). Thus W_{t+1} is increasing in m_t. The effect of an increase in m_t on the transition distribution is as follows. Consider $m_t' > m_t$. Then $F\left(m_{t+1}|m_t', \sigma^2\right) \leq F\left(m_{t+1}|m_t, \sigma^2\right)$ for any $\sigma^2 < \infty$. That is, $F\left(m_{t+1}|m_t', \sigma^2\right)$ first order stochastically dominates $F\left(m_{t+1}|m_t, \sigma^2\right)$. Thus,

$$\int f\left(m_{t+1}\right) dF\left(m_{t+1}|m_t', \sigma^2\right) \geq \int f\left(m_{t+1}\right) dF\left(m_{t+1}|m_t, \sigma^2\right) \tag{75}$$

for any f in the class of increasing functions. Since W_{t+1} is a member of this class, we conclude that (74) is increasing in m_t. Hence W_t is increasing in m_t for all $t \leq T$.

The boundedness assumption on q_τ guarantees that m_τ is bounded. Thus W_t is bounded. Boundedness and $\beta \in (0,1)$ are sufficient conditions for the optimal value function in the finite horizon problem to converge to the optimal value of the infinite horizon problem. ■

The following Lemma is useful for the proof of Lemma 4.

Lemma 7 *The joint value of employment, W_τ, is convex in m_τ for all $\tau > 0$.*

Proof. Consider once again the $T-$period model developed in the proof of Lemma 2. The value function W_T in the terminal period (71) is convex in m_T. The proof proceeds by induction. Take as the induction hypothesis that W_{t+1} is convex in m_{t+1}, $t < T$. Now consider W_t, given in (72). It is sufficient to establish that $E\left[W_{t+1}|m_t, s_t^2\right]$ is convex in m_t. From (9), m_{t+1} is linear in m_t. Since W_{t+1} is a convex increasing function of m_{t+1}, it follows that W_{t+1} is convex in m_t. It is useful to write $W_{t+1} = W_{t+1}(m_t)$ to illustrate this dependence. Convexity implies

$$\alpha W_{t+1}(m_t) + (1-\alpha) W_{t+1}(m_t') \geq W_{t+1}(\alpha m_t + (1-\alpha) m_t') \tag{76}$$

for all m_t, m_t' and $\alpha \in [0,1]$. Since the expectation operator preserves inequalities,

$$E\left[\alpha W_{t+1}(m_t) + (1-\alpha) W_{t+1}(m_t')\right] \geq E\left[W_{t+1}(\alpha m_t + (1-\alpha) m_t')\right] \tag{77}$$

and hence

$$\alpha E\left[W_{t+1}(m_t)\right] + (1-\alpha) E\left[W_{t+1}(m_t')\right] \geq E\left[W_{t+1}(\alpha m_t + (1-\alpha) m_t')\right] \tag{78}$$

for all m_t, m_t' and $\alpha \in [0,1]$. ■

Proof of Lemma 4. First note that $E\left[W_{\tau+1}|m_\tau, s_\tau^2\right] = \int W_{\tau+1} dF\left(m_{\tau+1}|m_\tau, s_\tau^2\right)$, where F is the normal distribution function given in (15) with mean m_τ and variance $v_{\tau+1}$ defined in (16). Differentiating $v_{\tau+1}$ with respect to s_τ^2 yields

$$\frac{\partial v_{\tau+1}}{\partial s_\tau^2} = s_\tau^2 \frac{s_\tau^2 + 2\sigma_e^2}{\left(s_\tau^2 + \sigma_e^2\right)} > 0. \tag{79}$$

Thus an increase in s_τ^2 constitutes a mean-preserving spread on $m_{\tau+1}$. Since $W_{\tau+1}$ is an increasing convex function of $m_{\tau+1}$, for any $\tilde{s}_\tau^2 > s_\tau^2$ we have

$$\begin{aligned} E\left[W_{\tau+1}|m_\tau, s_\tau^2\right] &= \int W_{\tau+1} dF\left(m_{\tau+1}|m_\tau, s_{\tau+1}^2\right) \\ &\leq \int W_{\tau+1} dF\left(m_{\tau+1}|m_\tau, \tilde{s}_{\tau+1}^2\right) \\ &= E\left[W_{\tau+1}|m_\tau, \tilde{s}_\tau^2\right]. \end{aligned} \tag{80}$$

■

The following lemmata are useful for the proof of Proposition 6.

Lemma 8 $\frac{\partial U}{\partial b} = 0, \frac{\partial U}{\partial a} \in \left(0, \frac{1}{1-\beta}\right)$

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

$$\begin{aligned} U &= h + \beta\pi \int_{\underline{b}}^{\bar{b}} J_0 dF_b^* + \beta(1-\pi) U \\ &= \frac{h + \beta\pi \int_{\underline{b}}^{\bar{b}} J_0 dF_b^*}{1 - \beta(1-\pi)} \end{aligned} \tag{81}$$

where $J_0 = E\left[\max\left\{J_1, U\right\}|0, \sigma_c^2 + \sigma_z^2\right]$ and F_b^* is defined Appendix B. The statement $\frac{\partial U}{\partial b} = 0$ is obvious, since U doesn't depend on b.

From (81) we have

$$\frac{\partial U}{\partial a} = \frac{\beta\pi}{1 - \beta\left(1 - \pi\right)} \int_{\underline{b}}^{\bar{b}} \frac{\partial J_0}{\partial a} dF_b^*. \tag{82}$$

We can rewrite J_0 as

$$J_0 = U + \int_{\bar{m}_1}^{\infty} \left(J_1 - U\right) dF_1 \tag{83}$$

where F_τ is shorthand for $F\left(m_\tau | m_{\tau-1}, s_{\tau-1}^2\right)$. Differentiating and applying Leibniz's Rule,

$$\begin{aligned} \frac{\partial J_0}{\partial a} &= \frac{\partial U}{\partial a} - \frac{\partial \bar{m}_1}{\partial a}\left(\bar{J}_1 - U\right) f\left(\bar{m}_1 | 0, \sigma_c^2 + \sigma_z^2\right) + \int_{\bar{m}_1}^{\infty} \frac{\partial\left(J_1 - U\right)}{\partial a} dF_1 \\ &= \frac{\partial U}{\partial a} + \int_{\bar{m}_1}^{\infty} \frac{\partial\left(J_1 - U\right)}{\partial a} dF_1 \end{aligned} \tag{84}$$

where $\bar{J}_\tau$ 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 (21) using Leibniz's Rule,

$$\frac{\partial\left(J_\tau - U\right)}{\partial a} = \delta - \delta\left(1 - \beta\right)\frac{\partial U}{\partial a} + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \frac{\partial\left(J_{\tau+1} - U\right)}{\partial a} dF_{\tau+1} \tag{85}$$

for all $\tau > 0$. Repeated substitution of (85) into (84) gives the forward recursion

$$\frac{\partial J_0}{\partial a} = \frac{\partial U}{\partial a}\left(1 - \delta\left(1 - \beta\right) Z\right) + \delta Z \tag{86}$$

where

$$Z = \sum_{\tau=1}^{\infty} \beta^{\tau-1} \int_{\bar{m}_1}^{\infty} \cdots \int_{\bar{m}_\tau}^{\infty} dF_\tau \cdots dF_1 \in \left(0, \frac{1}{1-\beta}\right) \tag{87}$$

Thus

$$\frac{\partial U}{\partial a} = \frac{\beta\pi}{1 - \beta\left(1 - \pi\right)}\left[\frac{\partial U}{\partial a}\left(1 - \delta\left(1 - \beta\right)\int_{\underline{b}}^{\bar{b}} Z dF_b^*\right) + \delta \int_{\underline{b}}^{\bar{b}} Z dF_b^*\right]. \tag{88}$$

Some omitted algebra reveals

$$\frac{\partial U}{\partial a} = \frac{1}{1-\beta}\left[\frac{\delta\beta\pi \int_{\underline{b}}^{\bar{b}} Z dF_b^*}{1 + \delta\beta\pi \int_{\underline{b}}^{\bar{b}} Z dF_b^*}\right] \in \left(0, \frac{1}{1-\beta}\right) \tag{89}$$

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

Lemma 9 *The joint value of continuing the employment relationship, $J_\tau + \Pi_\tau$, is strictly increasing in a and b.*

Proof. Return to the finite horizon model used in the proof of Lemma 2. In the terminal period,

$$J_T + \Pi_T = \mu + a + b + m_T \tag{90}$$

which is strictly increasing in a and b. Take as the induction hypothesis that $J_{t+1} + \Pi_{t+1}$ is strictly increasing in a and b, $t < T$. Then

$$J_t + \Pi_t = \mu + a + b + m_t + \beta E\left[\max\left\{J_{t+1} + \Pi_{t+1}, U\right\}|m_t, s_t^2\right]. \tag{91}$$

By Lemma 8, U is independent of b and increasing in a. Combining this and the induction hypothesis, the last term in (91) is increasing in a and b. Since the current period return $\mu + a + b + m_t$ is strictly increasing in a and b, so is $J_t + \Pi_t$. ■

Proof of Proposition 6. Rewrite the threshold value of beliefs as

$$\bar{m}_\tau = (1-\beta)\,U - \mu - a - b + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \left(U - J_{\tau+1} - \Pi_{\tau+1}\right) d\bar{F}_{\tau+1} \tag{92}$$

where $\bar{F}_{\tau+1} = F\left(m_{\tau+1}|\bar{m}_\tau, s_\tau^2\right)$. Let $x \in \{a, b\}$. Differentiating (92) using Leibniz's Rule,

$$\begin{aligned}
\frac{\partial \bar{m}_\tau}{\partial x} &= (1-\beta)\frac{\partial U}{\partial x} - 1 - \beta \frac{\partial \bar{m}_{\tau+1}}{\partial x}\left(U - \bar{J}_{\tau+1} - \bar{\Pi}_{\tau+1}\right) f\left(\bar{m}_{\tau+1}|\bar{m}_\tau, s_\tau^2\right) \\
&\quad + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \frac{\partial\left(U - J_{\tau+1} - \Pi_{\tau+1}\right)}{\partial x} d\bar{F}_{\tau+1} \\
&\quad + \beta \frac{\partial \bar{m}_\tau}{\partial x} \int_{\bar{m}_{\tau+1}}^{\infty} \left(U - J_{\tau+1} - \Pi_{\tau+1}\right)\left(\frac{m_{\tau+1} - \bar{m}_\tau}{s_\tau^2}\right) d\bar{F}_{\tau+1} \\
&= \frac{(1-\beta)\frac{\partial U}{\partial x} - 1 + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \frac{\partial\left(U - J_{\tau+1} - \Pi_{\tau+1}\right)}{\partial x} d\bar{F}_{\tau+1}}{1 + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \left(U - J_{\tau+1} - \Pi_{\tau+1}\right)\left(\frac{m_{\tau+1} - \bar{m}_\tau}{s_\tau^2}\right) d\bar{F}_{\tau+1}}
\end{aligned} \tag{93}$$

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 8,

$$\frac{\partial \bar{m}_\tau}{\partial b} = \frac{-1 - \beta \int_{\bar{m}_{\tau+1}}^{\infty} \frac{\partial\left(J_{\tau+1} + \Pi_{\tau+1}\right)}{\partial b} d\bar{F}_{\tau+1}}{1 + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \left(U - J_{\tau+1} - \Pi_{\tau+1}\right)\left(\frac{m_{\tau+1} - \bar{m}_\tau}{s_\tau^2}\right) d\bar{F}_{\tau+1}}. \tag{94}$$

Since $\frac{\partial(J_{\tau+1} + \Pi_{\tau+1})}{\partial b} > 0$ by Lemma 9, 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 $m_{\tau+1} \geq \bar{m}_\tau$ for $m_{\tau+1} \geq \bar{m}_{\tau+1}$ (with equality only when $m_{\tau+1} = \bar{m}_{\tau+1}$). Thus $\frac{\partial \bar{m}_\tau}{\partial b} < 0$.

Substituting $x = a$ into (93) gives

$$\frac{\partial \bar{m}_\tau}{\partial a} = \frac{(1-\beta)\frac{\partial U}{\partial a} - 1 + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \frac{\partial\left(U - J_{\tau+1} - \Pi_{\tau+1}\right)}{\partial a} d\bar{F}_{\tau+1}}{1 + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \left(U - J_{\tau+1} - \Pi_{\tau+1}\right)\left(\frac{m_{\tau+1} - \bar{m}_\tau}{s_\tau^2}\right) d\bar{F}_{\tau+1}}. \tag{95}$$

As in (94), the denominator is positive. To sign the numerator note that for all $s \geq 1$,

$$\begin{aligned} U - J_{\tau+s} - \Pi_{\tau+s} &= (1-\beta) U - \mu - a - b - m_{\tau+s} \\ &\quad + \beta \int_{\bar{m}_{\tau+s+1}}^{\infty} \left(U - J_{\tau+s+1} - \Pi_{\tau+s+1} \right) dF_{\tau+s+1} \end{aligned} \tag{96}$$

and differentiating gives the recursion

$$\frac{\partial \left(U - J_{\tau+s} - \Pi_{\tau+s} \right)}{\partial a} = (1-\beta) \frac{\partial U}{\partial a} - 1 + \beta \int_{\bar{m}_{\tau+s+1}}^{\infty} \frac{\partial \left(U - J_{\tau+s+1} - \Pi_{\tau+s+1} \right)}{\partial a} dF_{\tau+s+1}. \tag{97}$$

Repeated substitution of (97) into the numerator of (95) gives

$$\frac{\partial \bar{m}_{\tau}}{\partial a} = \frac{\left[(1-\beta) \frac{\partial U}{\partial a} - 1 \right] \bar{Z}_{\tau}}{1 + \beta \int_{\bar{m}_{\tau+1}}^{\infty} \left(U - J_{\tau+1} - \Pi_{\tau+1} \right) \left(\frac{m_{\tau+1} - \bar{m}_{\tau}}{s_{\tau}^2} \right) d\bar{F}_{\tau+1}} \tag{98}$$

where

$$\bar{Z}_{\tau} = 1 + \sum_{s=1}^{\infty} \beta^s \int_{\bar{m}_{\tau+1}}^{\infty} \int_{\bar{m}_{\tau+2}}^{\infty} \cdots \int_{\bar{m}_{\tau+s}}^{\infty} dF_{\tau+s} \cdots dF_{\tau+2} d\bar{F}_{\tau+1} > 0 \tag{99}$$

and we conclude that the numerator of (98) is negative, since $\frac{\partial U}{\partial a} < \frac{1}{1-\beta}$ by Lemma 8. ■

B Appendix: The Steady State

B.1 Flow-Balance Equations

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

$$\begin{aligned} & (1-u)\, l(a,b,k,\tau) \Pr(m_\tau < \bar{m}_\tau) \\ = \quad & (1-u)\, l(a,b,k,\tau)\, \Phi\left(\frac{\bar{m}_\tau}{V_\tau^{1/2}}\right) \end{aligned} \tag{100}$$

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

$$(1-u)\sum_{\tau=1}^{\infty} l(a,b,k,\tau)\, \Phi\left(\frac{\bar{m}_\tau}{V_\tau^{1/2}}\right) \tag{101}$$

and the aggregate flow into unemployment is

$$(1-u)\int_{\underline{a}}^{\bar{a}}\int_{\underline{b}}^{\bar{b}}\int_{0}^{\infty}\sum_{\tau=1}^{\infty} l(a,b,k,\tau)\, \Phi\left(\frac{\bar{m}_\tau}{V_\tau^{1/2}}\right) dkdbda. \tag{102}$$

Let f_a and $f_{b,k}$ denote the density functions associated with the distributions F_a and $F_{b,k}$ of worker and firm types . The flow of type a workers out of unemployment and into type (b,k) firms is

$$m(u,v)\, f_a(a)\, f_{b,k}(b,k). \tag{103}$$

Thus the aggregate flow out of unemployment is

$$m(u,v)\int_{\underline{a}}^{\bar{a}}\int_{\underline{b}}^{\bar{b}}\int_{0}^{\infty} f_a(a)\, f_{b,k}(b,k)\, dkdbda = m(u,v). \tag{104}$$

The steady state flow-balance condition is the equality of (101) and (103) for all worker types a and all firm types (b,k) . This implies the aggregate steady state flow-balance (102) $=$ (104) .

B.2 Unemployment and Vacancies

In the steady state, the probability λ that a randomly selected vacancy is filled is constant. Thus the equilibrium number of vacancies opened by each firm, i.e., the solution to (26), is also constant. Let $v^*_{b,k}$ denote the steady state number of vacancies opened by a type (b,k) firm, i.e.,

$$v^*_{b,k} = \arg\max_{v\in\mathbb{N}} \sum_{l=0}^{v} \binom{v}{l} \lambda^l (1-\lambda)^{v-l}\left[l\int_{\underline{a}}^{\bar{a}} \Pi_0 dF^*_a - \kappa(l)\right] - kv \tag{105}$$

when λ takes its steady state value and where F^*_a is the steady state distribution of unemployed worker types defined below. Thus the steady state number of vacancies opened by all type (b,k) firms is $\phi v^*_{b,k} f_{b,k}(b,k)$, and the steady state stock of vacancies in the economy is

$$v = \phi\int_{\underline{b}}^{\bar{b}}\int_{0}^{\infty} v^*_{b,k} f_{b,k}(b,k)\, dkdb. \tag{106}$$

The steady state level of unemployment is implicitly defined by the equality of (102) and (104) when v takes its steady state value.

Each open vacancy is associated with a firm type (b,k). Let $f^*_{b,k}(b,k)$ denote the steady state distribution of firm types among open vacancies. This is

$$f^*_{b,k}(b,k) = \phi \frac{v^*_{b,k}}{v} f_{b,k}(b,k). \tag{107}$$

Define the marginal distribution $f^*_b(b) = \int_0^\infty f^*_{b,k}(b,k)\,dk$ with associated 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:

$$l(a) = \int_{\underline{b}}^{\overline{b}} \int_0^\infty \sum_{\tau=1}^{\infty} l(a,b,k,\tau)\,dkdb. \tag{108}$$

Then the density function $f^*_a(a)$ associated with F^*_a is $f^*_a(a) = u^{-1}\left[f_a(a) - (1-u)\,l(a)\right]$.

B.3 Firm size

Let

$$l(b,k) = \int_{\underline{a}}^{\overline{a}} \sum_{\tau=1}^{\infty} l(a,b,k,\tau)\,da \tag{109}$$

be the density of employment at type (b,k) firms. Then the average size of type (b,k) firms is

$$\frac{(1-u)\,l(b,k)}{\phi f_{b,k}(b,k)}. \tag{110}$$

C Appendix: The Dense Sampling Algorithm

In labor market data, observations are connected by a sequence of workers and firms. Workers are connected to one another by a common employer. Firms are connected to one another by a common employee. Connectedness is crucial for identifying worker and firm effects in linear and mixed models of employment outcomes. The degree of connectedness depends both on the number of connected groups in the data and their size. See Searle (1987) for a statistical definition of connectedness. Abowd et al. (2002) discuss connectedness in the context of labor market data.

This section describes an algorithm for sampling highly connected work histories from longitudinal linked employer-employee data. The algorithm is designed to draw two disjoint samples of predictable size. The disjoint samples can be used for independent model estimation and validation. Of primary importance is the fact that the resultant samples have all the statistical properties of a simple random sample of workers employed at a point in time. Specifically, all such workers have an equal probability of being sampled. However, unlike a (naive) simple random sample of workers, the algorithm guarantees that all workers are connected to at least $n > 1$ others.

The basic idea is as follows. In a reference period, sample firms with probabilities that are proportional to employment. Next, sample workers within firms, with equal (firm-specific) probabilities. Roughly speaking, the probability of sampling a particular employee within a firm is inversely proportional to the firm's employment in the reference period. In a sample of dominant jobs,[45] the resulting probability of sampling any worker is a constant.

The samples' characteristics are determined by three parameters: $p \in [0, 1]$ determines a firm's probability of being sampled; $n \in \mathbb{N}$ determines the minimum level of connectedness and a worker's probability of being sampled within a firm; and $m \in [0, 1]$ determines an observation's probability of being sampled into disjoint sample $s \in \{1, 2\}$. Sample $s = 1$ is a $100mpn$ percent random sample of workers; sample $s = 2$ is a $100(1 - m)\,pn$ percent random sample of workers. When $m = \frac{1}{2}$, the disjoint samples are of equal size.

Let $\mathbf{S}$ denote the base sample from which the disjoint subsamples s will be drawn. Let t be the reference period, and let N_j denote firm j's employment in t. The algorithm relies on two assumptions:

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

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

Assumptions 1 and 2 are easily satisfied by imposing restrictions on the base sample $\mathbf{S}$. For example, restrict $\mathbf{S}$ to a sample of dominant jobs at firms with at least n employees in t.

Denote the probability that firm j is sampled into some s by $\pi(j)$, and let $\pi(j) = \min\{1, pN_j\}$. That is,

$$\pi(j) = \begin{cases} 1 & \text{if } N_j \geq p^{-1} \\ pN_j & \text{if } N_j < p^{-1} \end{cases}$$

so that a firm's probability of being sampled is proportional to employment in t.

[45] A worker's dominant job in a period is the employer at which he/she earned the most in that period. Each individual has only one dominant job in each period. Technically, the algorithm requires that each individual has only one employer per period. The dominant job restriction is a convenient way of guaranteeing this.

Sampling Rule 1 *If firm j is sampled in t, sample j into $s = 1$ with probability m; sample j into $s = 2$ with probability $(1 - m)$.*

Denote the probability that firm j is sampled into s by $\pi_s(j)$. Thus,

$$\begin{aligned}
\pi_1(j) &= \min\{m, mpN_j\} \\
&= \begin{cases} m & \text{if } N_j \geq p^{-1} \\ mpN_j & \text{if } N_j < p^{-1} \end{cases} \\
\pi_2(j) &= \min\{1 - m, (1 - m)pN_j\} \\
&= \begin{cases} 1 - m & \text{if } N_j \geq p^{-1} \\ (1 - m)pN_j & \text{if } N_j < p^{-1} \end{cases} .
\end{aligned}$$

Let $n_j = \max\{n, pnN_j\}$. That is,

$$n_j = \begin{cases} pnN_j & \text{if } N_j \geq p^{-1} \\ n & \text{if } N_j < p^{-1} \end{cases} . \tag{111}$$

Sampling Rule 2 *If firm j is sampled into s in t, draw a simple random sample of n_j workers employed at j in period t into s.*

Together, Rule 2 and the definition of n_j in (111) have several implications for the structure of the dense sample. First, it is clear that in each sample s, each worker is connected to at least n others: their fellow employees sampled from firm j in t. Consequently, increasing n increases the minimum degree of connectedness in each sample. Second, as shown in Figure 8 the firm-specific sampling rate n_j/N_j is nonincreasing in firm j's period t employment N_j.

Let $\pi_s(i|j)$ denote the probability that worker i is sampled into s, given that j has been sampled into s in t. This is

$$\begin{aligned}
\pi_s(i|j) &= n_j N_j^{-1} \\
&= \begin{cases} pn & \text{if } N_j \geq p^{-1} \\ nN_j^{-1} & \text{if } N_j < p^{-1} \end{cases} .
\end{aligned}$$

To determine an individual's unconditional probability of being sampled, we need to introduce some further notation. Let $\pi_s(j|i)$ denote the probability that firm j is sampled into s in t, given that employee i is sampled into s and $j = J(i,t)$. Assumption 1 implies $\pi_s(j|i) = 1$. Denote the unconditional probability that individual i is sampled into s by $\pi_s(i)$. By Bayes' rule,

$$\pi_s(i) = \frac{\pi_s(i|j)\,\pi_s(j)}{\pi_s(j|i)} = \pi_s(i|j)\,\pi_s(j)$$

so that

$$\begin{aligned}
\pi_1(i) &= \begin{cases} mpn & \text{if } N_j \geq p^{-1} \\ mpN_j n N_j^{-1} & \text{if } N_j < p^{-1} \end{cases} \\
&= mpn \qquad (112) \\
\pi_2(i) &= \begin{cases} (1 - m)pn & \text{if } N_j \geq p^{-1} \\ (1 - m)pN_j n N_j^{-1} & \text{if } N_j < p^{-1} \end{cases} . \\
&= (1 - m)pn \qquad (113)
\end{aligned}$$

A final Sampling Rule completely specifies the subsamples:

Sampling Rule 3 *If i is sampled into s, sample i's complete work history into s.*

Equations (112) and (113) demonstrate that both subsamples s have the properties of a simple random sample of individuals employed in t. That is, all individuals in $\mathbf{S}$ that are employed in t have equal probability of being sampled into each s. Furthermore, Assumption 1 and Rules 1 and 2 guarantee that the samples are disjoint: $\pi_s\left(i|i \text{ is sampled into } s'\right) = 0$ for $s, s' = 1, 2$ and $s \neq s'$. Finally, Rule 3 states that the subsamples consist of the complete work histories (in $\mathbf{S}$) of individuals sampled according to Rules 1 and 2.

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, March). 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, 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.

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.

Blume, L. E. and D. Easley (1998). Rational expectations and rational learning. In B. Allen and J. S. Jordan (Eds.), *Organizations with Incomplete Information: Essays in Economic Analysis: A Tribute to Roy Radner*, Chapter 3, pp. 61–109. Cambridge: Cambridge University Press.

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

Burdett, K. and M. G. Coles (1999). Long-term partnership formation: Marriage and employment. *The Economic Journal 109*, F307–F334.

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

Chamberlain, G. (1984). *Panel Data*, Chapter 22, pp. 1248–1318. New York: Elsevier.

Dostie, B. (2002). Controlling for demand side factors and job matching: Maximum likelihood estimates of the returns to seniority using matched employer-employee data. Mimeo.

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

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

Gautier, P. A. (2000). Unemployment and search externalities in a model with heterogeneous jobs and workers. Mimeo.

Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2002). Comparative advantage, learning, and sectoral wage discrimination. NBER Working Paper 8889.

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.

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.

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*. 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. (2002). Skill and luck in the theory of turnover. Mimeo.

Nagypal, E. (2000). Learning-by-doing versus learning about match quality: Can we tell them apart? Mimeo.

Nerlove, M. (1971). A note on error components models. *Econometrica 39*, 383–396.

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 (1998). A multivariate technique for multiply-imputing missing values using a sequence of regression models. Survey Research Center, University of Michigan.

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

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*. John Wiley and Sons.

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

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. Mimeo.

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.

Stinson, M. (2002). Estimating the relationship between employer-provided health insurance, worker mobility, and wages. Mimeo.

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

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

TABLE 1
PROPERTIES OF CONNECTED GROUPS OF WORKERS AND FIRMS

	Full Analysis Sample[a]	Dense Sample 1[b]	Dense Sample 2[b]	Simple Random Sample[c]
Number of Groups	84,708	1,140	1,081	9,457
Number of Workers	9,271,766	49,425	48,003	49,200
Number of Firms	573,237	27,421	27,555	40,064
Number of Worker-Firm Matches	15,305,508	92,539	90,500	93,182
Number of Matches in Smallest Group	5	5	5	1
Proportion of Matches in:				
Largest Group	99.06	67.25	68.82	59.37
Second Largest Group	0.0006	24.70	22.68	20.30
Third Largest Group	0.0003	0.04	0.04	0.06
Groups containing 5 or more matches	100	100	100	84.44
Groups containing only 1 match	0	0	0	5.50

[a] Results combined across three completed data implicates.

[b] One percent dense random samples of workers employed in 1997, drawn according to the dense sampling algorithm in Appendix C. Results are combined across three completed data implicates.

[c] One percent simple random sample of workers employed in 1997. Results are from one completed data implicate.

TABLE 2
SUMMARY STATISTICS
Estimates Combined Across 3 Completed Data Implicates

	Full Analysis Sample		Dense Sample 1		Dense Sample 2		Simple Random Sample	
Variable	Mean[a]	Std. Dev[b]	Mean[a]	Std. Dev[b]	Mean[a]	Std. Dev[b]	Mean[a]	Std. Dev[b]
Demographic Characteristics								
Male (Proportion)	0.560	0.496	0.564	0.496	0.584	0.493	0.569	0.495
Age (Years)	40.6	10.2	40.4	9.5	40.3	9.6	40.3	9.5
Men								
Nonwhite (Proportion)	0.209	0.574	0.210	0.573	0.203	0.553	0.210	0.570
Race Missing (Proportion)	0.036	0.250	0.034	0.243	0.036	0.244	0.035	0.245
Less than high school (Proportion)	0.119	0.445	0.110	0.428	0.115	0.429	0.109	0.424
High school (Proportion)	0.299	0.666	0.291	0.657	0.299	0.650	0.297	0.659
Some college (Proportion)	0.232	0.600	0.233	0.599	0.234	0.588	0.231	0.594
Associate or Bachelor Degree (Proportion)	0.247	0.617	0.256	0.623	0.247	0.601	0.258	0.622
Graduate or Professional Degree (Proportion)	0.103	0.416	0.110	0.428	0.105	0.411	0.105	0.417
Women								
Nonwhite (Proportion)	0.237	0.694	0.240	0.702	0.240	0.721	0.236	0.599
Race Missing (Proportion)	0.022	0.225	0.021	0.220	0.019	0.211	0.021	-0.011
Less than high school (Proportion)	0.094	0.453	0.085	0.434	0.090	0.456	0.088	0.283
High school (Proportion)	0.314	0.784	0.299	0.772	0.308	0.804	0.301	0.390
Some college (Proportion)	0.253	0.715	0.250	0.714	0.251	0.736	0.251	0.311
Associate or Bachelor Degree (Proportion)	0.259	0.723	0.278	0.748	0.268	0.757	0.270	0.423
Graduate or Professional Degree (Proportion)	0.080	0.418	0.088	0.440	0.083	0.438	0.090	0.308
Work History Characteristics								
Tenure (Years)	4.48	3.48	4.90	3.59	4.78	3.51	4.85	3.56
Job is Left Censored (Proportion)	0.331	0.470	0.358	0.479	0.342	0.474	0.347	0.476
Real Annualized Earnings (1990 Dollars)	53755	50804	57209	51196	56571	51980	56483	50074
Men								
Labor Market Experience (Years)	11.8	13.1	11.7	12.6	12.1	12.7	11.7	12.6
Initial Experience <0 (Proportion)	0.023	0.201	0.022	0.197	0.021	0.190	0.023	0.200
Worked 0 Full Quarters in Calendar Year (Proportion)	0.077	0.363	0.059	0.318	0.060	0.316	0.060	0.320
Worker 1 Full Quarter in Calendar Year (Proportion)	0.146	0.490	0.114	0.435	0.122	0.441	0.120	0.443
Worker 2 Full Quarters in Calendar Year (Proportion)	0.134	0.470	0.123	0.450	0.120	0.436	0.122	0.447
Worker 3 Full Quarters in Calendar Year (Proportion)	0.143	0.484	0.136	0.472	0.134	0.460	0.134	0.466
Worker 4 Full Quarters in Calendar Year (Proportion)	0.500	0.914	0.568	0.963	0.563	0.992	0.564	0.968
Women								
Labor Market Experience (Years)	9.5	13.0	9.3	12.6	8.8	12.4	9.2	12.5
Initial Experience <0 (Proportion)	0.023	0.227	0.022	0.226	0.021	0.221	0.022	0.223
Worked 0 Full Quarters in Calendar Year (Proportion)	0.070	0.393	0.053	0.346	0.055	0.359	0.056	0.357
Worker 1 Full Quarter in Calendar Year (Proportion)	0.136	0.538	0.108	0.486	0.113	0.509	0.111	0.496
Worker 2 Full Quarters in Calendar Year (Proportion)	0.129	0.526	0.117	0.505	0.114	0.510	0.117	0.507
Worker 3 Full Quarters in Calendar Year (Proportion)	0.141	0.548	0.129	0.529	0.132	0.548	0.128	0.529
Worker 4 Full Quarters in Calendar Year (Proportion)	0.524	0.958	0.592	1.003	0.586	1.033	0.588	1.009
Miscellany								
Year (Proportions)								
1991	0.094	0.293	0.080	0.271	0.079	0.270	0.079	0.270
1992	0.093	0.291	0.083	0.277	0.083	0.276	0.083	0.275
1993	0.096	0.294	0.089	0.285	0.089	0.285	0.089	0.284
1994	0.099	0.298	0.097	0.295	0.097	0.295	0.096	0.294
1995	0.101	0.302	0.105	0.306	0.105	0.307	0.104	0.305
1996	0.104	0.305	0.115	0.319	0.115	0.319	0.114	0.318
1997	0.106	0.308	0.131	0.337	0.131	0.338	0.138	0.345
1998	0.108	0.311	0.118	0.322	0.118	0.322	0.117	0.322
1999	0.107	0.309	0.108	0.310	0.108	0.311	0.107	0.309
Number of Observations	49,281,533		357,725		345,954		357,009	
Number of Workers	9,271,766		49,425		48,003		49,200	
Number of Firms	573,237		27,421		27,555		40,064	
Number of Worker-Firm Matches	15,305,508		92,539		90,500		93,182	

[a] Means are computed on each completed data implicate for each sample. The reported mean is the simple average 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 standard deviation is the square root of the simple average of the variances computed on each implicate.

TABLE 3
ESTIMATED VARIANCE COMPONENTS AND SUMMARY OF MODEL FIT STATISTICS
Dependent Variable: Log(Real Annualized Earnings), Combined Results From 3 Completed Data Implicates

	Fixed Person and Firm Effects		Random Person and Firm Effects		Random Person, Firm, and Match Effects		Random Person and Firm Effects, Unstructured Residual Covariance	
	Parameter Estimate[a,e]	Standard Error[b]	Parameter Estimate[a]	Standard Error[b]	Parameter Estimate[a]	Standard Error[b]	Parameter Estimate[a]	Standard Error[b]
No Correction for Truncation								
Variance of person effect (σ^2_α)	0.290	(0.002)	0.230	(0.005)	0.171	(0.003)	0.177	(0.002)
Variance of firm effect (σ^2_ψ)	0.077	(0.000)	0.153	(0.002)	0.074	(0.002)	0.076	(0.007)
Variance of match effect (σ^2_γ)					0.070	(0.002)		
Residual variance (σ^2_ε)	0.061	(0.000)	0.044	(0.001)	0.036	(0.000)	n/a	n/a
Corrected for Truncation								
Variance of person effect (σ^2_α)			0.229	(0.005)	0.171	(0.003)	0.176	(0.002)
Variance of firm effect (σ^2_ψ)			0.153	(0.002)	0.075	(0.002)	0.077	(0.006)
Variance of match effect (σ^2_γ)					0.069	(0.002)		
Residual variance (σ^2_ε)			0.044	(0.001)	0.036	(0.000)	n/a	n/a
Summary of Model Fit Statistics								
	Value[a]	Between-Implicate Std. Dev.[c]	Value[a]	Between-Implicate Std. Dev.[c]	Value[a]	Between-Implicate Std. Dev.[c]	Value[a]	Between-Implicate Std. Dev.[c]
No Correction for Truncation								
Log Likelihood			263423	(1051.0)	280342	(928.2)	313288	(1292.8)
AIC			-526760	(2102.0)	-560595	(1856.4)	-626163	(2585.5)
BIC			-526296	(2102.1)	-560121	(1856.4)	-623941	(2585.1)
Var(out-of-sample prediction error)[d]			0.090	(0.0012)	0.069	(0.0007)	0.096	(0.0004)
Corrected for Truncation								
Log Likelihood			263807	(1063.3)	280509	(938.4)	313395	(1298.8)
AIC			-527526	(2126.7)	-560927	(1876.7)	-626376	(2597.5)
BIC			-527051	(2126.8)	-560442	(1876.7)	-624143	(2597.1)
Var(out-of-sample prediction error)[d]			0.090	(0.0012)	0.069	(0.0007)	0.096	(0.0004)
Number of Observations	49,281,533	(9103.2)	357,725	(2363.5)	357,725	(2363.5)	357,725	(2363.5)
Number of Workers	9,271,766	(710.4)	49,425	(150.4)	49,425	(150.4)	49,425	(150.4)
Number of Firms	573,237	(118.1)	27,421	(12.6)	27,421	(12.6)	27,421	(12.6)
Number of Worker-Firm Matches	15,305,508	(3195.5)	92,539	(470.3)	92,539	(470.3)	92,539	(470.3)

[a] Simple average of parameter estimate across three completed data implicates.
[b] Square root of the total variance of parameter estimate across three completed data implicates, as defined in Rubin (1987).
[c] Square root of between-implicate variance.
[d] Computed on Dense Sample 2. Dependent variable is scaled to have unit variance.
[e] Simple variance of estimated person and firm effects, averaged across three completed data implicates.

TABLE 4
ESTIMATED VARIANCE COMPONENTS AND SUMMARY OF MODEL FIT STATISTICS
Dependent Variable: Real Annualized Earnings (Unit Variance Scale), Combined Results From 3 Completed Data Implicates

	Fixed Person and Firm Effects		Random Person and Firm Effects		Random Person, Firm, and Match Effects		Random Person and Firm Effects, Unstructured Residual Covariance	
	Parameter Estimate[a,e]	Standard Error[b]	Parameter Estimate[a]	Standard Error[b]	Parameter Estimate[a]	Standard Error[b]	Parameter Estimate[a]	Standard Error[b]
No Correction for Truncation								
Variance of person effect (σ^2_α)	0.720	(0.003)	0.551	(0.004)	0.496	(0.005)	0.402	(0.004)
Variance of firm effect (σ^2_ψ)	0.062	(0.000)	0.100	(0.003)	0.045	(0.002)	0.043	(0.002)
Variance of match effect (σ^2_γ)					0.080	(0.002)		
Residual variance (σ^2_ε)	0.160	(0.000)	0.151	(0.001)	0.137	(0.001)	n/a	n/a
Corrected for Truncation								
Variance of person effect (σ^2_α)			0.551	(0.036)	0.495	(0.005)	0.401	(0.003)
Variance of firm effect (σ^2_ψ)			0.099	(0.003)	0.045	(0.002)	0.043	(0.002)
Variance of match effect (σ^2_γ)					0.079	(0.002)		
Residual variance (σ^2_ε)			0.151	(0.001)	0.137	(0.001)	n/a	n/a
Summary of Model Fit Statistics								
	Value[a]	Between-Implicate Std. Dev.[c]	Value[a]	Between-Implicate Std. Dev.[c]	Value[a]	Between-Implicate Std. Dev.[c]	Value[a]	Between-Implicate Std. Dev.[c]
No Correction for Truncation								
Log Likelihood			66937	(874.9)	72661	(723.6)	106727	(1719.8)
AIC			-133787	(1749.7)	-145234	(1447.1)	-213043	(3439.7)
BIC			-133323	(1749.6)	-144759	(1447.0)	-210820	(3440.2)
Var(out-of-sample prediction error)[d]			0.131	(0.0006)	0.114	(0.0009)	0.131	(0.0012)
Corrected for Truncation								
Log Likelihood			67061	(869.4)	72736	(720.5)	106785	(1703.4)
AIC			-134034	(1738.8)	-145382	(1441.1)	-213156	(3406.9)
BIC			-133560	(1738.7)	-144896	(1440.9)	-210923	(3407.3)
Var(out-of-sample prediction error)[d]			0.131	(0.0006)	0.114	(0.0009)	0.131	(0.0009)
Number of Observations	49,281,533	(9103.2)	357725	(2363.5)	357725	(2363)	357725	(2363)
Number of Workers	9,271,766	(710.4)	49425	(150.4)	49425	(150)	49425	(150)
Number of Firms	573,237	(118.1)	27421	(12.6)	27421	(13)	27421	(13)
Number of Worker-Firm Matches	15,305,508	(3195.5)	92539	(470.3)	92539	(470)	92539	(470)

[1] The standard deviation of real annualized earnings in Dense Sample 1 is $51,195 (1990 Dollars).
[a] Simple average of parameter estimate across three completed data implicates.
[b] Square root of the total variance of parameter estimate across three completed data implicates, as defined in Rubin (1987).
[c] Square root of between-implicate variance.
[d] Computed on Dense Sample 2. Dependent variable is scaled to have unit variance.
[e] Simple variance of estimated person and firm effects, averaged across three completed data implicates.

TABLE 5
CORRELATIONS AMONG ESTIMATED EFFECTS
Dependent Variable: Log(Real Annualized Earnings), No Correction for Truncation
Results Combined From 3 Completed Data Implicates

Fixed Person and Firm Effects

	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1.000	0.738	0.659	0.335	0.448	0.176
Pure Person Effect (θ)	0.738	1.000	0.914	0.406	0.034	-0.297
Unobserved Component (α)	0.659	0.914	1.000	0.000	-0.005	-0.271
Observed Component (Uη)	0.335	0.406	0.000	1.000	0.094	-0.121
Pure Firm Effect (ψ)	0.448	0.034	-0.005	0.094	1.000	0.048
Time-Varying Covariates (Xβ)	0.176	-0.297	-0.271	-0.121	0.048	1.000

Random Person and Firm Effects

	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1.000	0.801	0.705	0.376	0.472	0.289
Pure Person Effect (θ)	0.801	1.000	0.908	0.408	0.027	0.020
Unobserved Component (α)	0.705	0.908	1.000	-0.013	-0.008	-0.029
Observed Component (Uη)	0.376	0.408	-0.013	1.000	0.083	0.110
Pure Firm Effect (ψ)	0.472	0.027	-0.008	0.083	1.000	0.042
Time-Varying Covariates (Xβ)	0.289	0.020	-0.029	0.110	0.042	1.000

Random Person, Firm, and Match Effects

	y	θ	α	Uη	ψ	γ	Xβ
Log Earnings (y)	1.000	0.829	0.735	0.378	0.536	0.615	0.290
Pure Person Effect (θ)	0.829	1.000	0.853	0.511	0.226	0.475	0.039
Unobserved Component (α)	0.735	0.853	1.000	-0.013	0.205	0.560	-0.032
Observed Component (Uη)	0.378	0.511	-0.013	1.000	0.095	-0.011	0.127
Pure Firm Effect (ψ)	0.536	0.226	0.205	0.095	1.000	0.166	0.039
Pure Match Effect (γ)	0.615	0.475	0.560	-0.011	0.166	1.000	-0.016
Time-Varying Covariates (Xβ)	0.290	0.039	-0.032	0.127	0.039	-0.016	1.000

Random Person and Firm Effects, Unstructured Residual Covariance

	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1.000	0.820	0.733	0.361	0.536	0.300
Pure Person Effect (θ)	0.820	1.000	0.867	0.489	0.222	0.018
Unobserved Component (α)	0.733	0.867	1.000	-0.011	0.201	-0.029
Observed Component (Uη)	0.361	0.489	-0.011	1.000	0.093	0.087
Pure Firm Effect (ψ)	0.536	0.222	0.201	0.093	1.000	0.041
Time-Varying Covariates (Xβ)	0.300	0.018	-0.029	0.087	0.041	1.000

TABLE 6
CORRELATIONS AMONG ESTIMATED EFFECTS
Dependent Variable: Log(Real Annualized Earnings), Corrected for Truncation
Results Combined From 3 Completed Data Implicates

Random Person and Firm Effects

	y	θ	α	$U\eta$	ψ	$X\beta$	$\beta_\lambda\lambda$
Log Earnings (y)	1	0.802	0.706	0.376	0.477	0.291	0.056
Pure Person Effect (θ)	0.802	1	0.908	0.407	0.032	0.024	0.105
Unobserved Component (α)	0.706	0.908	1	-0.012	-0.004	-0.028	0.111
Observed Component ($U\eta$)	0.376	0.407	-0.012	1	0.084	0.118	0.103
Pure Firm Effect (ψ)	0.477	0.032	-0.004	0.084	1	0.050	-0.064
Time-Varying Covariates ($X\beta$)	0.291	0.024	-0.028	0.118	0.050	1	0.131
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.056	0.105	0.111	0.103	-0.064	0.131	1

Random Person, Firm, and Match Effects

	y	θ	α	$U\eta$	ψ	γ	$X\beta$	$\beta_\lambda\lambda$
Log Earnings (y)	1	0.829	0.735	0.378	0.536	0.616	0.291	0.056
Pure Person Effect (θ)	0.829	1	0.854	0.510	0.225	0.476	0.042	0.113
Unobserved Component (α)	0.735	0.854	1	-0.013	0.204	0.561	-0.031	0.122
Observed Component ($U\eta$)	0.378	0.510	-0.013	1	0.096	-0.011	0.132	0.103
Pure Firm Effect (ψ)	0.536	0.225	0.204	0.096	1	0.165	0.043	-0.133
Pure Match Effect (γ)	0.616	0.476	0.561	-0.011	0.165	1	-0.013	0.160
Time-Varying Covariates ($X\beta$)	0.291	0.042	-0.031	0.132	0.043	-0.013	1	0.130
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.056	0.113	0.122	0.103	-0.133	0.160	0.130	1

Random Person and Firm Effects, Unstructured Residual Covariance

	y	θ	α	$U\eta$	ψ	$X\beta$	$\beta_\lambda\lambda$
Log Earnings (y)	1	0.820	0.733	0.362	0.537	0.302	0.056
Pure Person Effect (θ)	0.820	1	0.867	0.489	0.222	0.021	0.116
Unobserved Component (α)	0.733	0.867	1	-0.011	0.200	-0.028	0.123
Observed Component ($U\eta$)	0.362	0.489	-0.011	1	0.094	0.092	0.106
Pure Firm Effect (ψ)	0.537	0.222	0.200	0.094	1	0.045	-0.126
Time-Varying Covariates ($X\beta$)	0.302	0.021	-0.028	0.092	0.045	1	0.124
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.056	0.116	0.123	0.106	-0.126	0.124	1

TABLE 7
CORRELATIONS AMONG ESTIMATED EFFECTS
Dependent Variable: Real Annualized Earnings, No Correction for Truncation
Results Combined From 3 Completed Data Implicates

Fixed Person and Firm Effects

	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1.000	0.761	0.729	0.224	0.297	0.205
Pure Person Effect (θ)	0.761	1.000	0.937	0.350	0.023	-0.292
Unobserved Component (α)	0.729	0.937	1.000	0.000	0.011	-0.230
Observed Component (Uη)	0.224	0.350	0.000	1.000	0.036	-0.220
Pure Firm Effect (ψ)	0.297	0.023	0.011	0.036	1.000	0.063
Time-Varying Covariates (Xβ)	0.205	-0.292	-0.230	-0.220	0.063	1.000

Random Person and Firm Effects

	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1.000	0.846	0.812	0.248	0.400	0.262
Pure Person Effect (θ)	0.846	1.000	0.941	0.344	0.173	-0.058
Unobserved Component (α)	0.812	0.941	1.000	0.007	0.160	-0.012
Observed Component (Uη)	0.248	0.344	0.007	1.000	0.066	-0.139
Pure Firm Effect (ψ)	0.400	0.173	0.160	0.066	1.000	0.020
Time-Varying Covariates (Xβ)	0.262	-0.058	-0.012	-0.139	0.020	1.000

Random Person, Firm, and Match Effects

	y	θ	α	Uη	ψ	γ	Xβ
Log Earnings (y)	1.000	0.853	0.822	0.247	0.424	0.624	0.262
Pure Person Effect (θ)	0.853	1.000	0.928	0.378	0.274	0.507	-0.063
Unobserved Component (α)	0.822	0.928	1.000	0.007	0.266	0.544	-0.012
Observed Component (Uη)	0.247	0.378	0.007	1.000	0.075	0.009	-0.141
Pure Firm Effect (ψ)	0.424	0.274	0.266	0.075	1.000	0.266	0.026
Pure Match Effect (γ)	0.624	0.507	0.544	0.009	0.266	1.000	0.004
Time-Varying Covariates (Xβ)	0.262	-0.063	-0.012	-0.141	0.026	0.004	1.000

Random Person and Firm Effects, Unstructured Residual Covariance

	y	θ	α	Uη	ψ	Xβ
Log Earnings (y)	1.000	0.809	0.781	0.242	0.412	0.262
Pure Person Effect (θ)	0.809	1.000	0.910	0.420	0.257	-0.075
Unobserved Component (α)	0.781	0.910	1.000	0.006	0.250	-0.012
Observed Component (Uη)	0.242	0.420	0.006	1.000	0.073	-0.155
Pure Firm Effect (ψ)	0.412	0.257	0.250	0.073	1.000	0.038
Time-Varying Covariates (Xβ)	0.262	-0.075	-0.012	-0.155	0.038	1.000

TABLE 8
CORRELATIONS AMONG ESTIMATED EFFECTS
Dependent Variable: Real Annualized Earnings, Corrected For Truncation
Results Combined From 3 Completed Data Implicates

Random Person and Firm Effects

	y	θ	α	$U\eta$	ψ	$X\beta$	$\beta_\lambda\lambda$
Log Earnings (y)	1	0.846	0.812	0.248	0.402	0.263	0.015
Pure Person Effect (θ)	0.846	1	0.942	0.344	0.174	-0.057	0.031
Unobserved Component (α)	0.812	0.942	1	0.007	0.161	-0.012	0.033
Observed Component ($U\eta$)	0.248	0.344	0.007	1	0.067	-0.137	0.034
Pure Firm Effect (ψ)	0.402	0.174	0.161	0.067	1	0.025	-0.101
Time-Varying Covariates ($X\beta$)	0.263	-0.057	-0.012	-0.137	0.025	1	0.036
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.015	0.031	0.033	0.034	-0.101	0.036	1

Random Person, Firm, and Match Effects

	y	θ	α	$U\eta$	ψ	γ	$X\beta$	$\beta_\lambda\lambda$
Log Earnings (y)	1	0.853	0.822	0.248	0.424	0.624	0.262	0.015
Pure Person Effect (θ)	0.853	1	0.929	0.378	0.272	0.507	-0.063	0.031
Unobserved Component (α)	0.822	0.929	1	0.007	0.264	0.544	-0.011	0.033
Observed Component ($U\eta$)	0.248	0.378	0.007	1	0.075	0.009	-0.140	0.034
Pure Firm Effect (ψ)	0.424	0.272	0.264	0.075	1	0.263	0.030	-0.155
Pure Match Effect (γ)	0.624	0.507	0.544	0.009	0.263	1	0.006	0.056
Time-Varying Covariates ($X\beta$)	0.262	-0.063	-0.011	-0.140	0.030	0.006	1	0.035
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.015	0.031	0.033	0.034	-0.155	0.056	0.035	1

Random Person and Firm Effects, Unstructured Residual Covariance

	y	θ	α	$U\eta$	ψ	$X\beta$	$\beta_\lambda\lambda$
Log Earnings (y)	1	0.808	0.780	0.242	0.412	0.262	0.015
Pure Person Effect (θ)	0.808	1	0.910	0.420	0.257	-0.075	0.032
Unobserved Component (α)	0.780	0.910	1	0.006	0.249	-0.012	0.034
Observed Component ($U\eta$)	0.242	0.420	0.006	1	0.073	-0.154	0.035
Pure Firm Effect (ψ)	0.412	0.257	0.249	0.073	1	0.040	-0.155
Time-Varying Covariates ($X\beta$)	0.262	-0.075	-0.012	-0.154	0.040	1	0.032
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.015	0.032	0.034	0.035	-0.155	0.032	1

TABLE 9
PARAMETER ESTIMATES FROM THE PROBIT EQUATION
Combined Results From Three Completed Data Implicates
Standard Errors in Parentheses

	Variance of Person Effect (σ^2_ζ)	Variance of Firm Effect (σ^2_ξ)	Log Likelihood[a]
Tenure 1-2 Years	0.001	0.091	-117856.7
	(0.001)	(0.008)	(2125.5)
Tenure 3-4 Years	0.243	0.678	-101463.3
	(0.007)	(0.014)	(2008.6)
Tenure 5-6 Years	0.307	1.035	-76709.5
	(0.012)	(0.081)	(942.2)
Tenure 7-8 Years	0.294	1.207	-57831.4
	(0.019)	(0.038)	(143.9)
Tenure 9-10 Years	0.231	1.294	-41535.5
	(0.013)	(0.038)	(536.0)
Tenure 11-12 Years	0.149	1.342	-22809.7
	(0.020)	(0.046)	(1727.5)
Tenure 13-14 Years	0.097	1.285	-11970.8
	(0.026)	(0.063)	(1799.1)
Tenure 15+ Years	0.168	1.101	-6284.8
	(0.107)	(0.067)	(1164.4)

[a] Value in parentheses is the square root of the between-implicate variance of lnL.

TABLE 10
MIXED MODEL ESTIMATES OF RESIDUAL COVARIANCE WITHIN WORKER-FIRM MATCH
Dependent Variable: Log(Real Annualized Earnings), No Correction for Truncation, Combined Results From 3 Completed Data Implicates

Tenure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0.126																				
2	0.071	0.093																			
3	0.051	0.067	0.089																		
4	0.041	0.055	0.067	0.089																	
5	0.030	0.044	0.056	0.069	0.085																
6	0.023	0.037	0.049	0.060	0.068	0.087															
7	0.017	0.031	0.043	0.054	0.062	0.071	0.089														
8	0.014	0.027	0.038	0.051	0.058	0.066	0.075	0.093													
9	0.011	0.024	0.035	0.047	0.055	0.062	0.070	0.079	0.096												
10	0.006	0.021	0.032	0.044	0.051	0.059	0.067	0.074	0.083	0.101											
11		0.018	0.029	0.040	0.048	0.056	0.063	0.070	0.077	0.088	0.105										
12			0.023	0.035	0.044	0.051	0.059	0.067	0.073	0.080	0.089	0.105									
13				0.033	0.043	0.048	0.054	0.062	0.068	0.075	0.081	0.091	0.109								
14					0.037	0.042	0.046	0.054	0.061	0.068	0.074	0.082	0.096	0.116							
15						0.030	0.034	0.042	0.047	0.055	0.060	0.068	0.079	0.093	0.119						
16							0.015	0.022	0.027	0.035	0.040	0.047	0.056	0.071	0.087	0.124					
17								0.004	0.007	0.013	0.016	0.022	0.028	0.040	0.057	0.087	0.128				
18									0.001	0.006	0.007	0.011	0.016	0.025	0.037	0.061	0.093	0.111			
19										0.003	0.004	0.006	0.009	0.016	0.025	0.040	0.066	0.081	0.097		
20											0.003	0.004	0.006	0.011	0.018	0.030	0.046	0.059	0.070	0.086	
21												0.003	0.005	0.008	0.012	0.021	0.034	0.042	0.051	0.062	0.080

TABLE 11
MIXED MODEL ESTIMATES OF RESIDUAL COVARIANCE WITHIN WORKER-FIRM MATCH
Dependent Variable: Log(Real Annualized Earnings), Corrected for Truncation, Combined Results From 3 Completed Data Implicates

Tenure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0.126																				
2	0.071	0.093																			
3	0.051	0.067	0.088																		
4	0.040	0.055	0.066	0.088																	
5	0.030	0.044	0.055	0.068	0.085																
6	0.023	0.037	0.048	0.059	0.068	0.086															
7	0.017	0.031	0.043	0.054	0.061	0.071	0.089														
8	0.014	0.027	0.038	0.050	0.058	0.066	0.074	0.093													
9	0.011	0.024	0.034	0.047	0.054	0.062	0.069	0.079	0.096												
10	0.007	0.021	0.032	0.043	0.051	0.059	0.067	0.075	0.083	0.101											
11		0.019	0.029	0.040	0.048	0.056	0.064	0.071	0.078	0.088	0.106										
12			0.024	0.036	0.045	0.052	0.060	0.068	0.074	0.081	0.090	0.107									
13				0.034	0.044	0.049	0.055	0.063	0.069	0.076	0.083	0.094	0.112								
14					0.038	0.043	0.048	0.056	0.062	0.070	0.076	0.085	0.098	0.119							
15						0.031	0.036	0.043	0.049	0.056	0.061	0.070	0.081	0.096	0.121						
16							0.017	0.024	0.028	0.037	0.042	0.049	0.059	0.073	0.090	0.127					
17								0.005	0.009	0.014	0.018	0.023	0.030	0.042	0.059	0.089	0.130				
18									0.002	0.007	0.008	0.012	0.018	0.027	0.039	0.062	0.094	0.113			
19										0.004	0.005	0.007	0.010	0.017	0.026	0.042	0.067	0.082	0.098		
20											0.004	0.005	0.007	0.012	0.019	0.031	0.047	0.060	0.071	0.087	
21												0.004	0.005	0.009	0.013	0.022	0.035	0.042	0.052	0.063	0.081

TABLE 12
MIXED MODEL ESTIMATES OF RESIDUAL COVARIANCE WITHIN WORKER-FIRM MATCH
Dependent Variable: Real Annualized Earnings (Unit Variance Scale), No Correction for Truncation, Combined Results From 3 Completed Data Implicates

Tenure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0.183																				
2	0.069	0.164																			
3	0.060	0.116	0.195																		
4	0.060	0.115	0.167	0.289																	
5	0.054	0.099	0.152	0.226	0.302																
6	0.039	0.109	0.152	0.221	0.255	0.369															
7	0.056	0.116	0.163	0.235	0.264	0.327	0.443														
8	0.070	0.095	0.146	0.227	0.266	0.319	0.380	0.495													
9	0.075	0.108	0.153	0.227	0.265	0.309	0.370	0.435	0.551												
10	0.064	0.113	0.155	0.224	0.266	0.312	0.375	0.430	0.485	0.612											
11		0.114	0.157	0.231	0.285	0.324	0.391	0.443	0.495	0.570	0.699										
12			0.139	0.222	0.262	0.315	0.380	0.434	0.483	0.542	0.608	0.732									
13				0.207	0.306	0.322	0.371	0.426	0.476	0.541	0.613	0.657	0.883								
14					0.287	0.320	0.348	0.388	0.440	0.504	0.583	0.634	0.754	0.932							
15						0.294	0.327	0.360	0.395	0.467	0.530	0.589	0.699	0.752	1.032						
16							0.231	0.254	0.305	0.354	0.425	0.457	0.625	0.695	0.734	0.974					
17								0.151	0.204	0.225	0.271	0.324	0.408	0.515	0.537	0.677	0.800				
18									0.147	0.165	0.203	0.231	0.284	0.359	0.363	0.467	0.520	0.519			
19										0.095	0.119	0.134	0.165	0.211	0.212	0.275	0.309	0.309	0.349		
20											0.071	0.080	0.098	0.125	0.126	0.163	0.183	0.183	0.205	0.287	
21												0.047	0.058	0.074	0.074	0.096	0.107	0.107	0.120	0.168	0.266

TABLE 13
MIXED MODEL ESTIMATES OF RESIDUAL COVARIANCE WITHIN WORKER-FIRM MATCH
Dependent Variable: Real Annualized Earnings (Unit Variance Scale), Corrected for Truncation, Combined Results From 3 Completed Data Implicates

Tenure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0.183																				
2	0.069	0.164																			
3	0.060	0.116	0.194																		
4	0.060	0.115	0.167	0.289																	
5	0.054	0.099	0.152	0.226	0.302																
6	0.039	0.109	0.152	0.221	0.255	0.370															
7	0.056	0.117	0.163	0.235	0.264	0.328	0.443														
8	0.071	0.095	0.146	0.227	0.266	0.319	0.381	0.497													
9	0.076	0.109	0.154	0.228	0.265	0.310	0.371	0.436	0.553												
10	0.066	0.114	0.156	0.225	0.267	0.314	0.376	0.431	0.487	0.614											
11		0.117	0.159	0.233	0.286	0.326	0.393	0.445	0.497	0.572	0.701										
12			0.141	0.224	0.264	0.318	0.382	0.437	0.487	0.546	0.612	0.738									
13				0.210	0.309	0.325	0.374	0.430	0.480	0.545	0.619	0.663	0.890								
14					0.294	0.326	0.354	0.395	0.447	0.512	0.591	0.644	0.765	0.946							
15						0.304	0.337	0.370	0.407	0.480	0.544	0.604	0.719	0.773	1.059						
16							0.250	0.274	0.326	0.379	0.454	0.488	0.672	0.746	0.782	1.044					
17								0.173	0.230	0.252	0.299	0.364	0.454	0.569	0.587	0.742	0.873				
18									0.166	0.186	0.224	0.260	0.317	0.399	0.399	0.514	0.572	0.553			
19										0.108	0.132	0.152	0.186	0.235	0.234	0.304	0.341	0.331	0.361		
20											0.079	0.091	0.110	0.140	0.139	0.180	0.202	0.196	0.213	0.291	
21												0.054	0.065	0.082	0.082	0.106	0.119	0.115	0.125	0.170	0.267

TABLE 14

Minimum Distance Estimates of Structural Parameters, Scale Parameter[1] $\delta=1$
Combined Results From 3 Completed Data Implicates

Variance Parameter	Log(Real Annualized Earnings) Scale (1) Maximum Tenure = 21 years		Real Annualized Earnings (Unit Variance) Scale (2) Maximum Tenure = 21 years		(3) Maximum Tenure = 10 years	
	Parameter Estimate[a]	Standard Error[b]	Parameter Estimate[a]	Standard Error[b]	Parameter Estimate[a]	Standard Error[b]
No Correction for Truncation						
Variance of Measurement Error (σ^2_u)	0.050	(0.000)	0.201	(0.015)	0.093	(0.001)
Variance of Match Quality (σ^2_c)	0.052	(0.002)	0.383	(0.036)	0.580	(0.018)
Variance of Initial Signal (σ^2_z)	0.023	(0.002)	8.661	(3.027)	11.422	(1.506)
Variance of Production Outcomes (σ^2_e)	0.004	(0.006)	0.228	(0.014)	2.549	(0.117)
P-value from Chi-Square Test[c]	0.001		0.002		0.026	
Corrected for Truncation						
Variance of Measurement Error (σ^2_u)	0.050	(0.000)	0.202	(0.015)	0.092	(0.003)
Variance of Match Quality (σ^2_c)	0.053	(0.003)	0.393	(0.016)	0.573	(0.007)
Variance of Initial Signal (σ^2_z)	0.024	(0.003)	9.331	(2.363)	11.235	(1.497)
Variance of Production Outcomes (σ^2_e)	0.004	(0.006)	0.226	(0.014)	2.440	(0.119)
P-value from Chi-Square Test[c]	0.001		0.002		0.026	

[1] Parameters estimates can be rescaled for alternate values $0<\delta<1$ by dividing the pararameter estimate by δ^2,

[a] Average of parameter estimates across three completed data implicates.

[b] Square root of total variance of parameter estimate across three completed data implicates, as defined in Rubin (1987).

[c] Details on combining the chi-square test statistic across completed data implicates can be found in the text, or in Rubin (1987, pp. 76-79). The chi-square test statistic has 161 degrees of freedom for models labeled "Maximum Tenure = 21 years", and 51 degress of freedom for the model labeled "Maximum Tenure = 10 years".

Figure 1
Yearly Time Series of Average Real Annualized Earnings

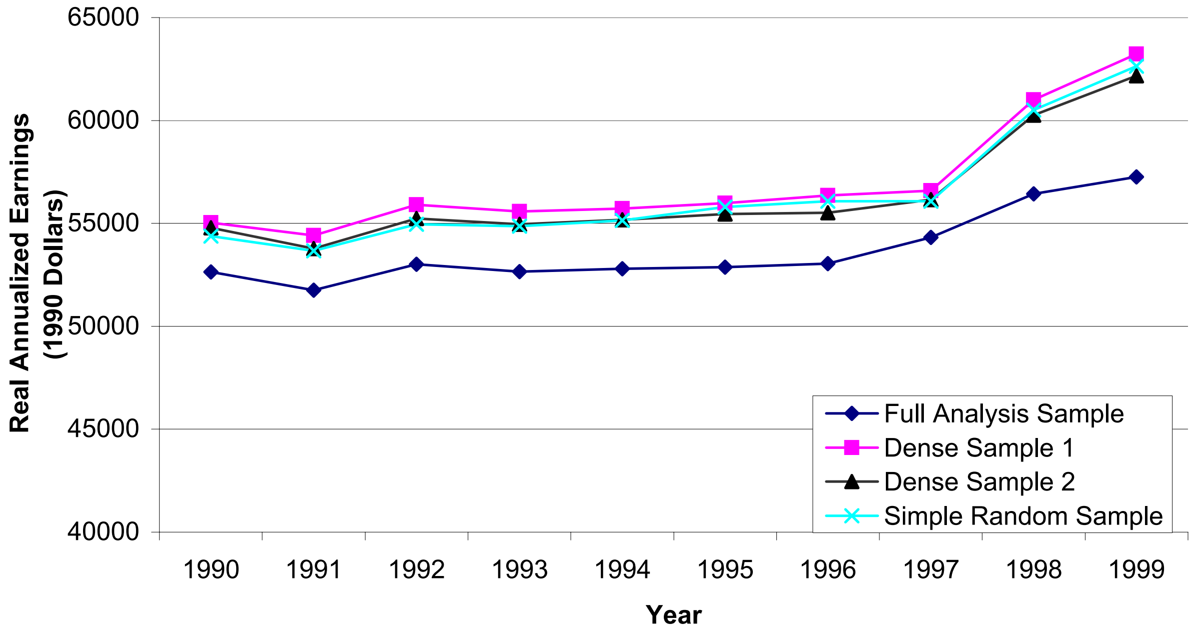

Figure 2
Yearly Time Series of Employment

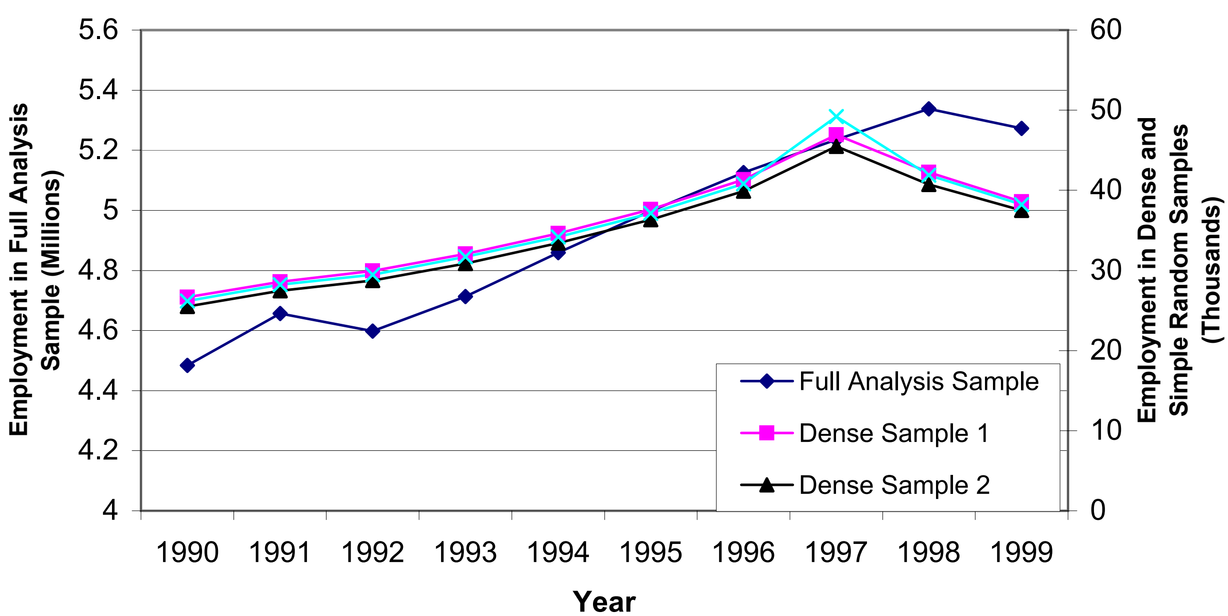

Figure 3
Estimated Sequence of Belief Variances, Log Earnings Scale

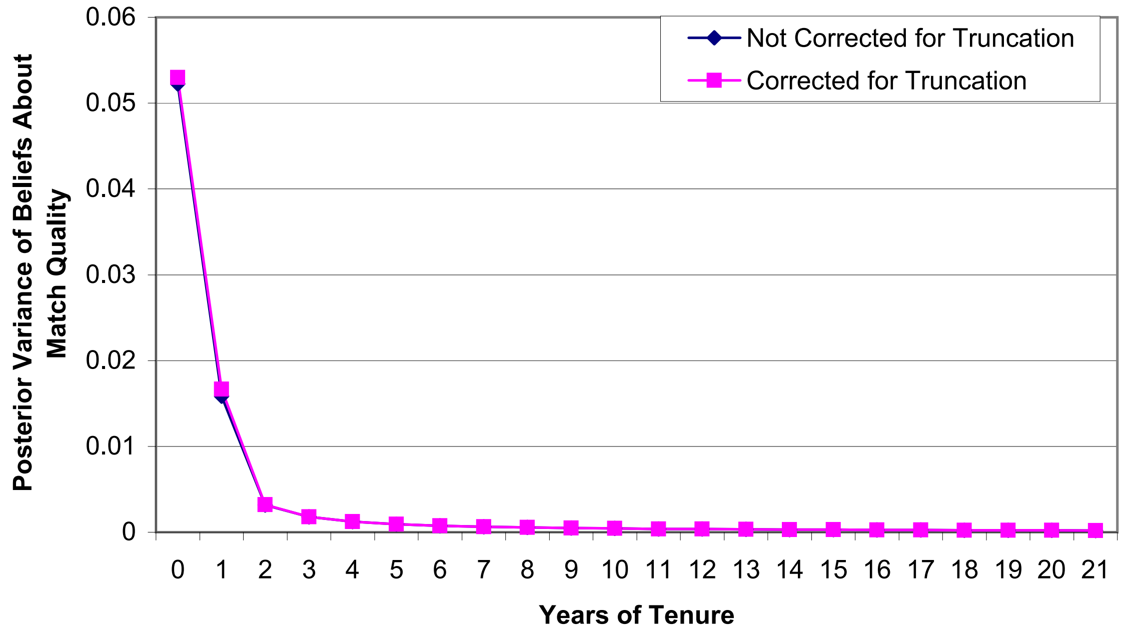

Figure 4
Estimated Sequence of Belief Variances, Earnings (Unit Variance) Scale

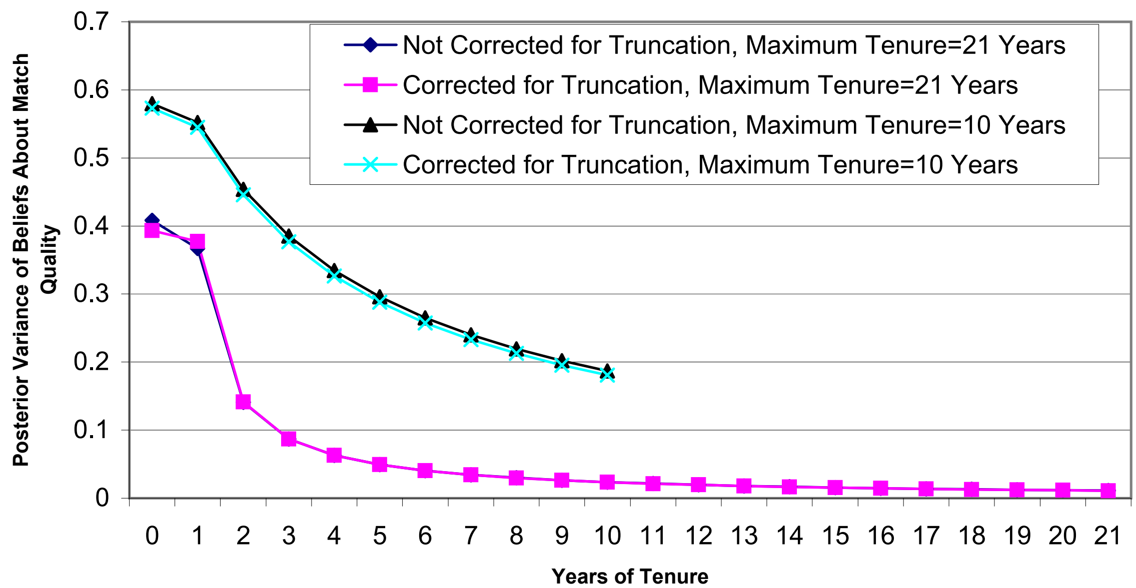

Figure 5
Estimated Relationship Between Person Effect and Job Duration
Model With Unrestricted Within-Match Residual Covariance
Right-Censored Spells Excluded From Regression (N=54,661)

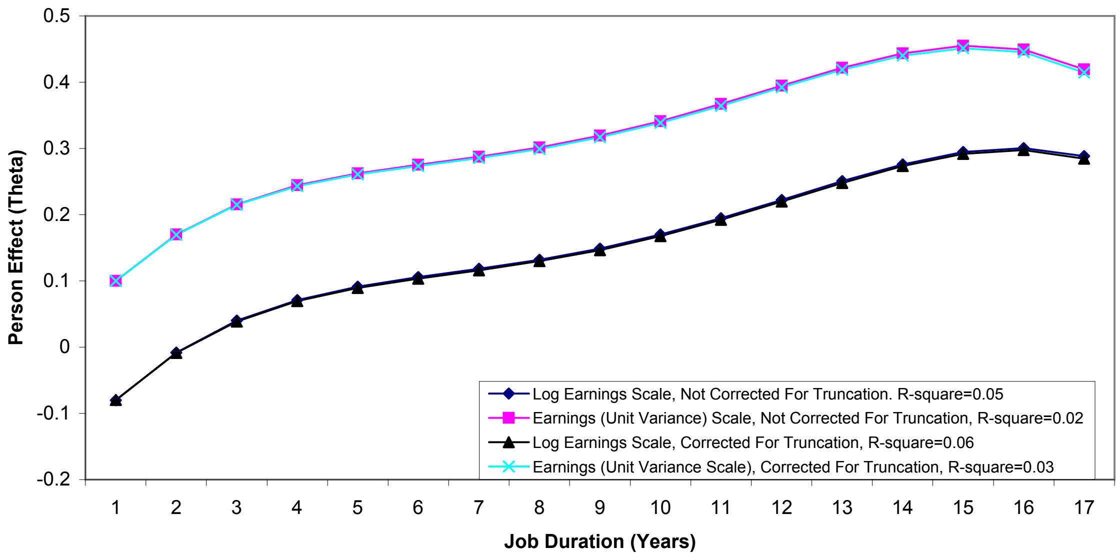

Figure 6
Estimated Relationship Between Firm Effect and Job Duration
Model With Unrestricted Within-Match Residual Covariance
Right-Censored Spells Excluded From Regression (N=54,661)

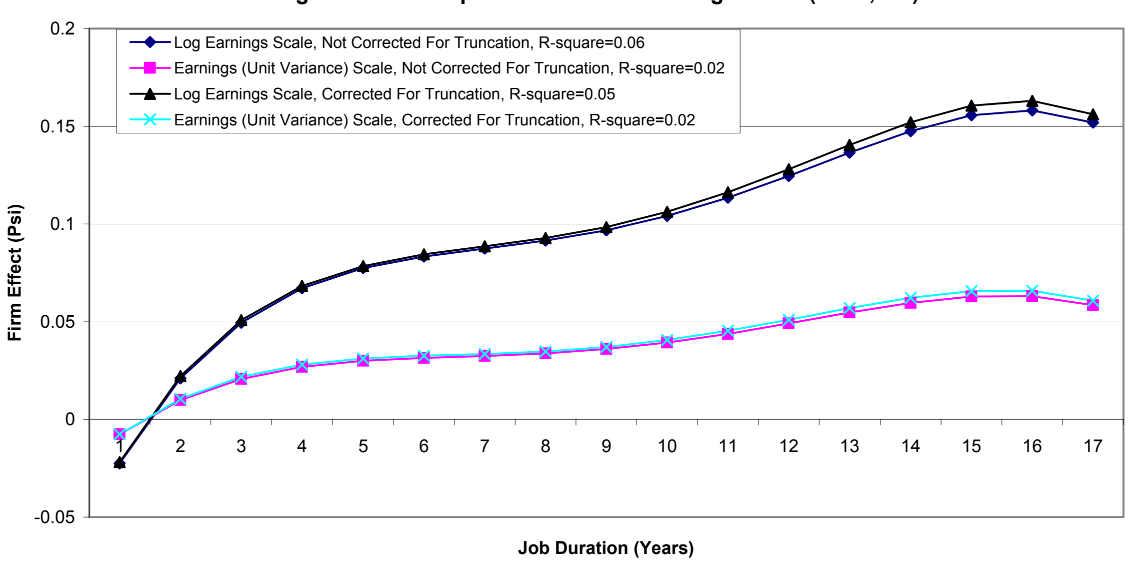

Figure 7
Estimated Relationship Between Firm Effect and Log(1997 Employment)
Model With Unrestricted Within-Match Residual Covariance (N=27,421)

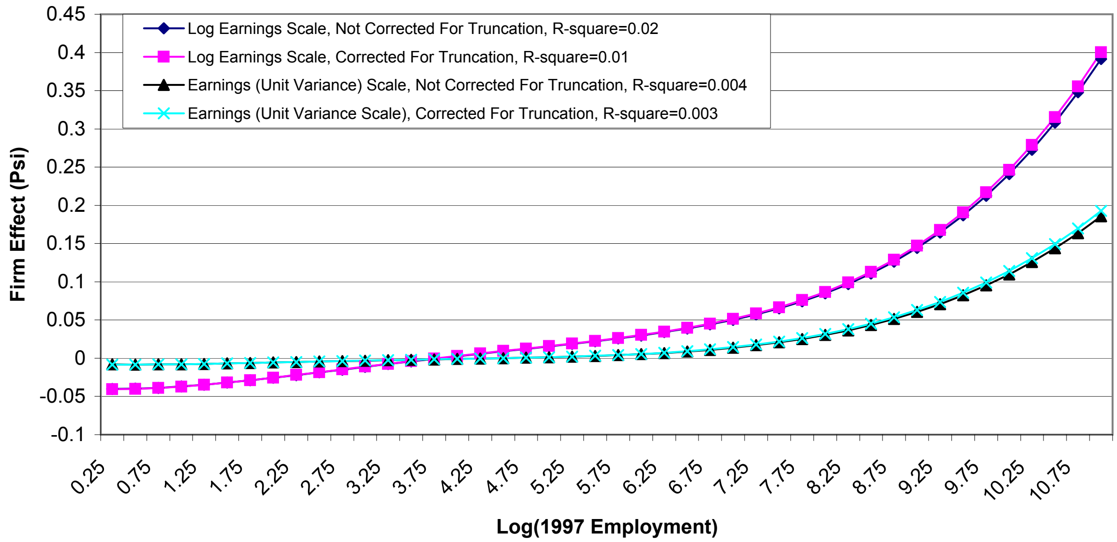

Figure 8
An Example of the Firm-Specific Sampling Rate in the Dense Sampling Algorithm