Longitudinal Employer - Household Dynamics

Technical paper No. TP-2002-20

Agent Heterogeneity and Learning: An Application to Labor Markets

Date	:	October 2002
Prepared by	:	Simon D Woodcock
Contact	:	Ronald Prevost (Ronald.C.Prevost@census.gov) U.S. Census Bureau, LEHD Program FB 2138-3 4700 Silver Hill Rd. Suitland, MD 20233 USA

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, and is released to inform interested parties of ongoing research and to encourage discussion of work in progress. This research is a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation 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(s) and do not represent the views of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau is preparing to support external researchers' use of these data; please contact Ronald Prevost (Ronald.C.Prevost@census.gov), U.S. Census Bureau, LEHD Program, FB 2138-3, 4700 Silver Hill Rd., Suitland, MD 20233, USA.

Agent Heterogeneity and Learning: An Application to Labor Markets*

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

Job Market Paper

October 30, 2002

*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. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the U.S. Census Bureau, or the National Science Foundation. Confidential data from the LEHD Program were used in this paper. The U.S. Census Bureau is preparing to support external researchers' use of these data under a protocol to be released in the near future; please contact Ron Prevost: Ronald.C.Prevost@census.gov.

†I would like to thank John M. Abowd, David Easley, George Jakubson, Martha Stinson, seminar participants at Cornell University and the 2002 CEA meetings, and members of the LEHD program staff 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 and their value of leisure when unemployed. 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 separation decision depends only on the posterior mean of beliefs and person and firm characteristics. These results have several implications for an empirical model of earnings with person and firm effects. 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 and test a variety of more parsimonious error structures, including the martingale structure implied by the learning process. I find considerable support for the matching model in these data.

1 Introduction

Economists have long recognized that observationally indistinguishable workers employed in seemingly identical firms earn different wages and have vastly different employment histories. At best, observable worker and firm characteristics explain 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 worker and firm 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. Furthermore, the separation decision depends only on the posterior mean of beliefs about match quality and worker and firm characteristics.

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 are a martingale. The latter implies a specific non-zero covariance for earnings residuals within a worker-firm match. In the most general empirical specification, I allow wage residuals to have a completely unstructured covariance within-match. I estimate and test a variety of more parsimonious error structures, including the martingale hypothesis of the matching model. 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] In general, 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 (forthcoming), 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 increase in the variance of earnings with labor market experience. The seminal Jovanovic (1979) matching model considered the case where identical workers and firms learn about the quality of a match. 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. Section 3 develops the econometric specification. Section 4 gives a detailed description of the data. Section 5 presents the results, and Section 6 concludes.

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.

[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.

Workers are identified by the continuous index i. In each period, workers are endowed with a single indivisible unit of labor that they supply to home production or production in a firm. They are heterogeneous in their marginal productivity when employed, denoted a_i. I will refer to a_i as worker quality. Assume

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

where F_a is a probability distribution known to all agents, and with support $[\overline{a}, \underline{a}]$. Note a_i is not a choice variable, and there is no human capital accumulation over the life cycle. Workers are heterogeneous in the value of home production when unemployed, denoted $h_i \in \mathbb{R}$.[2] Assume a_i and h_i are exogenous, known to the worker, and observable to the firm when worker and firm meet. Workers seek to maximize the expected present value of wages. Workers can be employed at only one firm each period.

Firms are identified by the continuous index j. They employ many workers. Firms operate in a competitive output market and produce a homogeneous good. The price of output is normalized to 1. Output can only be produced by a worker matched to a firm. Thus firms own some unmodeled input that is essential for production. Firms seek to maximize the expected net revenues of a match: the value of output minus a wage payment to the worker.

Firms are heterogeneous in their technology, denoted b_j, which affects the marginal productivity of all their employees. Assume b_j has support $\left[\overline{b}, \underline{b}\right]$ and

$$b_j \sim F_b \text{ iid across firms} \tag{2}$$

where F_b is a probability distribution known to all agents. I will refer to b_j as firm quality. Firms are heterogeneous in their cost of maintaining a vacancy, denoted $k_j \in \mathbb{R}$. Assume that firms know their own values of b_j and k_j, and that these parameters are observable by the worker when worker and firm meet. Both b_j and k_j are exogenous. Firms incur cost $c\left(l_j\right)$ to hire l_j workers in the current period. Assume c is continuous, increasing, and convex.

Unemployed workers are matched to firms with open vacancies. Search is undirected. The total number of matches formed in a period is given by $m\left(u, v\right)$ 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 $\pi \equiv m\left(u, v\right)/u$. Similarly, the probability that a randomly selected vacancy will be filled is $\lambda \equiv m\left(u, v\right)/v$. With a large number of workers and firms, all agents take u and v as given.

Worker-firm matches are heterogeneous in their marginal productivity, denoted c_{ij}. Assume

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

[2]Assume for simplicity that h_i includes all search costs, the value of leisure, etc.

I call c_{ij} match quality and use F_c to denote the normal distribution function in (3). The normality assumption follows Jovanovic (1979) and others. Match quality c_{ij} is unobserved. The worker and firm learn its value slowly. When a worker and firm first meet, they observe a noisy signal of match quality $x_{ij} = c_{ij} + z_{ij}$ where

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

Let F_z denote the normal distribution function in (4). The worker and firm form beliefs about the value of c_{ij} on the basis of a prior and the signal x_{ij}. They subsequently update their beliefs about c_{ij} 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_{ij} is unobserved, but is symmetric. That is, the worker and firm both know a_i and b_j and have common beliefs about c_{ij}.

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

$$q_{ij\tau} = \mu + a_i + b_j + c_{ij} + e_{ij\tau} \tag{5}$$

where τ indexes tenure (the duration of the match), μ is the grand mean of productivity (known to all agents), and $e_{ij\tau}$ is a match-specific idiosyncratic shock. As a normalization, I use tenure $\tau = 1$ to refer to the period in which the match forms, i.e., before any production has taken place. Assume

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

The linear production technology (5) generalizes that of Jovanovic (1979) to the case of heterogeneous workers and firms in a discrete time setting. Since a_i, b_j, and μ are known, agents extract the noisy signal of match quality $c_{ij} + e_{ij\tau}$ from production outcomes $q_{ij\tau}$.

Within-period timing is as follows:

1. Unemployed workers are randomly matched to a firm with an open vacancy. Upon meeting, agents observe a_i, b_j, and the signal x_{ij}.

2. Workers and firms decide whether or not to continue the match. The continuation decision is based on all current information about the match: a_i, h_i, b_j, k_j and current beliefs about c_{ij}. The current period wage $w_{ij\tau}$ is simultaneously determined by a Nash bargain.

3a. If agents decide to terminate the match, the worker enters unemployment and receives h_i. The firm incurs vacancy cost k_j.

3b. If agents decide to continue the match, output $q_{ij\tau}$ is produced and observed by both parties. Agents then update their beliefs about c_{ij} and the negotiated wage is paid to the worker.

4. Firms open new vacancies v_j.

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_i, b_j, c_{ij}, z_{ij}, and $e_{ij\tau}$ are governed by equations (1), (2), (3), (4), and (6). Recall a_i and b_j are learned when the match is formed. Agents update their beliefs about match quality using Bayes' rule when they acquire new information, i.e., upon observing the signal x_{ij} and production outcomes $q_{ij\tau}$.

After observing the signal x_{ij}, worker and firm posterior beliefs about c_{ij} are normally distributed with mean m_{ij1} and variance s_1^2 where

$$m_{ij1} = x_{ij}\left(\frac{\sigma_c^2}{\sigma_c^2+\sigma_z^2}\right) \tag{7}$$

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

In each subsequent period that the match persists, the worker and firm extract the signal $c_{ij}+e_{ij\tau}$ from observed output $q_{ij\tau}$. 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 $m_{ij\tau}$ and variance s_τ^2, where

$$\begin{aligned} m_{ij\tau} &= \left(\frac{m_{ij\tau-1}}{s_{\tau-1}^2}+\frac{c_{ij}+e_{ij\tau-1}}{\sigma_e^2}\right)\Big/\left(\frac{1}{s_{\tau-1}^2}+\frac{1}{\sigma_e^2}\right) \\ &= \left(\frac{x_{ij}}{\sigma_z^2}+\sum_{s=1}^{\tau-1}\frac{c_{ij}+e_{ijs}}{\sigma_e^2}\right)\Big/\left(\frac{\sigma_c^2+\sigma_z^2}{\sigma_c^2\sigma_z^2}+\frac{\tau-1}{\sigma_e^2}\right) \end{aligned} \tag{9}$$

$$\begin{aligned} \frac{1}{s_\tau^2} &= \frac{1}{s_{\tau-1}^2}+\frac{1}{\sigma_e^2} \\ &= \frac{\sigma_c^2+\sigma_z^2}{\sigma_c^2\sigma_z^2}+\frac{\tau-1}{\sigma_e^2}. \end{aligned} \tag{10}$$

Clearly the evolution of s_τ^2 is deterministic and does not depend on the value of the signals received. Equation (9) says that the updated posterior mean of beliefs $m_{ij\tau}$ is a precision-weighted average of the prior mean $m_{ij\tau-1}$ and the signal $c_{ij}+e_{ij\tau-1}$. 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,

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

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

which is a standard result for Bayesian learnings with "correct" priors (see e.g., Blume and Easley (1998)). These two equations imply that asymptotically, beliefs converge to point mass at true match quality.

In what follows, it will be of interest to describe the distribution of beliefs in the population. It is a standard result that the unconditional distribution of $m_{ij\tau}$ is normal with mean zero and variance V_τ, where

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

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.

Denote the complete set of tenure-τ information about the productivity of a match by $\Omega_{ij\tau} = (a_i, b_j, m_{ij\tau}, s_\tau^2)$. It is notationally convenient to describe the evolution of information about the productivity of a match using a transition distribution $F(\Omega_{ij,\tau+1} | \Omega_{ij\tau})$. Both a_i and b_j are fixed, so the transition distribution describes the probabilistic evolution of $m_{ij,\tau+1}$ given $m_{ij\tau}$, and the deterministic evolution of $s_{\tau+1}^2$ given s_τ^2. It is straightforward to show that

$$m_{ij,\tau+1} | \Omega_{ij\tau} \sim N\left(m_{ij\tau}, \frac{s_\tau^4}{s_\tau^2 + \sigma_e^2} \right). \quad (14)$$

More generally, for any $p > \tau$

$$m_{ijp} | \Omega_{ij\tau} \sim N\left(m_{ij\tau}, \frac{s_\tau^4 (p - \tau)}{s_\tau^2 (p - \tau) + \sigma_e^2} \right). \quad (15)$$

Note that (14) and (15) imply the posterior mean of beliefs is a martingale. Conditional on current information, expectations about future realizations of the random variable $m_{ij\tau}$ 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. In evaluating the expected joint surplus of the match, worker and firm expectations are taken with respect to tenure-τ information about the productivity of the match, $\Omega_{ij\tau}$. It follows that equilibrium wages map tenure-τ information about the match into payments from firm to worker. To reflect this, I write the tenure-τ equilibrium wage as $w_{ij\tau} = w(\Omega_{ij\tau})$ for each $\tau > 0$. I assume the function w is known to all agents.

It is a consequence of the Nash bargain that w is an equilibrium function in the following sense: it is chosen so that workers and firms agree about the set of acceptable job matches.[3]

[3]See Pissarides (1984).

It follows that for a match between worker i and firm j that has lasted τ periods, there is a reservation level of beliefs about match quality, $\bar{m}_{ij\tau}$, below which the match dissolves and above which the match continues.

In deriving the equilibrium wage and the expected joint surplus of the match, I use the following notation: $J\left[w\left(\Omega_{ij\tau}\right)\right]$ is the expected value to worker i of employment at firm j given the wage function w and information $\Omega_{ij\tau}$; U_i is the value of worker i's outside option (unemployment); $\Pi\left[w\left(\Omega_{ij\tau}\right)\right]$ is the expected value to firm j of net revenues from a match with worker i given w and $\Omega_{ij\tau}$; and V_j is the value of firm j's outside option (a vacancy). At tenure τ, the match continues if and only if

$$J\left[w\left(\Omega_{ij\tau}\right)\right]+\Pi\left[w\left(\Omega_{ij\tau}\right)\right] \geq U_i+V_j. \tag{16}$$

When (16) is satisfied, the equilibrium wage $w_{ij\tau}=w\left(\Omega_{ij\tau}\right)$ solves the Nash bargaining wage condition

$$J\left[w\left(\Omega_{ij\tau}\right)\right]-U_i=\delta\left(J\left[w\left(\Omega_{ij\tau}\right)\right]+\Pi\left[w\left(\Omega_{ij\tau}\right)\right]-U_i-V_j\right) \tag{17}$$

or equivalently,

$$(1-\delta)\left(J\left[w\left(\Omega_{ij\tau}\right)\right]-U_i\right)=\delta\left(\Pi\left[w\left(\Omega_{ij\tau}\right)\right]-V_j\right) \tag{18}$$

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

In solving for the equilibrium wage and employment condition, I use the following strategy. First, I make two conjectures regarding the structure of equilibrium wages and employment. Then I characterize the various value functions under the assumption that the conjectures are true. Finally, I show that equilibrium wages and employment satisfy the two conjectures. Conjecture 1 greatly simplifies expectations over future values of $w\left(\Omega_{ij\tau}\right)$. Recall that $m_{ij\tau}$ and s_τ^2 are the posterior mean and variance of tenure-τ beliefs about match quality, respectively.

Conjecture 1 *The equilibrium wage offer function w is linear in $m_{ij\tau}$ and independent of s_τ^2.*

Conjecture 2 concerns the reservation level of beliefs about match quality, $\bar{m}_{ij\tau}$. We shall see that as a consequence of the Nash bargain and the linearity property embodied in Conjecture 1, the match terminates when the posterior mean of beliefs about match quality falls below $\bar{m}_{ij\tau}$.

Conjecture 2 *The reservation level of beliefs about match quality, $\bar{m}_{ij\tau}$, is independent of tenure. That is, $\bar{m}_{ij\tau}=\bar{m}_{ij}$ for all $\tau>0$.*

I prove that Conjectures 1 and 2 are true in Propositions 6 and 7.

Before deriving the various value functions, I need to introduce some final notation. Let $G_p(x)=\Pr\left(m_{ijp}<x|\Omega_{ij\tau}\right)$ for $p>\tau$. Then $G_p\left(\bar{m}_{ijp}\right)$ is the subjective probability that the match will terminate at tenure p, given tenure-τ information about the productivity of the match. Note G_p is just the normal distribution given by (15). I use E_τ to denote an expectation taken with respect to tenure-τ information. Tenure-τ expectations of tenure $\tau+1$ quantities are taken with respect to the transition distribution of information, denoted $F\left(\Omega_{ij\tau+1}|\Omega_{ij\tau}\right)$, defined previously.

2.2.1 The Worker's Value of Employment and Unemployment

The expected value to worker i of employment with firm j at wage $w_{ij\tau} = w\left(\Omega_{ij\tau}\right)$ is today's wage payment plus the discounted expected value of employment next period, adjusted for the possibility that the match terminates. That is,

$$\begin{aligned} J\left[w\left(\Omega_{ij\tau}\right)\right] &= w_{ij\tau} + \beta\left[1 - G_{\tau+1}\left(\bar{m}_{ij,\tau+1}\right)\right] E_\tau J\left[w\left(\Omega_{ij,\tau+1}\right)\right] + \beta G_{\tau+1}\left(\bar{m}_{ij,\tau+1}\right) U_i \\ &= w_{ij\tau} + \beta\left[1 - G_{\tau+1}\left(\bar{m}_{ij,\tau+1}\right)\right] \int J\left[w\left(\Omega_{ij,\tau+1}\right)\right] dF\left(\Omega_{ij\tau+1}|\Omega_{ij\tau}\right) \\ &\quad + \beta G_{\tau+1}\left(\bar{m}_{ij,\tau+1}\right) U_i \end{aligned} \tag{19}$$

When Conjecture 1 is true, $E_\tau w_{ijp} = w_{ij\tau}$ for all $p > \tau$. When Conjecture 2 is also true,

$$\begin{aligned} E_\tau J\left[w\left(\Omega_{ij,\tau+1}\right)\right] &= E_\tau w_{ij,\tau+1} + \beta\left[1 - G_{\tau+2}\left(\bar{m}_{ij}\right)\right] E_\tau E_{\tau+1} J\left[w\left(\Omega_{ij,\tau+2}\right)\right] + \beta G_{\tau+2}\left(\bar{m}_{ij}\right) U_i \\ &= w_{ij\tau} + \beta\left[1 - G_{\tau+2}\left(\bar{m}_{ij}\right)\right] E_\tau J\left[w\left(\Omega_{ij,\tau+2}\right)\right] + \beta G_{\tau+2}\left(\bar{m}_{ij}\right) U_i. \end{aligned} \tag{20}$$

Forward recursion on (19) and (20) gives

$$\begin{aligned} J\left[w\left(\Omega_{ij\tau}\right)\right] &= w_{ij\tau}\left(1 + \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} \prod_{p=\tau+1}^{s} \left[1 - G_p\left(\bar{m}_{ij}\right)\right]\right) \\ &\quad + U_i \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} G_s\left(\bar{m}_{ij}\right) \prod_{p=\tau+1}^{s-1} \left[1 - G_p\left(\bar{m}_{ij}\right)\right]. \end{aligned} \tag{21}$$

Equation (21) says that the expected value of employment is a weighted average of the current wage and the value of unemployment. The weights are discounted employment and separation hazards at each future tenure τ, given current beliefs about match quality.

Deriving the value of unemployment is rather tedious and not particularly instructive. Thus I relegate it to Appendix A. When Conjectures 1 and 2 are true, the value of being unemployed today and behaving optimally thereafter is

$$U_i = \frac{h_i + \beta\pi\left(\sum_{\tau=1}^{\infty} \beta^{\tau-1} \int w\left(\Omega_{ij1}^0\right) \prod_{s=1}^{\tau} \left[1 - G_s\left(\bar{m}_{ij}\right)\right] dF_b\right)}{1 - \beta\left(1-\pi\right) - \pi \sum_{\tau=1}^{\infty} \beta^{\tau} \int G_\tau\left(\bar{m}_{ij}\right) \prod_{s=1}^{\tau-1} \left[1 - G_s\left(\bar{m}_{ij}\right)\right] dF_b} \tag{22}$$

where $\Omega_{ij1}^0 = \left(a_i, b_j, 0, s_1^2\right)$ reflects agents' prior beliefs about match quality, and $\pi = m\left(u,v\right)/u$ is the probability that an unemployed worker is matched to a firm. There is no closed form expression for U_i under the general distribution F_b of firm quality.[4] The numerator in Equation (22) is the sum of the value of home production today and the discounted expected value of employment in subsequent periods when the identity of the employing firm is unknown. That is, before firm quality and the signal x_{ij} are known. The denominator normalizes to account for the possibility of re-entering unemployment at each future tenure.

In each period, workers are either unemployed or employed at a single firm. Let $u_i = 1$ if worker i is unemployed, and zero otherwise. In each period, the number of unemployed in the economy is simply $u = \int_0^1 u_i di$.

2.2.2 The Firm's Value of Employment and of a Vacancy

I now turn to the firm's value of employment. The value to firm j of employing worker i at wage $w_{ij\tau} = w(\Omega_{ij\tau})$ is today's expected net revenues plus the discounted expected value of employment next period, adjusted for the possibility that the match terminates. Thus,

$$\begin{aligned}
\Pi\left[w\left(\Omega_{ij\tau}\right)\right] &= E_\tau q_{ij\tau} - w_{ij\tau} + \beta\left[1 - G_{\tau+1}\left(\bar{m}_{ij,\tau+1}\right)\right] E_\tau \Pi\left[w\left(\Omega_{ij,\tau+1}\right)\right] + \beta G_{\tau+1}\left(\bar{m}_{ij,\tau+1}\right) V_j \\
&= \mu + a_i + b_j + m_{ij\tau} - w_{ij\tau} \\
&\quad + \beta\left[1 - G_{\tau+1}\left(\bar{m}_{ij,\tau+1}\right)\right] \int \Pi\left[w\left(\Omega_{ij,\tau+1}\right)\right] dF\left(\Omega_{ij\tau+1} | \Omega_{ij\tau}\right) \qquad (23) \\
&\quad + \beta G_{\tau+1}\left(\bar{m}_{ij,\tau+1}\right) V_j. \qquad (24)
\end{aligned}$$

Applying Conjectures 1 and 2, forward recursion analogous to (20) and (21) gives

$$\begin{aligned}
\Pi\left[w\left(\Omega_{ij\tau}\right)\right] &= \left(\mu + a_i + b_j + m_{ij\tau} - w_{ij\tau}\right)\left(1 + \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} \prod_{p=\tau+1}^{s} \left[1 - G_p\left(\bar{m}_{ij}\right)\right]\right) \\
&\quad + V_j \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} G_s\left(\bar{m}_{ij}\right) \prod_{p=\tau+1}^{s-1} \left[1 - G_p\left(\bar{m}_{ij}\right)\right]. \qquad (25)
\end{aligned}$$

Equation (25) says that the value of employing worker i is a weighted average of the expected net revenues accruing to the match (expected output minus the current wage) and the value of a vacancy. Like the worker's value of employment, the weights are discounted employment and separation hazards at each future tenure τ, given current beliefs about match quality.

I derive the value of a vacancy in Appendix A. When Conjectures 1 and 2 are true, this is

$$V_j = \frac{\beta\lambda\left(\sum_{\tau=1}^{\infty} \beta^{\tau-1} \int \left[\mu + a_i + b_j - w\left(\Omega_{ij1}^0\right)\right] \prod_{s=1}^{\tau} \left[1 - G_s\left(\bar{m}_{ij}\right)\right] dF_a\right) - k_j}{1 - \beta(1-\lambda) - \lambda \sum_{\tau=1}^{\infty} \beta^{\tau} \int G_\tau\left(\bar{m}_{ij}\right) \prod_{s=1}^{\tau-1} \left[1 - G_s\left(\bar{m}_{ij}\right)\right] dF_a} \qquad (26)$$

[4] There is also no closed form expression for U_i when firm quality is normally distribued. A closed form may exist for other distributions.

where $\lambda = m(u, v)/v$ is the probability that a vacancy is filled. Like the worker's value of unemployment U_i, there is no closed form expression for V_j when worker quality is drawn from the general distribution F_a. The numerator in (26) is the discounted expected value of employing a worker next period before the identity of the matching worker i is known (that is, before worker quality and the signal x_{ij} are known) net of the cost of maintaining the vacancy. The denominator normalizes the value to reflect the possibility of the match terminating at each future tenure.

2.2.3 The Firm's Decision to Open Vacancies

The production technology (5) implies that employees of firm j 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 firm j opens v_j vacancies, we can model the number that are filled, l_j, as a binomial process. It follows that the number of vacancies opened by firm j in a given period solves[5]

$$\max_{v_j \in \mathbb{N}} \sum_{l_j=0}^{v_j} \binom{v_j}{l_j} \lambda^{l_j} (1-\lambda)^{v_j - l_j} \left[l_j \Pi_0(b_j) - c(l_j) \right] - k_j v_j \tag{27}$$

where $\Pi_0(b_j)$ is the expected present value of net revenues from a match for a firm of quality b_j, before the identity of the matching worker i is known. I derive $\Pi_0(b_j)$ in Appendix A.

Note that firm size (employment) is indeterminate. However, increasing and convex hiring costs c guarantee the solution to (27) is well defined and the number of vacancies opened in any period by firm j is finite. At any point in time, the total number of vacancies in the economy is just $v = \int_0^{\phi} v_j dj$.

2.2.4 The Equilibrium Wage

With expressions for the value functions in hand, I can now prove the two Conjectures and derive the equilibrium wage and employment conditions. Define the following terms:

$$A_{ij\tau} = 1 + \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} \prod_{p=\tau+1}^{s} \left[1 - G_p(\bar{m}_{ij})\right] \tag{28}$$

$$B_{ij\tau} = 1 - \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} G_s(\bar{m}_{ij}) \prod_{p=\tau+1}^{s-1} \left[1 - G_p(\bar{m}_{ij})\right]. \tag{29}$$

Substituting the definitions of $A_{ij\tau}$ and $B_{ij\tau}$ into (21), we see the worker's net surplus from the match is:

$$J\left[w(\Omega_{ij\tau})\right] - U_i = w_{ij\tau} A_{ij\tau} - U_i B_{ij\tau}. \tag{30}$$

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

Similarly, the firm's net surplus from the match is:

$$\Pi\left[w\left(\Omega_{ij\tau}\right)\right]-V_j=\left(\mu+a_i+b_j+m_{ij\tau}-w_{ij\tau}\right)A_{ij\tau}-V_jB_{ij\tau}. \tag{31}$$

The following lemmata establish properties of $A_{ij\tau}$ and $B_{ij\tau}$, and are required to prove Conjectures 1 and 2. The proofs of Lemma 3 and Lemma 4 are in Appendix B.

Lemma 3 *Define $A_{ij\tau}$ as in* (28)*. Then $A_{ij\tau}$ is bounded from below by* 1 *and from above by* $(1-\beta)^{-1}$.

Lemma 4 *Define $A_{ij\tau}$ and $B_{ij\tau}$ as in* (28) *and* (29)*. Then*

$$B_{ij\tau}=(1-\beta)\,A_{ij\tau}$$

for all i,j, and $\tau>0$.

Lemma 5 *Define $B_{ij\tau}$ as in* (29)*. Then $B_{ij\tau}$ is bounded from below by* $(1-\beta)$ *and from above by* 1.

Proof. Follows immediately from Lemma 3 and Lemma 4. ■

We can now prove Conjectures 1 and 2, which are restated as Propositions 6 and 7.

Proposition 6 *The equilibrium wage offer function w is linear in $m_{ij\tau}$ and independent of* s^2_τ.

Proof. Substituting (30) and (31) into the Nash bargaining wage condition (18) gives the equilibrium wage:

$$w_{ij\tau}=\delta\left(\mu+a_i+b_j+m_{ij\tau}\right)+\left(\left(1-\delta\right)U_i-\delta V_j\right)\frac{B_{ij\tau}}{A_{ij\tau}} \tag{32}$$

which is well defined by Lemma 3 and Lemma 5. Applying Lemma 4 gives

$$w_{ij\tau}=\delta\left(\mu+a_i+b_j+m_{ij\tau}\right)+\left(\left(1-\delta\right)U_i-\delta V_j\right)\left(1-\beta\right) \tag{33}$$

for all i,j and $\tau>0$. ■

As conjectured, (33) verifies 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 $a_i=b_j=0$ for all i,j. 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 (33) is linear in expected marginal product, $\mu + a_i + b_j + m_{ij\tau}$, and in the posterior mean of beliefs about match quality, $m_{ij\tau}$. A stronger result is that when workers capture all the quasi–rents associated with the match, that is as $\delta \to 1$, so that firms earn zero expected profit; and when vacancies are costless (as in Jovanovic (1979)); then the equilibrium wage converges to $w'_{ij\tau} = \mu + a_i + b_j + m_{ij\tau}$.[6] 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 (33).

Note we can rewrite the Nash bargained wage in (33) as

$$\begin{aligned} w_{ij\tau} &= \delta\mu + \left[\delta a_i + (1-\beta)(1-\delta)U_i\right] + \left[\delta b_j - (1-\beta)\delta V_j\right] + \delta m_{ij\tau} \\ &= \delta\mu + \theta_i + \psi_j + \delta m_{ij\tau} \qquad (34) \end{aligned}$$

where

$$\theta_i = \delta\left[a_i - (1-\beta)U_i\right] + (1-\beta)U_i \qquad (35)$$

$$\psi_j = \delta\left[b_j - (1-\beta)V_j\right]. \qquad (36)$$

The expression (34) shows that equilibrium wages are linear in a worker-specific component θ_i, a firm-specific component ψ_j, and the posterior mean of beliefs about match quality. I make use of this fact in developing the empirical strategy that follows. In light of this, I refer to θ_i and ψ_j as empirical person and firm effects. Equation (36) illustrates that the firm effect is simply the worker's share δ of the firm's contribution to the joint surplus accruing to the match. Similarly, equation (35) demonstrates that the person effect is the worker's share of his contribution to the joint surplus, plus compensation $(1-\beta)U_i$ for forgoing his next-best alternative.

The following Proposition establishes that the reservation level of beliefs about match quality depends only on worker and firm characteristics, and is independent of tenure.

Proposition 7 *The reservation level of beliefs about match quality, $\bar{m}_{ij\tau}$, is independent of tenure. That is, $\bar{m}_{ij\tau} = \bar{m}_{ij}$ for all $\tau > 0$.*

Proof. The reservation level of beliefs about match quality at tenure τ, $\bar{m}_{ij\tau}$, is the level at which parties to the match are indifferent between continuing the match and allowing it to dissolve. That is, the value at which the joint surplus from the match is zero. Thus $\bar{m}_{ij\tau}$ is defined by

$$J\left[w\left(\bar{\Omega}_{ij\tau}\right)\right] + \Pi\left[w\left(\bar{\Omega}_{ij\tau}\right)\right] = U_i + V_j \qquad (37)$$

where $\bar{\Omega}_{ij\tau} = \left(a_i, b_j, \bar{m}_{ij\tau}, s_\tau^2\right)$. Substituting (30) and (31) into (37) yields

$$\left(\mu + a_i + b_j + \bar{m}_{ij\tau}\right) A_{ij\tau} = \left(U_i + V_j\right) B_{ij\tau}$$

[6] This follows from the observation that as $\delta \to 1$ and with $k_j = 0$, the equilibrium value of a vacancy V_j must also be zero.

and applying Lemma 4,

$$\begin{aligned} \bar{m}_{ij\tau} &= (U_i + V_j)(1-\beta) - \mu - a_i - b_j \\ &\equiv \bar{m}_{ij} \end{aligned} \tag{38}$$

for all $\tau > 0$. ■

A final Proposition establishes the relationship between the various components of the equilibrium wage: the empirical person and firm effects θ_i and ψ_j, and the posterior mean of beliefs about match quality $m_{ij\tau}$. I make considerable use of Proposition 8 in the empirical specification that follows.

Proposition 8 *At each tenure $\tau > 0$, $E(m_{ij\tau}\theta_i) = E(m_{ij\tau}\psi_j) = 0$ for all i, j.*

Proof. Follows immediately from the definition of $m_{ij\tau}$ in (7) and (9), the definitions of θ_i and ψ_j in (35) and (36), and the distributional assumptions on a_i, b_j, c_{ij}, z_{ij}, and $e_{ij\tau}$ in (1), (2), (3), (4), and (6). ■

The intuition behind Proposition 8 is simple. The empirical person effect θ_i is a nonlinear function of the worker's marginal product a_i and value of home production h_i. The empirical firm effect is a nonlinear function of the firm's marginal product b_j and cost of maintaining a vacancy k_j. At each tenure, the posterior mean of beliefs about match quality $m_{ij\tau}$ is a function only of priors and the signals of match quality observed up to that point. Priors are common to all agents, and the signals of match quality are independent of a_i, h_i, b_j and k_j, and thus independent of θ_i and ψ_j.

Proposition 8 is particularly convenient for developing an empirical specification based on the equilibrium wage function (34). It says that in an empirical earnings model with fixed or random person and firm effects, we can treat the posterior mean of beliefs as a normally distributed statistical residual with a non-zero covariance within a worker-firm match. Normality follows from the Bayesian updating process (see Section 2.1). The specific form of the within-match residual covariance follows also from the Bayesian learning process, and is discussed at length in Section 3.3.

2.3 Discussion

In developing the matching model, I have alluded several times to the empirical specification that is developed in the next Section. Before turning to empirics, however, it is useful to discuss various predictions that stem from the matching model with regards to equilibrium wages, mobility, turnover, and firm size.

First and foremost, the model predicts that wages are linear in person- and firm-specific components. In keeping with the empirical literature, I have called these empirical person and firm effects, and denoted them θ_i and ψ_j. They are functions of the random variables a_i, h_i, b_j, and k_j, and thus are random variables themselves.

Equilibrium wages are also linear in the posterior mean beliefs about match quality $m_{ij\tau}$. This has a number of implications for the equilibrium distribution of wages and their evolution within a worker-firm match. First, since $m_{ij\tau}$ 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, $m_{ij\tau}$ is a martingale (see Section 2.1). Thus the model predicts that within a worker-firm match, wages are also a martingale. The martingale property is common to most learning models, see e.g. Farber and Gibbons (1996) and Gibbons et al. (2002).[7] Econometric implications of the martingale hypothesis are discussed in Section 3.3. However, the martingale structure also has a number of economic consequences. First, recalling the definition of $m_{ij\tau}$ in equation (9), shocks to beliefs about match quality (z_{ij} and $e_{ij\tau}$) are permanent. Within a worker-firm match, these are the only shocks to wages.[8] Thus wage shocks are permanent. Second, on average, wage shocks diminish with tenure. To see this, recall that $m_{ij\tau}$ is a precision-weighted average of $m_{ij\tau-1}$ and the signal $c_{ij} + e_{ij\tau}$. The precision of the shocks is constant, but the precision of beliefs increases with tenure. Thus as agents learn about match quality, each successive shock receives smaller weight in the updating process.[9] Third, within a worker-firm match the variance of earnings increases with tenure. This arises because the variance of $m_{ij\tau}$ 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 which declines with tenure, and the variance of the posterior mean of beliefs V_τ, defined in equation (13), which increases with tenure.[10] That the variance of earnings increases with tenure is broadly consistent with the empirical finding that the variance of earnings increases with labor market experience (see e.g., Mincer (1974)).

Under the matching model, employment relationships terminate when the posterior mean of beliefs about match quality fall below the match-specific threshold $\bar{m}_{ij}$, defined in (38). Comparative statics on (38) characterize the relationship between employment duration and the worker and firm marginal products a_i and b_j. Unfortunately the worker and firm outside options U_i and V_j are themselves complex functions of a_i and b_j, which complicates signing

[7] In these two papers, firms and matches are homogeneous. Workers vary in their ability, which is unknown 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] Extending the model to include aggregate and/or firm-specific shocks to productivity is left for future research.

[9] Asymptotically, shocks receive zero weight in the update.

[10] 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. However, since the posterior mean of beliefs $m_{ij\tau}$ is a function of all the signals received, its variance increases with tenure. Each signal is a random variable (with common variance). Thus each successive signal contributes to the variance of $m_{ij\tau}$.

the derivatives. Nevertheless, (38) makes clear that the effects of a_i and b_j on expected duration are symmetric.

Comparative statics on (38) with respect to the empirical person and firm effects θ_i and ψ_j are pertinent to empirical work. It is easy to show that

$$\frac{\partial \bar{m}_{ij}}{\partial \psi_j} = -\frac{1}{\delta} < 0 \tag{39}$$

$$\left.\frac{\partial \bar{m}_{ij}}{\partial \theta_i}\right|_{a_i=0} = -\frac{1}{1+\delta} < 0. \tag{40}$$

Note that lower values of $\bar{m}_{ij}$ 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_j. A corollary is that firms with larger values of ψ_j experience less turnover than firms with smaller effects. Since firms with large firm effects are better able to retain workers, *ceteris paribus* these firms will have larger employment at any point in time than firms with smaller values. This is consistent with the empirical finding that conditional on observable worker and firm characteristics, larger firms pay more on average than smaller firms (see e.g. Brown and Medoff (1989)). Abowd et al. (1999) find that estimated firm-size wage effects are well explained by firm-size category average person and firm effects.[11]

Similar predictions arise from equation (40). On average, workers with high values of θ_i enjoy longer job duration and change jobs less often. Lillard (1999) finds a comparable result in NLSY data.[12] Combining (39) and (40), empirically we should expect the duration-weighted correlation between θ_i and ψ_j to be positive. Abowd et al. (2002) find the reverse is true in France and in the State of Washington.[13]

The model also predicts that in a cross-section, workers with longer job tenure will on average be observed to earn higher wages 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.[14] The result stems

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

[12] 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 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, the finds the reverse in 1996 SIPP data.

[13] In earlier work that estimated approximate values of fixed worker and firm effects, Abowd et al. (1999) found a positive duration-weighted correlation between estimated person and firm effects, as predicted by the matching model.

[14] 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 Robert A (1987), and Topel

from two observations. First, higher values of θ_i and ψ_j are on average associated with longer duration. Second, conditional on θ_i and ψ_j, longer lasting matches are on average associated with higher match quality c_{ij}, and thus with larger values of $m_{ij\tau}$ at each tenure. Since θ_i, ψ_j, and $m_{ij\tau}$ all enter positively into wages, the result follows.

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 $m_{ij\tau}$ enters the equilibrium wage, but the employment relationship only continues as long as $m_{ij\tau} \geq \bar{m}_{ij}$. Thus the observed distribution of earnings is truncated: earnings outcomes are only observed if $w_{ij\tau} \geq \mu + \theta_i + \psi_j + \bar{m}_{ij}$.

3 Empirical Specification

The empirical specification is based primarily on the equilibrium wage function (34). 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. The equilibrium wage equation (34) is linear in person- and firm- specific components, which suggests adopting a similar approach. I depart from this 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 noted in Section 2.3, the empirical person and firm effects θ_i and ψ_j are random variables. Second, it is the distribution of the person and firm effects that is of primary interest. Their realization for specific workers and firms is of secondary importance. A mixed model specification estimates parameters of the distribution of the random effects. Given these, one can compute Best Linear Unbiased Predictors (BLUPs) of the realized random effects. Third, the data can be considered a random sample of workers and firms from a larger population. Thus the person and firm effects are also 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 θ_i and ψ_j.

The second important departure from earlier empirical work is to explicitly account for the structure implied by the learning process on earnings residuals. There are several aspects of this structure to accommodate. Recall the equilibrium wage function (34), into which θ_i,

---

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. Note it is not "productive" human capital: 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.

ψ_j, and $\delta m_{ij\tau}$ enter linearly. Proposition 8 established that $E(m_{ij\tau}\theta_i) = E\left(m_{ij\tau}\psi_j\right) = 0$ for all i, j, and $\tau > 0$. Furthermore, $m_{ij\tau}$ is a normally distributed random variable with zero expected value in the population, and satisfying

$$
\begin{aligned}
E\left[m_{ij\tau}m_{i'j'\tau'}\right] &= 0 \text{ for } i \neq i' \text{ or } j \neq j' &(41)\\
E\left[m_{ij\tau}m_{ij\tau'}\right] &= V_\tau \text{ for } \tau \leq \tau' &(42)
\end{aligned}
$$

where V_τ is the unconditional variance of $m_{ij\tau}$ defined in equation (13).[15] Taken together, these facts about $m_{ij\tau}$ imply that it is appropriate to treat the term $\delta m_{ij\tau}$ in (34) as a normally distributed statistical residual with a non-zero covariance within a worker-firm match. This imposes considerable structure on the residuals. First, from equation (41) residuals are uncorrelated across worker-firm matches. Second, residuals within a worker-firm match have the martingale covariance structure given by equation (42). Farber and Gibbons (1996) obtain a similar result. Further details on the martingale covariance structure are in Section 3.3. Finally, as discussed at the end of the previous Section, worker-firm matches terminate when $m_{ij\tau}$ falls below the match-specific threshold $\bar{m}_{ij}$, implying that the observed distribution of earnings residuals (and hence earnings) is truncated. The specification I now develop accounts for all these aspects of the learning process.

An empirical specification for earnings based on (34) takes the form

$$w_{ijt} = \mu + x_{it}'\beta + \theta_i + \psi_j + \varepsilon_{ijt} \tag{43}$$

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;[16] β is a vector of returns to time-varying characteristics; θ_i is the pure person effect; ψ_j is the pure firm effect for the firm j at which worker i was employed in t (denoted $j = J(i,t)$); and ε_{ijt} is a statistical residual. As in Section 2, i indexes people and j indexes firms. Note that in Section 2, τ was used to index tenure; here t indexes calendar time.[17] We can further decompose the pure person effect θ_i into components observed and unobserved by the econometrician as

$$\theta_i = \alpha_i + u_i'\eta \tag{44}$$

[15] The fact that $Cov(m_{ij\tau}, m_{ij\tau'}) = Var(m_{ij\tau})$ for $\tau \leq \tau'$ is a standard property of Bayesian learning. The only information common to $m_{ij\tau}$ and $m_{ij\tau'}$ 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 3.3.

[16] The inclusion of a set of time-varying individual covariates represents a slight deviation from the theoretical model. In this application, covariates x_{it} include time effects, labor market experience, and controls for attachment to the labor force. These are well known determinants of earnings. Extending the theoretical model to include aggregate shocks (time effects) and time-varying individual productivity (experience) is left for future research.

[17] The inclusion of time-varying covariates x_{it} in (43) 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.

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 measures returns to time-invariant characteristics.[18].

Let N^* denote the total number of observations; N the number of workers; J the number of firms; q the number of time-varying covariates; and p the number of time-invariant person characteristics. Rewriting (43) and (44) in matrix notation, we have

$$w = X\beta + U\eta + D\alpha + F\psi + \varepsilon \tag{45}$$

where w is the $N^* \times 1$ vector of earnings outcomes, X is the $N^* \times q$ matrix of time-varying covariates (including the constant term); β is the $q \times 1$ vector of returns to time-varying covariates; U is the $N^* \times p$ matrix of time-invariant person characteristics; η is the $p \times 1$ vector of returns to time-invariant person characteristics; 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 firm effects; and ε is the $N^* \times 1$ vector of residuals.

Abowd and Kramarz (1999) discuss fixed and mixed model specifications based on equations like (45). 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 specifications.

3.1 The Fixed Model

The fixed model is completely specified by (45) and the following assumptions on ε_{ijt} :

$$E\left[\varepsilon_{ijt}|i,j,t,x,u\right] = 0 \tag{46}$$

$$E\left[\varepsilon\varepsilon'\right] = \sigma_\varepsilon^2 I_{N^*} \tag{47}$$

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

$$\begin{bmatrix} X'X & X'U & X'D & X'F \\ U'X & U'U & U'D & U'F \\ D'X & D'U & D'D & D'F \\ F'X & F'U & F'D & F'F \end{bmatrix} \begin{bmatrix} \beta \\ \eta \\ \alpha \\ \psi \end{bmatrix} = \begin{bmatrix} X'w \\ U'w \\ D'w \\ F'w \end{bmatrix}. \tag{48}$$

In the data described in Section 4, the cross product matrix on the left hand side of (48) is of sufficiently high dimension to preclude estimation using standard software packages.

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

Instead, I obtain the full least squares solutions $\hat{\beta}$, $\hat{\eta}$, $\hat{\alpha}$, and $\hat{\psi}$ using the least squares conjugate gradient algorithm of Abowd et al. (2002). This algorithm exploits the sparse structure of the cross products matrix after blocking on connected groups of workers and firms.[19] 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 α and ψ in detail. I apply their procedure to obtain unique estimates of α and ψ subject to the restriction that the overall and group means of the person and firm effects 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 (45). The particular specification that I consider treats β and η as fixed, and α and ψ as random. The model is completely specified by (45) and the assumption

$$\begin{bmatrix} \alpha \\ \psi \\ \varepsilon \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 I_N & 0 & 0 \\ 0 & \sigma_\psi^2 I_J & 0 \\ 0 & 0 & R \end{bmatrix} \right). \tag{49}$$

It is worth noting that unlike the usual random effects specification considered in the econometric literature, (45) and (49) *do not* assume that the random person and firm effects are orthogonal to the design of the fixed effects (X and U). Such an assumption is almost always violated in economic data. More general specifications than (49) 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.

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 in the data. Suppose the data are arranged in lexicon order, that is by t within j within i. In the balanced data case, where there are $\bar{\tau}$ observations on each worker-firm match, the various specifications I consider have a residual covariance that can be written

$$R = I_M \otimes W \tag{50}$$

where W is a $\bar{\tau} \times \bar{\tau}$ positive-semidefinite matrix of within-match residual covariances. I discuss the several forms of W that I estimate in Sections 3.3 and 3.5. The extension to

[19] See Searle (1987) for a general 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.

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.[20] 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 general form of the residual covariance is

$$R = I_M \otimes S_{ij}' W S_{ij}. \tag{51}$$

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 the residual covariance R : ANOVA, Maximum Likelihood (ML), and Restricted Maximum Likelihood (REML). ANOVA and ML methods are familiar to most economists; REML less so.[21] 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. 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 turn out to be equivalent to residuals obtained after fitting the fixed effects via OLS. The linear combinations $k'w$ are chosen so that

$$k' \left(\begin{bmatrix} X & U \end{bmatrix} \begin{bmatrix} \beta \\ \eta \end{bmatrix} \right) = 0 \;\; \forall \beta, \eta \tag{52}$$

which implies

$$k' \left(\begin{bmatrix} X & U \end{bmatrix} \right) = 0. \tag{53}$$

Thus k' projects onto the space orthogonal to $\begin{bmatrix} X & U \end{bmatrix}$, and must therefore be of the form

$$\begin{aligned} k' &= c' \left[I_{N^*} - \begin{bmatrix} X & U \end{bmatrix} \left(\begin{bmatrix} X' \\ U' \end{bmatrix} \begin{bmatrix} X & U \end{bmatrix} \right)^- \begin{bmatrix} X' \\ U' \end{bmatrix} \right] && (54) \\ &\equiv c' M_{XU} && (55) \end{aligned}$$

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

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

[21] 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.

for arbitrary c', and where A^{-} denotes the generalized inverse of A. When $\begin{bmatrix} X & U \end{bmatrix}$ has rank $r \leq q+p$, there are only $N^{*}-r$ linearly independent vectors k' satisfying (52).

Define $K' = TM_{XU}$ with rows k' satisfying (52), and where K' and T have full row rank $N^{*}-r$. REML estimation proceeds by performing maximum likelihood on $K'w$. For $w \sim N(X\beta + U\eta, \mathbf{V})$ it follows that

$$K'w \sim N(0, K'\mathbf{V}K) \tag{56}$$

where $\mathbf{V} = DD'\sigma_{\alpha}^{2} + FF'\sigma_{\psi}^{2} + R$ is the covariance of earnings implied by (49). The REML log-likelihood (i.e., the log-likelihood of $K'w$) is therefore

$$\log L_{REML} = -\frac{1}{2}(N^{*}-r)\log 2\pi - \frac{1}{2}\log|K'\mathbf{V}K| - \frac{1}{2}w'K(K'\mathbf{V}K)^{-1}K'w. \tag{57}$$

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'. Second, REML estimates are invariant to the value of the fixed effects. 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 (57). The AI algorithm is a variant of Fisher scoring. Unlike the method of scoring, which uses the expected information matrix to compute parameter updates, AI uses a computationally convenient average of the expected and observed information matrices in the update.[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 (57), 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 a random effect γ. 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 $0.5(1-\Pr(\chi_{1}^{2} \leq \Lambda))$.

[22]The usual statistical definition of balanced data can be found in Searle (1987). Longitudinal linked data on employers and employees is balanced if we observe each worker is 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 (57). The observed information matrix is the inverse of the negative Hessian. Newton's method uses the observed information matrix to compute parameter updates in maximizing the log-likelohood.

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. Henderson, in Henderson et al. (1959) derived a system of equations that simultaneously yield the BLUE of the fixed effects (β and η) and BLUP of the random effects. These equations have become known as the mixed model equations or Henderson equations. Define

$$G = \begin{bmatrix} \sigma_\alpha^2 I_N & 0 \\ 0 & \sigma_\psi^2 I_J \end{bmatrix}. \tag{58}$$

The mixed model equations are

$$\begin{bmatrix} \begin{bmatrix} X' \\ U' \end{bmatrix} R^{-1} \begin{bmatrix} X & U \end{bmatrix} & \begin{bmatrix} X' \\ U' \end{bmatrix} R^{-1} \begin{bmatrix} D & F \end{bmatrix} \\ \begin{bmatrix} D' \\ F' \end{bmatrix} R^{-1} \begin{bmatrix} X & U \end{bmatrix} & \begin{bmatrix} D' \\ F' \end{bmatrix} R^{-1} \begin{bmatrix} D & F \end{bmatrix} + G^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\beta} \\ \tilde{\eta} \\ \tilde{\alpha} \\ \tilde{\psi} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} X' \\ U' \end{bmatrix} R^{-1} w \\ \begin{bmatrix} D' \\ F' \end{bmatrix} R^{-1} w \end{bmatrix} \tag{59}$$

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 (59) requires estimates of R and G. I apply common practice, and use the REML estimates $\tilde{G}$ and $\tilde{R}$ for this purpose.

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 (59) converge to the normal equations (48). Thus the mixed model estimates $(\tilde{\beta}, \tilde{\eta}, \tilde{\alpha}, \tilde{\psi})$ converge to the least squares solutions $(\hat{\beta}, \hat{\eta}, \hat{\alpha}, \hat{\psi})$.

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

$$E\left(\begin{bmatrix} \tilde{\alpha} \\ \tilde{\psi} \end{bmatrix} - \begin{bmatrix} \alpha \\ \psi \end{bmatrix}\right)' A \left(\begin{bmatrix} \tilde{\alpha} \\ \tilde{\psi} \end{bmatrix} - \begin{bmatrix} \alpha \\ \psi \end{bmatrix}\right) \tag{60}$$

where A is any positive definite symmetric matrix. They are *linear* in w, and *unbiased* in the sense $E(\tilde{\alpha}) = E(\alpha)$ and $E(\tilde{\psi}) = E(\psi)$.

3.3 The Martingale Hypothesis

Having discussed fixed and mixed model estimation in some detail, I now turn to implications of the matching model for earnings residuals. In the matching model, agents update their beliefs about match quality using Bayes' Rule. As a consequence, the posterior mean of beliefs about match quality $m_{ij\tau}$ is a martingale. This 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 m_{ij}, $m'_{ij} =$

$[m_{ij1}\ \cdots\ m_{ij\bar{\tau}}]$.[25] Overlaying V with classical measurement error as in Farber and Gibbons (1996), the martingale structure implies

$$V = \begin{bmatrix} V_1 + \sigma_u^2 & V_1 & V_1 & \cdots & V_1 \\ V_1 & V_2 + \sigma_u^2 & V_2 & \cdots & V_2 \\ V_1 & V_2 & V_3 + \sigma_u^2 & \cdots & V_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ V_1 & V_2 & V_3 & \cdots & V_{\bar{\tau}} + \sigma_u^2 \end{bmatrix} \tag{61}$$

where σ_u^2 is the variance of measurement error, and where the terms V_τ for $\tau = 1, ..., \bar{\tau}$ are the unconditional variance of $m_{ij\tau}$ in (13).

Some aspects of V are worthy of note. First, concentrating on the lower triangle of (61), off-diagonal elements within each column are equal. The reason for this 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 just the variance of the signals received up to tenure τ, which is $Var\left(m_{ij\tau}\right) \equiv V_\tau$. Second, diagonal elements of V differ from off-diagonal elements in the same column by the variance of measurement error. Finally, it can be shown that $V_{\tau+1} > V_\tau$. This is because as tenure increases, so does the number of observed signals. Each signal enters the posterior mean of beliefs (recall (9)), so its variance increases with the number of signals.

The structural parameters σ_c^2, σ_z^2, and σ_e^2 enter into each V_τ. Thus they can be recovered (up to a factor of proportionality, δ^2) from an estimate of the within-match residual covariance. I test the martingale hypothesis and recover the structural parameters σ_c^2, σ_z^2, σ_e^2, and σ_u^2 using a two-step procedure. The first step is to obtain an estimate of the within-match residual covariance W. Under the fixed model specification, the estimate $\hat{W}$ and its covariance are computed directly from the estimated residuals (see Abowd and Card (1989) for methods and formulae).[26] Under the mixed model specification, the estimate $\tilde{W}$ is obtained by estimating the earnings equation (43) with an unstructured within-match residual covariance. That is, the only restriction placed on W during mixed model estimation is symmetry and positive semi-definiteness. An estimate of the covariance 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), I then fit the martingale covariance $\delta^2 V$ to the estimate of W by equally weighted minimum distance (EWMD).[27] This yields estimates of the structural

[25] In the unbalanced data case where a match between worker i and firm j lasts $\tau_{ij} \leq \bar{\tau}$ periods, $m'_{ij} = \left[m_{ij1}\ \cdots\ m_{ij\tau_{ij}}\right]$. The residual covariance is $R = I_M \otimes \delta^2 S'_{ij} V S_{ij}$. The selection matrix S_{ij} was defined earlier. It selects those rows and columns of V corresponding to observed earnings outcomes on the match between worker i and firm j.

[26] Estimates obtained on $\hat{W}$ will be available in a future version of this paper.

[27] Optimal minimum distance estimation, as discussed in Hansen (1982) and Chamberlain (1984), proved

parameters up to a factor of proportionality: the square of the bargaining strength parameter δ. I test the martingale hypothesis with the usual χ^2 test of overidentifying restrictions, using the test statistic of Newey (1985).[28]

3.4 Accommodating Residual Truncation

Under the matching model, a match between worker i and firm j terminates when $m_{ij\tau}$ falls below the match-specific threshold $\bar{m}_{ij}$ defined in (38). This implies the distribution of earnings residuals is truncated. Only earnings observations such that $m_{ij\tau} \geq \bar{m}_{ij}$ are observed. Rewrite the definition of $\bar{m}_{ij}$ in (38) as $\bar{m}_{ij} = -\mu - \zeta_i - \nu_j$ where $\zeta_i = a_i - (1-\beta) U_i$ and $\nu_j = b_j - (1-\beta) V_j$. Conditional on the match between worker i and firm j surviving to tenure τ, the marginal probability of observing the earnings outcome $w_{ij\tau}$ is

$$\begin{aligned} \Pr(m_{ij\tau} \geq \bar{m}_{ij}) &= 1 - \Phi\left(\frac{-\mu - \zeta_i - \nu_j}{V_\tau^{1/2}}\right) \\ &= \Phi\left(\frac{\mu + \zeta_i + \nu_j}{V_\tau^{1/2}}\right) \end{aligned} \tag{62}$$

where Φ is the standard normal CDF, and the latter equality follows from symmetry of the normal distribution about its mean. Equation (62) suggests a simple means of correcting the distribution of earnings residuals to account for truncation. Specifically,

$$\begin{aligned} E\left[w_{ijt} | m_{ij\tau} \geq \bar{m}_{ij}\right] &= \mu + x_{it}'\beta + \theta_i + \psi_j + V_\tau^{1/2} \frac{\phi\left(\frac{\mu+\zeta_i+\nu_j}{V_\tau^{1/2}}\right)}{\Phi\left(\frac{\mu+\zeta_i+\nu_j}{V_\tau^{1/2}}\right)} \\ &= \mu + x_{it}'\beta + \theta_i + \psi_j + V_\tau^{1/2}\lambda_{ij\tau} \end{aligned} \tag{63}$$

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

A simple but inefficient truncation correction can be based solely on the marginal distribution of $m_{ij\tau}$: estimate a sequence of $\bar{\tau}$ continuation probits with (fixed or random) person and firm effects. Using results from these, construct an estimate of $\lambda_{ij\tau}$ and include it as an additional regressor in the earnings equation (43). One problem with this approach is that the estimated person and firm effects vary across the probit equations. A simple alternative is to pool the probit equations, but in this case the variance term V_τ is restricted to be the same at each tenure. Of course the efficient solution is to estimate instead a $\bar{\tau}$-variate probit with (fixed or random) person and firm effects with the full covariance structure V given in (61), ignoring the measurement error terms. However, multivariate probits are notoriously

infeasible. Under both fixed and mixed model specifications, the covariance of W was poorly conditioned, and did not invert.

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

difficult to estimate, especially when the dimension $\bar{\tau}$ is large; even more so with the large number of effects to be estimated. I opt for a compromise where the person and firm effects are constrained to be equal across probit equations, and allow limited variation across tenure in the variance terms V_τ.

For a match between worker i and firm j, define the $\bar{\tau} \times 1$ vector d_{ij}. The τ^{th} element of d_{ij} is 1 if the worker is employed at firm j at tenure τ and 0 otherwise. I treat the person- and firm-specific mobility effects ζ_i and ν_j from equation (62) as random, and estimate a multivariate probit on d_{ij} using Average Information REML applied to the method of Schall (1991).[29] The estimating equation is based on

$$\begin{aligned} \Pr\left(d_{ij1}=1, d_{ij2}=1, ..., d_{ij\bar{\tau}}=1\right) &= \Phi_{\bar{\tau}}\left(\mu+\zeta_i+\nu_j, \bar{V}\right) && (64) \\ \zeta_i &\sim N\left(0, \sigma_\zeta^2\right) && (65) \\ \nu_j &\sim N\left(0, \sigma_\nu^2\right) && (66) \end{aligned}$$

where $\Phi_{\bar{\tau}}$ denotes the $\bar{\tau}$-variate normal CDF, and $\bar{V}$ is a restricted form of the covariance V of the vector of belief terms m_{ij}:

$$\bar{V} = \begin{bmatrix} \bar{V}_1 & 0 \\ 0 & \bar{V}_2 \end{bmatrix} \otimes I_{\bar{\tau}/2}. \quad (67)$$

$\bar{V}_1$ and $\bar{V}_2$ are scalars. The restricted covariance $\bar{V}$ has the following properties. Since the off-diagonal elements are zero, the specification is based on the marginal distribution of the belief terms $m_{ij\tau}$, rather than their joint distribution. All cross-equation correlation is built in via the random effects ζ_i and ν_j. Additionally, the diagonal elements of $\bar{V}$ are restricted to be equal in each period of the first half of the observed tenure distribution, and equal in each period of the second half of the observed tenure distribution.[30] It is important to note that these restrictions are only imposed on the probit covariance. I subsequently recover the unrestricted matrix V from earnings residuals, as described in Section 3.3.

With estimates of $\bar{V}_1$, $\bar{V}_2$, and the realized random effects $\tilde{\zeta}_i$ and $\tilde{\nu}_j$ in hand, I construct an estimate $\tilde{\lambda}_{ij\tau}$ of the Inverse Mills' Ratio term (63) for each worker-firm match at each observed tenure. The estimate $\tilde{\lambda}_{ij\tau}$ is included as an additional time-varying covariate in the earnings equation (43).

3.5 Alternate Covariance Structures

In addition to the martingale covariance structure predicted by the matching model, I test a variety of more parsimonious forms for W in the mixed model. The simplest of these is

[29] 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 $\Phi_{\bar{\tau}}$. The process requires an iterative reweighting of the design matrices of fixed and random effects in the linearized system, see Schall (1991) for details.

[30] These can be interpreted as restrictions on the learning process.

$W = \sigma_\varepsilon^2 I_{\bar{T}}$, i.e., a homoskedastic error with no within-match covariance between earnings residuals. Additionally, I estimate and test several models with one and two parameter residual covariances: an AR(1) residual, AR(2), MA(1), MA(2), and ARMA(1,1). In the context of the matching model, these structures can be interpreted as a simplified learning process, whereby beliefs about match quality evolve as an AR(1), AR(2), etc.

For completeness, I also test for the presence of a random match effect γ_{ij} in the earnings equation. The match effect is the interaction between person and firm effects. Conditional on observables x_{it} and u_i, a random match effect implies earnings are equicorrelated within-match. Like the random person and firm effects, the match effect is not assumed orthogonal to x_{it} and u_i.

3.6 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 rather than levels, partly because earnings are typically found to be approximately lognormal, but also to alleviate heteroskedasticity. As a compromise, I opt to model earnings in both logs and levels. I model earnings in levels to test the martingale hypothesis and other predictions of the matching model, and model earnings in logs to facilitate comparison with earlier work. The log-linear specification can be interpreted as an approximation to the earnings equation in levels as follows. Rewrite (43) 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

$$\begin{aligned} \ln w_{ijt} &\approx \ln \mu + x'_{it}\frac{\beta}{\mu} + \frac{\theta_i}{\mu} + \frac{\psi_j}{\mu} + \frac{\varepsilon_{ijt}}{\mu} \\ &= \mu^* + x'_{it}\beta^* + \theta_i^* + \psi_j^* + \varepsilon_{ijt}^* \end{aligned} \tag{68}$$

where the first line of (68) uses the first-order Taylor series approximation around $x = 0$, $\ln(1 + x) \approx x$, and where $\mu^* = \ln \mu$, $\beta^* = \beta/\mu$, $\theta_i^* = \theta/\mu$, $\psi_j^* = \psi/\mu$, and $\varepsilon_{ijt}^* = \varepsilon_{ijt}/\mu$.

4 Data

Estimating the models described in the previous section requires longitudinal linked data on employers and employees: data with repeated observations on both workers and firms. Data used in this application are from the Longitudinal Employer-Household Dynamics (LEHD) program database, under development at the U.S. Census Bureau. The LEHD database includes data from eight states: California, Florida, Illinois, Maryland, Minnesota, North

Carolina, Pennsylvania, and Texas. Together, these states represent about 50 percent of U.S. employment. In this paper, I use data from two of the eight 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. In particular, the universe of firms subject to UI system reporting varies slightly across states. These and other characteristics of UI wage records are detailed elsewhere, for example Stevens (2002) and Burgess et al. (2000). With the UI wage records as its frame, the LEHD data comprise the universe of employers required to file UI system wage reports — that is, all employment covered by the UI system in the eight participating states. The data span the first quarter 1990 through the fourth quarter 1999.[31]

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. In the LEHD database, additional demographic characteristics are integrated with these data from various internal Census Bureau sources. Such characteristics include sex, race, and date of birth.[32]

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.

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 earnings (as reported in the UI system wage records) were largest in t.[33] About 87 percent of the UI system wage records correspond to employment at a dominant employer. The second concept is *full quarter employment.* In quarter q, individual i is identified as having worked a full quarter at SEIN $j = J(i, q)$ if there are UI wage records for which $J(i, q-1) = J(i, q) = J(i, q+1)$. That is, if individual i was employed at the SEIN in both the previous and subsequent quarter.

4.1 Sample Construction

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

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

[32] Based on an exact match to administrative data sources.

[33] The dominant employer is identified on the basis of actual reported earnings, not the annualized earnings measure discussed in Section 4.2.

sample period,[34] 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 prior to estimation, reducing the analysis sample to 49.3 million annual earnings records.

Using the method of Abowd et al. (2002), estimation of the fixed model is possible on the entire analysis sample. Unfortunately, estimating the mixed model on the full sample remains computationally infeasible. Drawing an appropriate random sample of observations for estimating the mixed model is not a simple task. 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 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. I demonstrate in Appendix C 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 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 the dense random sampling algorithm of Appendix C. Each worker is connected to at least $n = 5$ others.[35] 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 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 groups. The simple random sample is much less connected. About 80 percent of jobs are contained in the two largest groups, and 84 percent in groups containing at least 5 worker-firm matches. This is in contrast to the full analysis sample and dense samples, in which all jobs are in groups with at least 5 worker-firm matches by construction. In the simple random sample, fully 5.5 percent of jobs are connected to no other.

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

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

4.2 Variable Creation and Missing Data Imputation

Time-varying covariates X in all earnings regressions 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).[36]

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. In keeping with the statistical literature on multiple imputation, I refer to these as completed data implicates.

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 was 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 real average earnings in the full quarters, the individual was presumed to have worked a full quarter.[37] That is, reported earnings were treated as full-quarter earnings (12.5 percent of the analysis sample). If on the other hand reported earnings were less than 80 percent of real average average earnings in the full quarters, earnings were imputed to the full-quarter level (7.5 percent

[36] As described in Section 4.2.3, the initial potential experience measure was set to zero in this case.

[37] 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.

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 were imputed to the full-quarter level, three imputed values were drawn from the posterior predictive distribution under a diffuse prior. After constructing the real full-quarter earnings measure, the quarterly measures were 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 reported 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 were drawn from the normal approximation (at the mode) to the posterior predictive distribution under a diffuse prior. Prior to estimation of the earnings model, education was collapsed 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 was negative, potential experience was set to zero. In each subsequent quarter, labor market experience is accumulated using the individual's realized labor market history. Note that since initial experience depends on the multiply-imputed education measure, calculated labor market experience varies across the three completed data implicates.

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 1990 quarter 1. For this state, all active jobs in this quarter were presumed left-censored. In the second state, the data series begins in 1985 quarter 1. All jobs in that state that were active in 1985 quarter 1 were presumed left-censored. Under this definition, left censored spells comprise 33 percent of jobs in the full analysis sample.

For spells that are not left-censored, tenure was set to 1 in the first quarter for which there was a UI system wage record, and subsequently accumulated using the individual's employment history at that employer. For left-censored spells, tenure as of the first quarter 1990 was 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 reported 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 was accumulated using the individual's employment history at that employer.

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 on 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 analytical result in Appendix C that they are equivalent. 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 higher labor force attachment than the average individual in the full analysis sample: individuals in the dense and simple random 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. These differences are all very slight, however.

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.[38] As a consequence, the dense and simple random samples are

[38] Recall these are random samples of individuals employed in 1997. There are no individuals in the dense

indistinguishable, but their employment series differ somewhat from that of the full analysis sample.

5 Results

The econometric analysis described in Section 3 is performed 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 very little variation in the estimated covariate effects across specifications. All estimated covariate effects are quite 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 of primary interest, estimated on the natural logarithm of annualized earnings. The four models are: 1) the fixed model; 2) the mixed model with random person and firm effects and a spherical error ($R = \sigma_\varepsilon^2 I_{N*}$); 3) the mixed model with random person, firm, and match effects and a spherical error; and 4) the 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 calculated 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, employment at a firm with an estimated firm effect one standard deviation above the mean 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.[40] This is perhaps not surprising, given that selection bias is frequently found to be quite small in earnings data. The second item of note is that in each of the models, the variance component associated with the

and simple random samples who were not employed in that year.

[39]I have not applied the truncation correction to the fixed model. Doing so would require computing realized random effects $\tilde{\zeta}_i$ and $\tilde{\nu}_j$ for each worker and firm in the full analysis sample (about 10 million effects total). Although this is technically feasible, it is extraordinarily demanding from a computational standpoint. I leave this exercise for future research.

[40]Standard errors of the estimated variance components are very small. Thus, some of the differences between parameter estimates with and without the truncation correction are statistically significant at conventional levels. However, it is difficult to argue that the differences are economically significant, given that they are all in the 3rd decimal place.

person effect (σ^2_α) is considerably larger than the variance component associated with the firm effect (σ^2_ψ). That is, with respect to earnings, individuals are more heterogeneous than firms. This is consistent with Abowd et al. (1999) and others, who find unobserved individual heterogeneity to be a much more important determinant of earnings 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 also yields a fairly large estimate of σ^2_α (0.23), though the estimate of σ^2_ψ is 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 estimate 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 employment at a firm whose ψ_j is one standard deviation above the mean increases earnings by 0.28 log points.

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 (59) on Dense Sample 2, using the estimated 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 (parameters are re-scaled accordingly). Having solved the mixed model equations, I compute the variance of the estimated residuals, reported in Table 3 as the variance of out-of-sample predication error. By this measure, the mixed model specification with a match effect has the smallest out-of-sample prediction error; the mixed model with the unstructured within-match residual covariance has the largest. This is perhaps not surprising, since the latter model attributes more earnings variation to the residual than other specifications.

I test for the presence of a match effect in the mixed model with a spherical error. The p-value for this test is extremely small ($< 10^{-12}$), 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 are not nested hypotheses. The AIC and BIC statistics indicate the model with W unrestricted fits the data better than the match effect specification.

Table 4 reproduces the information 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 very little influence on the estimated variance components; the estimates of σ^2_α are (much) 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 person and firm 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 economically significant.

As in Table 3, the mixed model with unstructured W obtains the best fit to earnings levels using in-sample measures, and the worst fit using out-of-sample measures. Once again, the mixed model with a match effect and spherical error has the smallest out-of-sample prediction error. However, on the unit variance scale, the prediction error is nearly twice that obtained in the log specification. It is no surprise that the linear model fits the logarithm of earnings better than it does earnings levels.

Tables 5 and 6 present information comparable to that in Tables 3 and 4 under a variety of alternate within-match residual covariance structures. The estimated variance components are comparable to those in Tables 3 and 4. Not surprisingly, none of these more parsimonious models fit as well as the specification with W unrestricted. On earnings logs, the ARMA(1,1) specification obtains the next best fit using in-sample measures. This suggests shocks to log earnings contain both permanent and transitory components, which is at odds with the learning model of Section 2. I perform REMLRTs of the ARMA(1,1) specification against the ARMA(1,0) and ARMA(0,1) null hypotheses. In each case, I reject the null at the 1 percent level. The AR(2) specification gives the best fit to earnings levels. Unlike the case of log earnings, this suggests that earnings shocks are permanent, and is consistent with the learning model. On the basis of a REMLRT, I reject the null of an AR(1) residual versus the AR(2) hypothesis.

Tables 7-10 present correlations among the estimated effects for the specifications presented in Tables 3 and 4. Similar correlation tables for models with the alternate residual covariance structures presented in Tables 5 and 6 are available from the author on request. Table 7 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 predicted 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 a 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 8 presents correlations among estimated effects for the mixed model specifications after correcting for truncation. Once again, correcting for truncation has little effect on the results. The truncation correction term $\beta_\lambda \lambda_{ij}$ is positively correlated with earnings (0.235 in all specifications), exhibits a small positive correlation with θ_i (about 0.1), and a small negative correlation with ψ_j (between -0.004 and -0.06 depending on specification).

Tables 9 and 10 reproduce the information in Tables 7 and 8 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 quite strong: between 0.17 and 0.27 depending on the specification.

5.1.2 The Continuation Probit

Table 11 presents parameter estimates in the probit model used to correct for truncation. The longest observed tenure in Dense Sample 1 was 21 years.[41] Thus the dependent vector d_{ij} is a (21×1) vector of employment outcomes, where the τ^{th} entry takes value 1 if the individual i was employed at firm j at tenure τ, and zero otherwise.[42] The residual variance term $\bar{V}_1$ corresponds to the first 10 years of job tenure; $\bar{V}_2$ corresponds to job tenure between 11 and 21 years. The estimates of $\bar{V}_1$ and $\bar{V}_2$ are reported on the scale of the data d_{ij}. Though $\bar{V}_1$ and $\bar{V}_2$ are rather crude estimates of terms in the martingale covariance V, they have the property $\bar{V}_2 > \bar{V}_1$ predicted by the matching model.

The person effect ζ_i was not identified for 28 percent of individuals in Dense Sample 1. This situation arose when an individual only held one job in the sample period, and the job spell was right-censored. In these cases, the person effect was restricted to equal the unconditional mean of ζ_i (zero).

Unlike the models estimated on earnings data, the variance component σ^2_ν associated with the firm effect in the probit equation is much larger than the variance component σ^2_ζ associated with the person effect. This suggests that firms exhibit greater heterogeneity in their separation policies than workers do.

---

[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 on left-censored job spells.

[42] Entries in d_{ij} that corresponded to years outside the sample period 1990-1999 were set to missing. The estimation algorithm treats the missing observations as missing at random. See Rubin (1976) for a definition.

5.2 Testing the Martingale Hypothesis

I perform a formal test of the matching model's predictions for the within-match residual covariance on the mixed model estimate of the unstructured residual covariance W. Estimates of W are presented in Tables 12-15. Table 12 presents the estimate of W obtained on log earnings without correcting for truncation. Estimates in Table 12 exhibit a number of the properties of the martingale covariance structure overlaid with classical measurement error given in (61): 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 covariances, the elements in Table 12 consistently decline. This is consistent with the ARMA(1,1) specification that fit these data well.

Table 13 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 12 and 13 suggests that there is less within-column decay in the autocovariances corrected for truncation than without the correction.

In Tables 14 and 15 I present estimates of W obtained on earnings levels, with and without correcting for truncation. They are virtually identical. Once again, on a casual level 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. This is particularly true for the first 10 years of tenure.

I formally test the martingale covariance structure predicted by the matching model by fitting (13) and (61) to estimates of W by equally weighted minimum distance (EWMD). Table 16 presents the resulting estimates of structural parameters and p-values from the chi-squared test of over-identifying restrictions.[43] Recall that the structural parameters σ_u^2, σ_c^2, σ_z^2, and σ_e^2 are only identified up to a factor of proportionality: the square of the bargaining

[43] 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, and let s_d^2 denote the simple variance of the statistics d_m. Define

$$\hat{r}_m = \frac{\left(1 + M^{-1}\right) s_d^2}{2\bar{d}_m + \left(4\bar{d}_m^2 - 2ks_d^2\right)^{1/2}} \tag{69}$$

and

$$\hat{v} = (M-1)\left(1 + \hat{r}_m^{-1}\right)^2. \tag{70}$$

The quantity $\hat{r}_m$ is a method of moments estimator of the relative increase in variance of the test statistic strength parameter δ. The estimates in Table 16 are presented on the scale of the data, i.e., for $\delta = 1$. They can be re-scaled to other values of $0 < \bar{\delta} < 1$ quite easily: the re-scaled parameter $\bar{\sigma}^2$ is $\bar{\sigma}^2 = \sigma^2/\bar{\delta}^2$.

The first column of Table 16 presents the results obtained on log earnings. Given that estimates of W obtained with and without the truncation correction were virtually identical, it is no surprise that the minimum distance estimates are almost identical 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 about the same magnitude as 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 0.01, 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_{ij}, 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 below 0.01. However, these results should be viewed cautiously, since the p-value of the χ^2 test of overidentifying restrictions is 0.0014. Though the log specification is only an approximation to the matching model, this is nevertheless quite strong evidence against the martingale prediction of the matching model.

Column (2) of Table 16 presents results estimated on earnings levels. The estimated variance of match quality is quite large: about 0.36, which is more than a third of the variance of annual earnings. The estimate of σ_z^2 (6.77) indicates the initial signal of match quality conveys little 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 (0.0012), we are inclined to reject the martingale hypothesis. However, recall from Tables 14 and 15 that for the first 10 years of tenure, the casual inspection indicated $\tilde{W}$ was highly consistent with the martingale structure. In column (3) of Table 16 I present results estimated on only the first 10 rows and columns of $\tilde{W}$. Unlike estimates obtained using the full 21 years of tenure data, these indicate that the variance of match quality is quite large (0.733) and that learning about match quality is quite slow (see Figure 4). Without the truncation correction, the p-value of the χ^2 test is 0.1447. Thus we cannot reject the martingale hypothesis implied by the matching model at even the 10 percent level. After correcting for truncation, the p-value falls to 0.0764 and we cannot reject the martingale hypothesis at the 5 percent level. Though far from conclusive evidence in favor of the matching model, these results certainly provide some support to the martingale hypothesis. However, they also raise the possibility that imputing initial tenure for left-censored job

due to nonresponse. The test statistic

$$\hat{D}_m = \frac{\frac{\bar{d}_m}{k} - \frac{M-1}{M+1}\hat{r}_m}{1 + \hat{r}_m} \tag{71}$$

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

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 the imputation.[44] I will address this issue more fully in a subsequent version of this paper by re-estimating W using only jobs that are not left-censored.

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; we observed a positive correlation between the estimated person and firm effects; and limited evidence in favor of the martingale 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 higher values of θ_i and ψ_j should on average be associated with longer job duration, and that larger firms should have larger estimated firm effects.

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. Figures 5 and 6 present the fitted curves. As the matching model predicted, higher 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 the 1997 employment to the estimated firm effects. Again, I focus on firm effects estimated in the mixed model with the unstructured within-match residual covariance. Results obtained on the other specifications are qualitatively the same. Figure 7 presents the fitted curve. 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. The equilibrium wage is linear in a person-specific component, a firm-specific component, and the posterior mean of beliefs about match quality. I showed that the matching model has considerable predictions for empirical person and firm effects and

[44]See footnote 41.

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. The most stringent test of the matching model, a test of the martingale 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 good workers and good firms last longer than other matches. Agents need to enter into a match to learn one another's type. By sampling a number of employment relationships, 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 extend the matching model to include aggregate productivity shocks. Another is to consider the case where a worker's productivity varies over time. Together, these amount to including a set of time-varying covariates in the theoretical model. On the empirical front, it would be of considerable value to refine the imputation of initial tenure for left-censored jobs. Relaxing the covariance structure in the probit equation would also be a worthy endeavor. An alternative would be to estimate a simultaneous duration model. This remains beyond the scope of current computational methods, however.

A Appendix: Omitted Derivations of Value Functions

A.1 The Value of Unemployment

I begin by deriving the value of unemployment. Let $J_0(a_i)$ denote the expected present value of wages of a worker of quality a_i who was unemployed last period and who is about to draw a match. In other words, $J_0(a_i)$ is the expected value of $J[w(\Omega_{ij1})]$ before the match is formed. That is, before the identity of the matching firm j is known, and before the signal x_{ij} is observed. Then

$$\begin{aligned}
J_0(a_i) &= E_0\left[1-G_1(\bar{m}_{ij})\right]w_{ij1} + U_i E_0 G_1(\bar{m}_{ij}) \\
&\quad +\beta E_0\left[1-G_1(\bar{m}_{ij})\right]\left[1-G_2(\bar{m}_{ij})\right]J\left[w(\Omega_{ij2})\right] + \beta U_i E_0\left[1-G_1(\bar{m}_{ij})\right]G_2(\bar{m}_{ij}) \\
&= \int\left[1-G_1(\bar{m}_{ij})\right]\int\int w_{ij1}\,dF_z dF_c dF_b + U_i\int G_1(\bar{m}_{ij})\,dF_b \\
&\quad +\beta\int\left[1-G_1(\bar{m}_{ij})\right]\left[1-G_2(\bar{m}_{ij})\right]\int\int J\left[w(\Omega_{ij2})\right]dF_z dF_c dF_b \\
&\quad +\beta U_i\int\left[1-G_1(\bar{m}_{ij})\right]G_2(\bar{m}_{ij})\,dF_b
\end{aligned} \tag{72}$$

which accounts for the possibility that the signal x_{ij} is too low for the match to persist. Let $\Omega^0_{ij\tau} = (a_i, b_j, 0, s^2_\tau)$. When w is linear in $m_{ij\tau}$ we can write

$$\begin{aligned}
E_0 w_{ij1} &= \int\int\int w_{ij1}\,dF_z dF_c dF_b \\
&= \int w\left(\Omega^0_{ij1}\right)dF_b.
\end{aligned} \tag{73}$$

When w is also independent of s^2_τ,

$$\begin{aligned}
E_0 w_{ij\tau} &= \int w\left(\Omega^0_{ij\tau}\right)dF_b \\
&= \int w\left(\Omega^0_{ij1}\right)dF_b \text{ for all } \tau > 0.
\end{aligned}$$

Consequently when Conjectures 1 and 2 are true,

$$\begin{aligned}
& E_0 \left[1 - G_1 \left(\bar{m}_{ij}\right)\right] \left[1 - G_2 \left(\bar{m}_{ij}\right)\right] J \left[w \left(\Omega_{ij2}\right)\right] \\
= & \int \left[1 - G_1 \left(\bar{m}_{ij}\right)\right] \left[1 - G_2 \left(\bar{m}_{ij}\right)\right] \int \int J \left[w \left(\Omega_{ij2}\right)\right] dF_z dF_c dF_b \\
= & \int \left[1 - G_1 \left(\bar{m}_{ij}\right)\right] \left[1 - G_2 \left(\bar{m}_{ij}\right)\right] \int \int \left[\begin{array}{c} w \left(\Omega_{ij2}\right) \\ +\beta \left[1 - G_3 \left(\bar{m}_{ij}\right)\right] E_0 J \left[w \left(\Omega_{ij3}\right)\right] \\ +\beta G_3 \left(\bar{m}_{ij}\right) U_i \end{array} \right] dF_z dF_c dF_b \\
= & \int \left[1 - G_1 \left(\bar{m}_{ij}\right)\right] \left[1 - G_2 \left(\bar{m}_{ij}\right)\right] w \left(\Omega_{ij1}^0\right) dF_b \\
& + \beta \int \left[1 - G_1 \left(\bar{m}_{ij}\right)\right] \left[1 - G_2 \left(\bar{m}_{ij}\right)\right] \left[1 - G_3 \left(\bar{m}_{ij}\right)\right] \int \int E_0 J \left[w \left(\Omega_{ij3}\right)\right] dF_z dF_c dF_b \\
& + \beta U_i \int \left[1 - G_1 \left(\bar{m}_{ij}\right)\right] \left[1 - G_2 \left(\bar{m}_{ij}\right)\right] G_3 \left(\bar{m}_{ij}\right) dF_b. \qquad (74)
\end{aligned}$$

Forward recursion on (72) gives:

$$\begin{aligned}
J_0 \left(a_i\right) = & \sum_{\tau=1}^{\infty} \beta^{\tau-1} \int w \left(\Omega_{ij1}^0\right) \prod_{s=1}^{\tau} \left[1 - G_s \left(\bar{m}_{ij}\right)\right] dF_b \\
& + U_i \sum_{\tau=1}^{\infty} \beta^{\tau-1} \int G_\tau \left(\bar{m}_{ij}\right) \prod_{s=1}^{\tau-1} \left[1 - G_s \left(\bar{m}_{ij}\right)\right] dF_b. \qquad (75)
\end{aligned}$$

Recall that U_i is the value of the worker's outside option – that is, the value of being unemployed today and behaving optimally thereafter. Thus $U_i = h_i + \beta \pi J_0 \left(a_i\right) + \beta \left(1 - \pi\right) U_i$ where $\pi = \frac{1}{u} m \left(u, v\right)$ is the worker's probability of drawing a match. Using (75),

$$U_i = \frac{h_i + \beta \pi \left(\sum_{\tau=1}^{\infty} \beta^{\tau-1} \int w \left(\Omega_{ij1}^0\right) \prod_{s=1}^{\tau} \left[1 - G_s \left(\bar{m}_{ij}\right)\right] dF_b\right)}{1 - \beta \left(1 - \pi\right) - \pi \sum_{\tau=1}^{\infty} \beta^{\tau} \int G_\tau \left(\bar{m}_{ij}\right) \prod_{s=1}^{\tau-1} \left[1 - G_s \left(\bar{m}_{ij}\right)\right] dF_b}. \qquad (76)$$

A.2 The Value of a Vacancy

I now turn to the value of a vacancy. Let $\Pi_0 \left(b_j\right)$ denote the expected present value of net revenues from a match for a firm of quality b_j before the identity of the matching worker i

is known. Then

$$\begin{aligned}
\Pi_0(b_j) &= E_0\left[1-G_1(\bar{m}_{ij})\right](q_{ij1}-w_{ij1})+E_0V_jG_1(\bar{m}_{ij}) \\
&\quad +\beta E_0\left[1-G_1(\bar{m}_{ij})\right]\left[1-G_2(\bar{m}_{ij})\right]\Pi\left[w(\Omega_{ij2})\right]+\beta V_jE_0\left[1-G_1(\bar{m}_{ij})\right]G_2(\bar{m}_{ij}) \\
&= \int\left[1-G_1(\bar{m}_{ij})\right]\int\int(q_{ij1}-w_{ij1})\,dF_zdF_cdF_a+V_j\int G_1(\bar{m}_{ij})\,dF_a \\
&\quad +\beta\int\left[1-G_1(\bar{m}_{ij})\right]\left[1-G_2(\bar{m}_{ij})\right]\int\int\Pi\left[w(\Omega_{ij2})\right]dF_zdF_cdF_a \\
&\quad +\beta V_j\int\left[1-G_1(\bar{m}_{ij})\right]G_2(\bar{m}_{ij})\,dF_a
\end{aligned}\tag{77}$$

which reflects the possibility that the signal x_{ij} is too low for the match to persist. Following Section A.1, we can write

$$\begin{aligned}
E_0w_{ij\tau} &= \int w\left(\Omega^0_{ij\tau}\right)dF_a \\
&= \int w\left(\Omega^0_{ij1}\right)dF_a \text{ for all } \tau>0.
\end{aligned}\tag{78}$$

and solve $\Pi_0(b_j)$ recursively to obtain

$$\begin{aligned}
\Pi_0(b_j) &= \sum_{\tau=1}^{\infty}\beta^{\tau-1}\int\left[\mu+a_i+b_j-w\left(\Omega^0_{ij1}\right)\right]\prod_{s=1}^{\tau}\left[1-G_s(\bar{m}_{ij})\right]dF_a \\
&\quad +V_j\sum_{\tau=1}^{\infty}\beta^{\tau-1}\int G_\tau(\bar{m}_{ij})\prod_{s=1}^{\tau-1}\left[1-G_s(\bar{m}_{ij})\right]dF_a.
\end{aligned}\tag{79}$$

Recall V_j is the value of firm j's outside option – that is, the value of a vacancy. Thus $V_j=-k_j+\beta\lambda\Pi_0(b_j)+\beta(1-\lambda)V_j$, where k_j is firm j's cost of maintaining a vacancy and $\lambda=\frac{1}{v}m(u,v)$ is the probability of a given vacancy being filled. Hence,

$$V_j=\frac{\beta\lambda\left(\sum_{\tau=1}^{\infty}\beta^{\tau-1}\int\left[\mu+a_i+b_j-w\left(\Omega^0_{ij1}\right)\right]\prod_{s=1}^{\tau}\left[1-G_s(\bar{m}_{ij})\right]dF_a\right)-k_j}{1-\beta(1-\lambda)-\lambda\sum_{\tau=1}^{\infty}\beta^{\tau}\int G_\tau(\bar{m}_{ij})\prod_{s=1}^{\tau-1}\left[1-G_s(\bar{m}_{ij})\right]dF_a}.\tag{80}$$

B Appendix: Omitted Proofs

Proof of Lemma 3. Note that each $G_p(\bar{m}_{ij})$ is a probability, so that $G_p(\bar{m}_{ij}) \in [0,1]$ for all $p > \tau$. Now notice that $A_{ij\tau}$ is strictly decreasing in each $G_p(\bar{m}_{ij})$. Thus the lower bound is given by

$$A_{ij\tau}|_{G_p(\bar{m}_{ij})=1\ \forall p>\tau} = 1$$

and the upper bound by

$$\begin{aligned} A_{ij\tau}|_{G_p(\bar{m}_{ij})=0\ \forall p>\tau} &= 1 + \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} && (81) \\ &= \frac{1}{1-\beta}. && (82) \end{aligned}$$

■

Proof of Lemma 4. To simplify notation, let $g_s = G_{\tau+s}(\bar{m}_{ij,\tau+s})$ for $s = 1, 2, 3, \ldots$. From the definitions in (28) and (29) — and without using Conjecture 1 or Conjecture 2, define

$$\begin{aligned} A_{ij\tau} &= 1 + \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} \prod_{p=\tau+1}^{s} [1 - G_p(\bar{m}_{ijp})] \\ &= 1 + \beta(1-g_1) + \beta^2(1-g_1)(1-g_2) + \beta^3(1-g_1)(1-g_2)(1-g_3) + \ldots \\ B_{ij\tau} &= 1 - \sum_{s=\tau+1}^{\infty} \beta^{s-\tau} G_s(\bar{m}_{ij}) \prod_{p=\tau+1}^{s-1} [1 - G_p(\bar{m}_{ij})] \\ &= 1 - \beta g_1 - \beta^2(1-g_1)g_2 - \beta^3(1-g_1)(1-g_2)g_3 + \ldots \end{aligned}$$

Now define the terms

$$\begin{aligned} a_0 &= 1 \\ a_s &= \beta^s \prod_{p=1}^{s} (1-g_p) \text{ for } s = 1, 2, 3, \ldots \end{aligned}$$

so that $A_{ij\tau} = \sum_{s=0}^{\infty} a_s$. Similarly, define

$$\begin{aligned} b_0 &= 1 \\ b_1 &= -\beta g_1 \\ b_s &= -\beta^s g_s \prod_{p=1}^{s-1} (1-g_p) \text{ for } s = 2, 3, 4, \ldots \end{aligned}$$

so that $B_{ij\tau} = \sum_{s=0}^{\infty} b_s$. Now notice that

$$\begin{aligned} a_1 &= \beta + b_1 \\ a_2 &= \beta^2 + \beta b_1 + b_2 \\ a_3 &= \beta^3 + \beta^2 b_1 + \beta b_2 + b_3 \end{aligned}$$

and so on, so that

$$\begin{aligned} a_s &= \beta^s + \beta^{s-1} b_1 + \beta^{s-2} b_2 + \beta^{s-3} b_3 + ... + \beta b_{s-1} + b_s \\ &= \sum_{p=0}^{s} \beta^{s-l} b_p \end{aligned}$$

for $s = 0, 1, 2, 3, ...$. Thus,

$$\begin{aligned} A_{ij\tau} &= \sum_{s=0}^{\infty} \sum_{p=0}^{s} \beta^{s-p} b_p \\ &= 1 + (\beta + b_1) + \left(\beta^2 + \beta b_1 + b_2\right) + \left(\beta^3 + \beta^2 b_1 + \beta b_2 + b_3\right) + ... \\ &= \left(1 + \beta + \beta^2 + \beta^3 + ...\right) + b_1 \left(1 + \beta + \beta^2 + \beta^3 + ...\right) + b_2 \left(1 + \beta + \beta^2 + \beta^3 + ...\right) + ... \\ &= \frac{1}{1-\beta} + \frac{b_1}{1-\beta} + \frac{b_2}{1-\beta} + ... \\ &= \frac{1}{1-\beta} \sum_{s=0}^{\infty} b_s \\ &= \frac{B_{ij\tau}}{1-\beta} \end{aligned}$$

■

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.

[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.

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.

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,

$$n_j = \begin{cases} pnN_j & \text{if } N_j \geq p^{-1} \\ n & \text{if } N_j < p^{-1} \end{cases}. \tag{83}$$

Sampling Rule 2 *If firm j is sampled into s in t, sample n_j period t employees of j into s.*

Together, Rule 2 and the definition of n_j in (83) 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

$$\begin{aligned}
\pi_1(i) &= \begin{cases} mpn & \text{if } N_j \geq p^{-1} \\ mpN_j n N_j^{-1} & \text{if } N_j < p^{-1} \end{cases} \\
&= mpn \qquad (84) \\
\pi_2(i) &= \begin{cases} (1-m)\,pn & \text{if } N_j \geq p^{-1} \\ (1-m)\,pN_j n N_j^{-1} & \text{if } N_j < p^{-1} \end{cases}. \\
&= (1-m)\,pn \qquad (85)
\end{aligned}$$

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 (84) and (85) 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: $\pi_s(i|i \text{ is sampled into } s') = 0$ 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, March). Econometric analysis of linked employer-employee data. *Labour Economics*, 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. 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 S. Robert A (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, June). Long-term partnership formation: Marriage and employment. *The Economic Journal 109*, F307–F334.

Burgess, S., J. Lane, and D. Stevens (2000). Job flows, worker flows and churning. *Journal of Labor Economics*, 473–502.

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

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

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

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

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

Gilmour, A. R., R. Thompson, and B. R. Cullis (1995, December). 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* (2), 192–218.

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

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

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.

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

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

Pissarides, C. (1984). Efficient job rejection. *The Economic Journal 94*(Supplement), 97–108.

Postel-Vinay, F. and J.-M. Robin (forthcoming). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*.

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.

Rubin, D. R. (1976). Inference and missing data. *Biometrika 63*(3), 581–592.

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, March). 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, March). State UI wage records: Description, access and use. Mimeo.

Stinson, M. (2002, September). Estimating the relationship between employer-provided health insurance, worker mobility, and wages. unpublished manuscript, Cornell University.

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

Topel, R. H. and M. P. Ward (1992, May). Job mobility and the careers of young men. *Quarterly Journal of Economics*, 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 Worker-Firm Matches in Smallest Group	5	5	5	1
Percentage of Worker-Firm 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
Parameter 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		0.230	(0.005)	0.171	(0.003)	0.177	(0.002)
Variance of firm effect (σ^2_ψ)	0.077		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		0.044	(0.001)	0.036	(0.000)	n/a	n/a
Corrected for Truncation								
Variance of person effect (σ^2_α)			0.226	(0.005)	0.164	(0.003)	0.169	(0.001)
Variance of firm effect (σ^2_ψ)			0.152	(0.002)	0.072	(0.002)	0.075	(0.007)
Variance of match effect (σ^2_γ)					0.071	(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)	0.134	(0.0005)
Corrected for Truncation								
Log Likelihood			263793	(1071.7)	280996	(943.4)	313900	(1283.0)
AIC			-527497	(2143.5)	-561901	(1886.8)	-627386	(2566.0)
BIC			-527023	(2143.7)	-561416	(1886.8)	-625153	(2565.7)
Var(out-of-sample prediction error)[d]			0.090	(0.0012)	0.069	(0.0007)	0.137	(0.0006)
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		0.551	(0.004)	0.496	(0.005)	0.402	(0.004)
Variance of firm effect (σ^2_ψ)	0.062		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		0.151	(0.001)	0.137	(0.001)	n/a	n/a
Corrected for Truncation								
Variance of person effect (σ^2_α)			0.550	(0.004)	0.492	(0.005)	0.399	(0.005)
Variance of firm effect (σ^2_ψ)			0.099	(0.003)	0.044	(0.002)	0.042	(0.002)
Variance of match effect (σ^2_γ)					0.081	(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)	0.243	(0.0059)
Corrected for Truncation								
Log Likelihood			66986	(872.9)	72753	(718.5)	106872	(1720.6)
AIC			-133884	(1745.8)	-145415	(1437.1)	-213330	(3441.3)
BIC			-133409	(1745.7)	-144930	(1436.9)	-211097	(3441.8)
Var(out-of-sample prediction error)[d]			0.131	(0.0006)	0.114	(0.0009)	0.245	(0.0057)
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
ESTIMATED VARIANCE COMPONENTS AND SUMMARY OF MODEL FIT STATISTICS
UNDER ALTERNATE RESIDUAL COVARIANCE STRUCTURES
Dependent Variable: Log(Real Annualized Earnings), Combined Results From 3 Completed Data Implicates
Standard Errors Omitted, All Parameter Estimates Significant at the 1 Percent Level

	Random Person and Firm Effects, AR(1) Residual	Random Person, Firm, and Match Effects, AR(1) Residual	Random Person and Firm Effects, AR(2) Residual	Random Person and Firm Effects, MA(1) Residual	Random Person, Firm, and Match Effects, MA(1) Residual	Random Person and Firm Effects, MA(2) Residual	Random Person, Firm, and Match Effects, MA(2) Residual	Random Person and Firm Effects, ARMA(1,1) Residual
	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]
No Correction for Truncation								
Variance of person effect (σ^2_α)	0.181	0.173	0.174	0.218	0.172	0.205	0.172	0.173
Variance of firm effect (σ^2_ψ)	0.079	0.074	0.073	0.134	0.074	0.116	0.074	0.072
Variance of match effect (σ^2_γ)		0.035			0.063		0.057	
Residual variance (σ^2_ε)	0.097	0.070	0.106	0.051	0.041	0.059	0.048	0.107
AR(1) tern (φ_1)	0.770	0.679	0.675					0.861
AR(2) term (φ_2)			0.145					
MA(1) tern (υ_1)				-0.535	-0.451	-0.634	-0.541	0.193
MA(2) tern (υ_2)						-0.396	-0.296	
Corrected for Truncation								
Variance of person effect (σ^2_α)	0.175	0.167	0.167	0.213	0.165	0.200	0.166	0.166
Variance of firm effect (σ^2_ψ)	0.077	0.071	0.073	0.133	0.072	0.115	0.072	0.070
Variance of match effect (σ^2_γ)		0.037			0.065		0.058	
Residual variance (σ^2_ε)	0.098	0.070	0.107	0.051	0.041	0.059	0.048	0.108
AR(1) tern (φ_1)	0.772	0.678	0.675					0.863
AR(2) term (φ_2)			0.146					
MA(1) tern (υ_1)				-0.536	-0.452	-0.635	-0.541	0.194
MA(2) tern (υ_2)						-0.397	-0.296	
Summary of Model Fit Statistics								
	Value[a]	Value[a]	Value[a]	Value[a]	Value[a]	Value[a]	Value[a]	Value[a]
No Correction for Truncation								
Log Likelihood	303323	304125	304523	288585	297529	296547	301655	304583
AIC	-606558	-607386	-608957	-577082	-594969	-593004	-603218	-609077
BIC	-606083	-606901	-608471	-576611	-594487	-592522	-602725	-608591
Var(out-of-sample prediction error)[d]	0.129	0.092	0.139	0.095	0.072	0.100	0.075	0.140
Corrected for Truncation								
Log Likelihood	303888	304333	305108	289008	298157	297003	302267	305171
AIC	-607687	-608573	-610123	-577927	-596223	-593914	-604441	-610251
BIC	-607201	-608077	-609627	-577445	-595731	-593421	-603937	-609754
Var(out-of-sample prediction error)[d]	0.131	0.092	0.141	0.095	0.072	0.101	0.076	0.143

[a] Simple average of parameter estimate across three completed data implicates.
[e] Simple variance of estimated person and firm effects, averaged across three completed data implicates.
[d] Computed on Dense Sample 2. Dependent variable is scaled to have unit variance.

TABLE 6
ESTIMATED VARIANCE COMPONENTS AND SUMMARY OF MODEL FIT STATISTICS
UNDER ALTERNATE RESIDUAL COVARIANCE STRUCTURES

Dependent Variable: Real Annualized Earnings (Unit Variance Scale)[1], Combined Results From 3 Completed Data Implicates
Standard Errors Omitted, All Parameter Estimates Significant at the 1 Percent Level

	Random Person and Firm Effects, AR(1) Residual	Random Person, Firm, and Match Effects, AR(1) Residual	Random Person and Firm Effects, AR(2) Residual	Random Person and Firm Effects, MA(1) Residual	Random Person, Firm, and Match Effects, MA(1) Residual	Random Person and Firm Effects, MA(2) Residual	Random Person, Firm, and Match Effects, MA(2) Residual	Random Person and Firm Effects, ARMA(1,1) Residual
	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]	Parameter Estimate[a]
No Correction for Truncation								
Variance of person effect (σ^2_α)	0.515	0.514	0.500	0.540	0.503	0.531	0.507	0.497
Variance of firm effect (σ^2_ψ)	0.045	0.045	0.039	0.074	0.045	0.060	0.045	0.038
Variance of match effect (σ^2_γ)		0.001			0.059		0.040	
Residual variance (σ^2_ε)	0.214	0.214	0.231	0.165	0.153	0.182	0.172	0.233
AR(1) tern (φ_1)	0.585	0.585	0.521					0.759
AR(2) term (φ_2)			0.150					
MA(1) tern (υ_1)				-0.416	-0.384	-0.482	-0.456	0.234
MA(2) tern (υ_2)						-0.284	-0.259	
Corrected for Truncation								
Variance of person effect (σ^2_α)	0.511	0.511	0.496	0.538	0.499	0.528	0.504	0.494
Variance of firm effect (σ^2_ψ)	0.044	0.044	0.038	0.074	0.044	0.059	0.044	0.038
Variance of match effect (σ^2_γ)		0.001			0.060		0.040	
Residual variance (σ^2_ε)	0.215	0.214	0.231	0.165	0.153	0.182	0.172	0.233
AR(1) tern (φ_1)	0.586	0.585	0.522					0.760
AR(2) term (φ_2)			0.150					
MA(1) tern (υ_1)				-0.416	-0.384	-0.483	-0.456	0.235
MA(2) tern (υ_2)						-0.284	-0.260	
Summary of Model Fit Statistics								
	Value[a]	Value[a]	Value[a]	Value[a]	Value[a]	Value[a]	Value[a]	Value[a]
No Correction for Truncation								
Log Likelihood	94710	94710	96268	86051	88054	92120	92778	96246
AIC	-189331	-189330	-192446	-172014	-176018	-184149	-185465	-192402
BIC	-188856	-188844	-191961	-171539	-175532	-183664	-184969	-191917
Var(out-of-sample prediction error)[d]	0.151	0.151	0.159	0.136	0.120	0.142	0.128	0.160
Corrected for Truncation								
Log Likelihood	94813	94813	96381	86120	88150	92204	92878	96360
AIC	-189535	-189534	-192671	-172150	-176209	-184315	-185663	-192628
BIC	-189050	-189038	-192175	-171664	-175713	-183819	-185156	-192132
Var(out-of-sample prediction error)[d]	0.152	0.151	0.160	0.140	0.120	0.142	0.128	0.161

[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.
[d] Computed on Dense Sample 2. Dependent variable is scaled to have unit variance.

TABLE 7
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 8
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	Uη	ψ	Xβ	$\beta_\lambda\lambda$
Log Earnings (y)	1.000	0.800	0.704	0.375	0.432	0.290	0.235
Pure Person Effect (θ)	0.800	1.000	0.908	0.407	0.004	0.018	0.127
Unobserved Component (α)	0.704	0.908	1.000	-0.012	-0.028	-0.030	0.107
Observed Component (Uη)	0.375	0.407	-0.012	1.000	0.072	0.110	0.069
Pure Firm Effect (ψ)	0.432	0.004	-0.028	0.072	1.000	0.000	-0.061
Time-Varying Covariates (Xβ)	0.290	0.018	-0.030	0.110	0.000	1.000	0.197
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.235	0.127	0.107	0.069	-0.061	0.197	1.000

Random Person, Firm, and Match Effects

	y	θ	α	Uη	ψ	γ	Xβ	$\beta_\lambda\lambda$
Log Earnings (y)	1.000	0.826	0.732	0.377	0.502	0.613	0.291	0.235
Pure Person Effect (θ)	0.826	1.000	0.851	0.514	0.218	0.480	0.037	0.103
Unobserved Component (α)	0.732	0.851	1.000	-0.013	0.202	0.566	-0.034	0.077
Observed Component (Uη)	0.377	0.514	-0.013	1.000	0.086	-0.011	0.127	0.069
Pure Firm Effect (ψ)	0.502	0.218	0.202	0.086	1.000	0.180	-0.009	-0.012
Pure Match Effect (γ)	0.613	0.480	0.566	-0.011	0.180	1.000	-0.030	-0.013
Time-Varying Covariates (Xβ)	0.291	0.037	-0.034	0.127	-0.009	-0.030	1.000	0.192
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.235	0.103	0.077	0.069	-0.012	-0.013	0.192	1.000

Random Person and Firm Effects, Unstructured Residual Covariance

	y	θ	α	Uη	ψ	Xβ	$\beta_\lambda\lambda$
Log Earnings (y)	1.000	0.817	0.730	0.361	0.504	0.301	0.235
Pure Person Effect (θ)	0.817	1.000	0.864	0.493	0.215	0.019	0.102
Unobserved Component (α)	0.730	0.864	1.000	-0.011	0.198	-0.030	0.078
Observed Component (Uη)	0.361	0.493	-0.011	1.000	0.085	0.090	0.068
Pure Firm Effect (ψ)	0.504	0.215	0.198	0.085	1.000	0.003	-0.004
Time-Varying Covariates (Xβ)	0.301	0.019	-0.030	0.090	0.003	1.000	0.164
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.235	0.102	0.078	0.068	-0.004	0.164	1.000

TABLE 9
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 10
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	Uη	ψ	Xβ	$\beta_\lambda\lambda$
Log Earnings (y)	1.000	0.846	0.812	0.248	0.392	0.262	0.141
Pure Person Effect (θ)	0.846	1.000	0.941	0.344	0.168	-0.059	0.086
Unobserved Component (α)	0.812	0.941	1.000	0.007	0.157	-0.012	0.072
Observed Component (Uη)	0.248	0.344	0.007	1.000	0.063	-0.139	0.053
Pure Firm Effect (ψ)	0.392	0.168	0.157	0.063	1.000	0.009	0.008
Time-Varying Covariates (Xβ)	0.262	-0.059	-0.012	-0.139	0.009	1.000	0.142
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.141	0.086	0.072	0.053	0.008	0.142	1.000

Random Person, Firm, and Match Effects

	y	θ	α	Uη	ψ	γ	Xβ	$\beta_\lambda\lambda$
Log Earnings (y)	1.000	0.852	0.821	0.247	0.420	0.624	0.262	0.141
Pure Person Effect (θ)	0.852	1.000	0.928	0.378	0.274	0.509	-0.064	0.080
Unobserved Component (α)	0.821	0.928	1.000	0.007	0.267	0.546	-0.012	0.065
Observed Component (Uη)	0.247	0.378	0.007	1.000	0.072	0.009	-0.142	0.053
Pure Firm Effect (ψ)	0.420	0.274	0.267	0.072	1.000	0.273	0.015	0.047
Pure Match Effect (γ)	0.624	0.509	0.546	0.009	0.273	1.000	0.000	-0.005
Time-Varying Covariates (Xβ)	0.262	-0.064	-0.012	-0.142	0.015	0.000	1.000	0.141
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.141	0.080	0.065	0.053	0.047	-0.005	0.141	1.000

Random Person and Firm Effects, Unstructured Residual Covariance

	y	θ	α	Uη	ψ	Xβ	$\beta_\lambda\lambda$
Log Earnings (y)	1.000	0.808	0.780	0.242	0.405	0.262	0.141
Pure Person Effect (θ)	0.808	1.000	0.911	0.419	0.257	-0.076	0.073
Unobserved Component (α)	0.780	0.911	1.000	0.006	0.253	-0.012	0.057
Observed Component (Uη)	0.242	0.419	0.006	1.000	0.068	-0.156	0.052
Pure Firm Effect (ψ)	0.405	0.257	0.253	0.068	1.000	0.027	0.032
Time-Varying Covariates (Xβ)	0.262	-0.076	-0.012	-0.156	0.027	1.000	0.124
Inverse Mills Ratio ($\beta_\lambda\lambda$)	0.141	0.073	0.057	0.052	0.032	0.124	1.000

TABLE 11

PARAMETER ESTIMATES FROM THE PROBIT EQUATION[a]

Combined Results From Three Completed Data Implicates

	Parameter Estimate	Standard Error
Variance of Person Effect (σ^2_ζ)	0.187	(0.007)
Variance of Firm Effect (σ^2_ν)	0.721	(0.012)
V_1[b]	0.821	(0.003)
V_2[c]	0.938	(0.068)
lnL[d]	-259780	(1450.2)

[a] The dependent variable is a (21 x 1) vector d_{ij} of employment outcomes for the match between worker *i* and firm *j*. The τ^{th} entry is 1 if the individual was employed at firm *j* at tenure τ years, and 0 otherwise.

[b] Estimated residual variance in the first 10 years of job tenure, on the scale of the data

[c] Estimated residual variance for 11-21 years of tenure, on the scale of the data

[d] Item in column labeled "Standard Error" is square root of between-implicate variance of lnL.

TABLE 12
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 13
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	0.127																				
2	0.072	0.095																			
3	0.052	0.069	0.090																		
4	0.042	0.056	0.068	0.090																	
5	0.031	0.046	0.057	0.070	0.087																
6	0.024	0.038	0.050	0.061	0.070	0.088															
7	0.019	0.033	0.045	0.056	0.063	0.073	0.091														
8	0.015	0.028	0.040	0.052	0.060	0.068	0.077	0.095													
9	0.013	0.026	0.037	0.049	0.057	0.064	0.072	0.082	0.099												
10	0.009	0.023	0.034	0.046	0.053	0.062	0.069	0.077	0.086	0.104											
11		0.021	0.031	0.043	0.050	0.059	0.066	0.073	0.080	0.091	0.109										
12			0.027	0.039	0.049	0.056	0.064	0.072	0.078	0.085	0.094	0.112									
13				0.037	0.047	0.052	0.058	0.067	0.073	0.080	0.087	0.099	0.117								
14					0.041	0.046	0.051	0.059	0.066	0.073	0.079	0.089	0.103	0.123							
15						0.034	0.039	0.046	0.052	0.060	0.065	0.074	0.086	0.100	0.126						
16							0.019	0.026	0.031	0.040	0.045	0.053	0.062	0.076	0.093	0.129					
17								0.007	0.011	0.016	0.019	0.026	0.033	0.044	0.061	0.091	0.131				
18									0.003	0.008	0.009	0.014	0.019	0.028	0.040	0.063	0.095	0.114			
19										0.004	0.006	0.008	0.011	0.018	0.027	0.042	0.067	0.083	0.098		
20											0.004	0.006	0.008	0.013	0.019	0.031	0.047	0.060	0.071	0.087	
21												0.004	0.006	0.009	0.014	0.022	0.035	0.043	0.052	0.063	0.081

TABLE 14
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 15

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	0.070	0.165																			
3	0.061	0.117	0.196																		
4	0.061	0.116	0.168	0.290																	
5	0.054	0.100	0.153	0.227	0.304																
6	0.039	0.110	0.154	0.223	0.257	0.371															
7	0.056	0.117	0.165	0.236	0.266	0.329	0.445														
8	0.071	0.096	0.147	0.229	0.268	0.321	0.383	0.498													
9	0.076	0.110	0.155	0.229	0.267	0.312	0.373	0.438	0.555												
10	0.065	0.115	0.157	0.226	0.268	0.315	0.378	0.433	0.488	0.616											
11		0.116	0.160	0.234	0.288	0.328	0.395	0.447	0.499	0.574	0.704										
12			0.143	0.225	0.266	0.319	0.384	0.439	0.489	0.548	0.614	0.740									
13				0.211	0.310	0.326	0.376	0.431	0.482	0.547	0.620	0.665	0.891								
14					0.292	0.325	0.353	0.394	0.447	0.511	0.590	0.643	0.763	0.941							
15						0.300	0.333	0.366	0.402	0.474	0.538	0.598	0.709	0.762	1.042						
16							0.237	0.261	0.312	0.361	0.434	0.466	0.635	0.706	0.745	0.985					
17								0.157	0.210	0.232	0.279	0.333	0.417	0.524	0.547	0.687	0.809				
18									0.151	0.170	0.208	0.237	0.290	0.366	0.370	0.474	0.527	0.525			
19										0.098	0.122	0.138	0.169	0.215	0.216	0.279	0.313	0.313	0.352		
20											0.073	0.082	0.101	0.128	0.128	0.165	0.186	0.185	0.207	0.290	
21												0.049	0.059	0.075	0.076	0.097	0.109	0.108	0.121	0.169	0.268

TABLE 16

EWMD Estimates of Structural Parameters, Scale Parameter[1] $\delta=1$

Combined Results From 3 Completed Data Implicates

	Log(Real Annualized Earnings) Scale		Real Annualized Earnings (Unit Variance) Scale			
	(1)		(2)		(3)	
	Maximum Tenure = 21 years		Maximum Tenure = 21 years		Maximum Tenure = 10 years	
Variance Parameter	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	(0.001)	0.224	(0.022)	0.085	(0.012)
Variance of Match Quality (σ^2_c)	0.049	(0.003)	0.360	(0.057)	0.733	(0.095)
Variance of Initial Signal (σ^2_z)	0.022	(0.005)	6.770	(3.880)	15.75	(3.04)
Variance of Production Outcomes (σ^2_e)	0.010	(0.001)	0.322	(0.214)	4.983	(1.775)
P-value from Chi-Square Test[c]	0.0014		0.0012		0.1447	
Corrected for Truncation						
Variance of Measurement Error (σ^2_u)	0.050	(0.001)	0.224	(0.022)	0.084	(0.006)
Variance of Match Quality (σ^2_c)	0.052	(0.003)	0.360	(0.033)	0.734	(0.095)
Variance of Initial Signal (σ^2_z)	0.025	(0.006)	6.694	(3.596)	15.542	(2.986)
Variance of Production Outcomes (σ^2_e)	0.009	(0.002)	0.287	(0.153)	4.947	(1.746)
P-value from Chi-Square Test[c]	0.0014		0.0013		0.0764	

[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

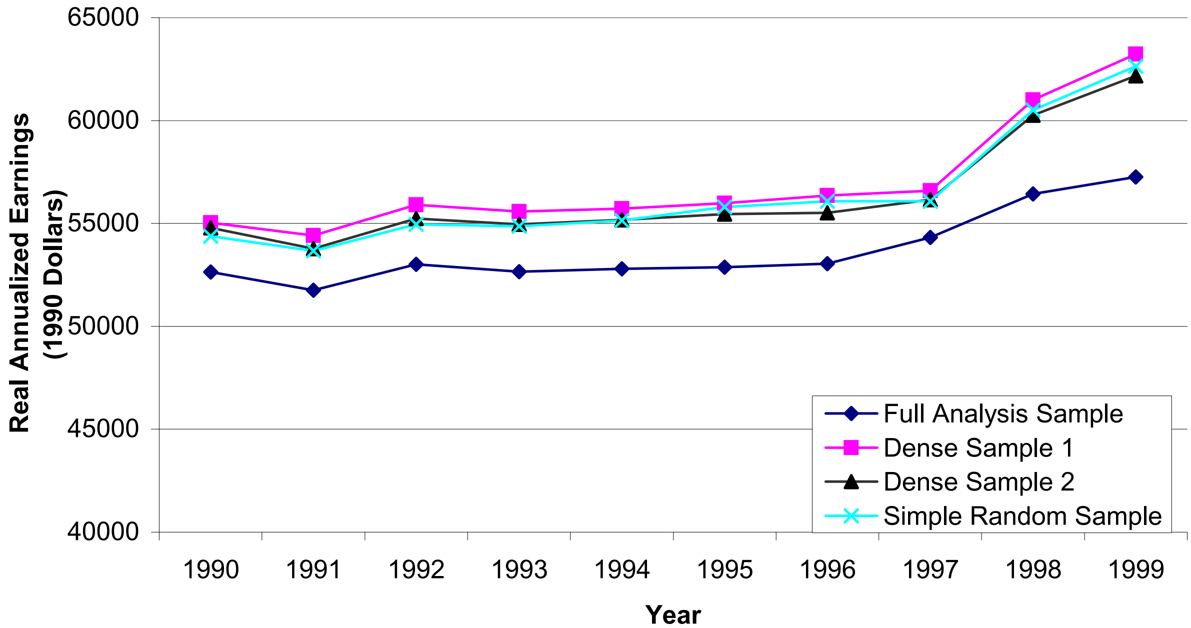
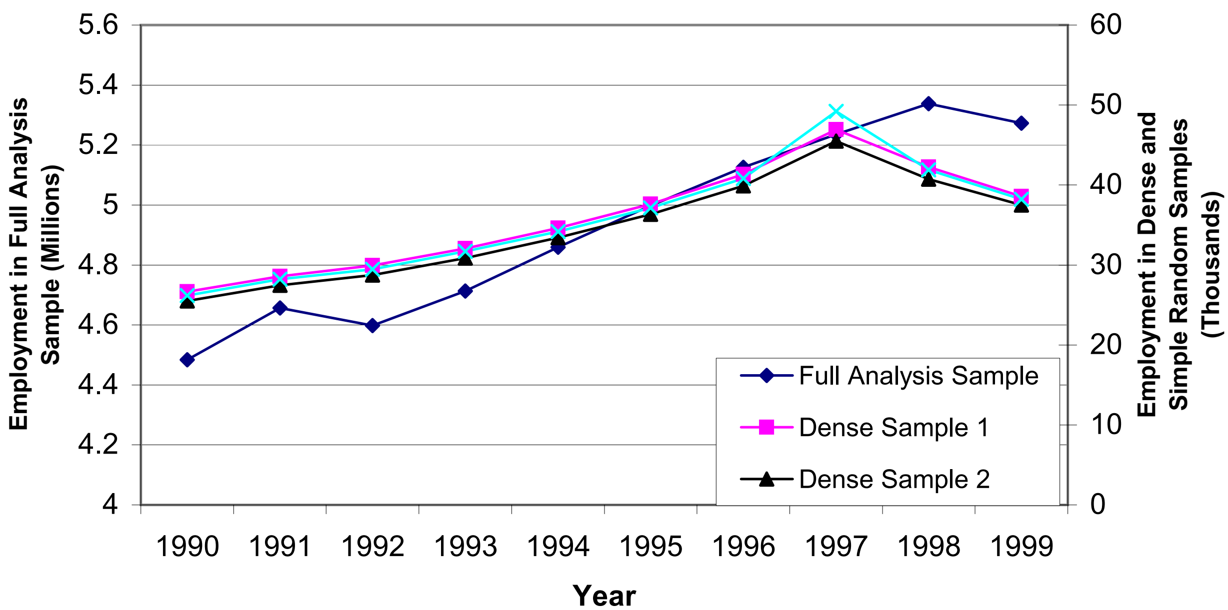


Figure 2
Yearly Time Series of Employment



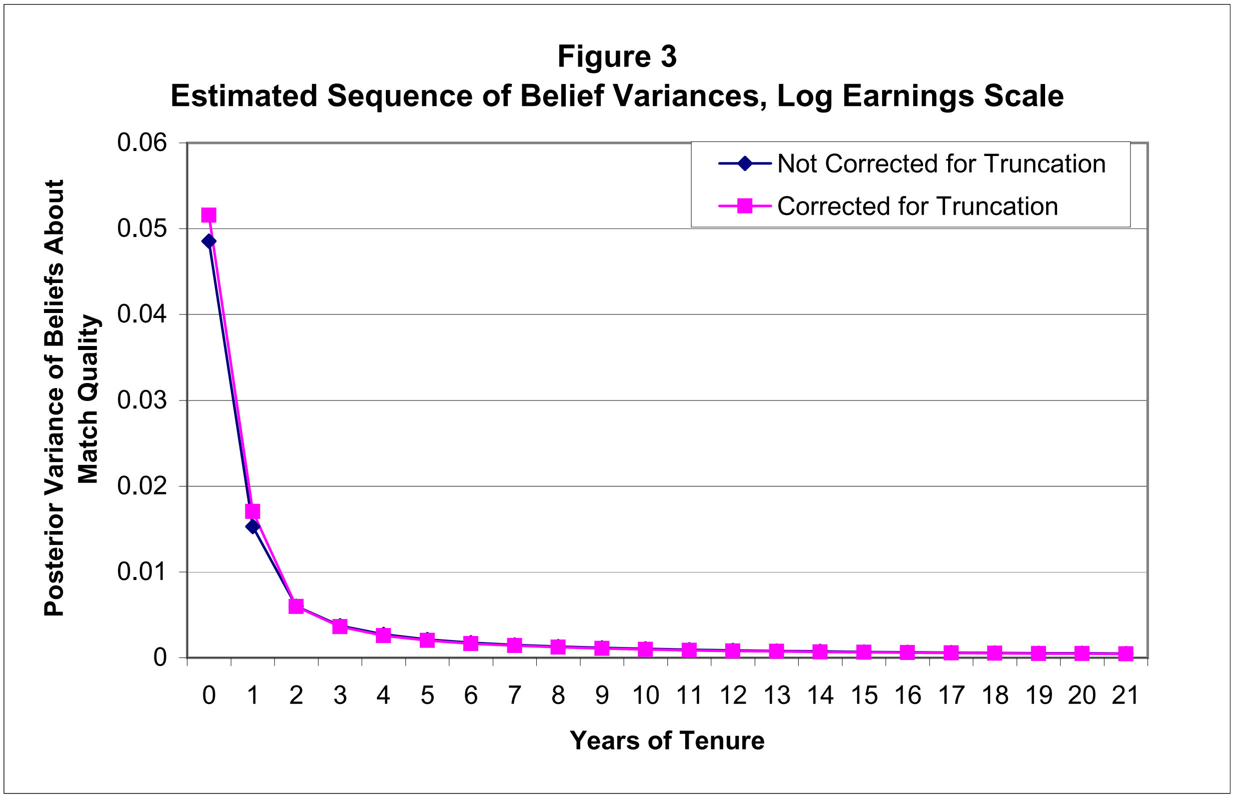


Figure 3
Estimated Sequence of Belief Variances, Log Earnings Scale
Posterior Variance of Beliefs About Match Quality
Years of Tenure
Not Corrected for Truncation
Corrected for Truncation
0.06
0.05
0.04
0.03
0.02
0.01
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

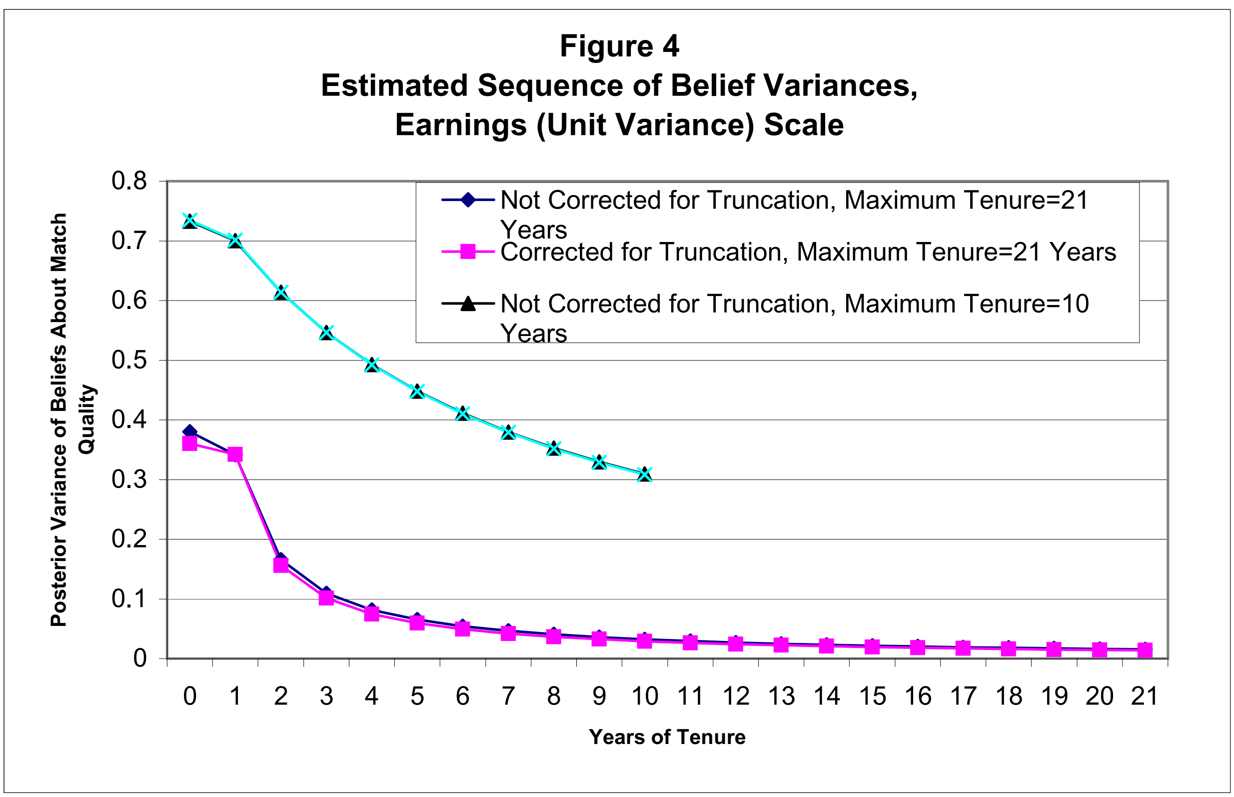


Figure 4
Estimated Sequence of Belief Variances,
Earnings (Unit Variance) Scale
Posterior Variance of Beliefs About Match Quality
Years of Tenure
Not Corrected for Truncation, Maximum Tenure=21 Years
Corrected for Truncation, Maximum Tenure=21 Years
Not Corrected for Truncation, Maximum Tenure=10 Years
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

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)

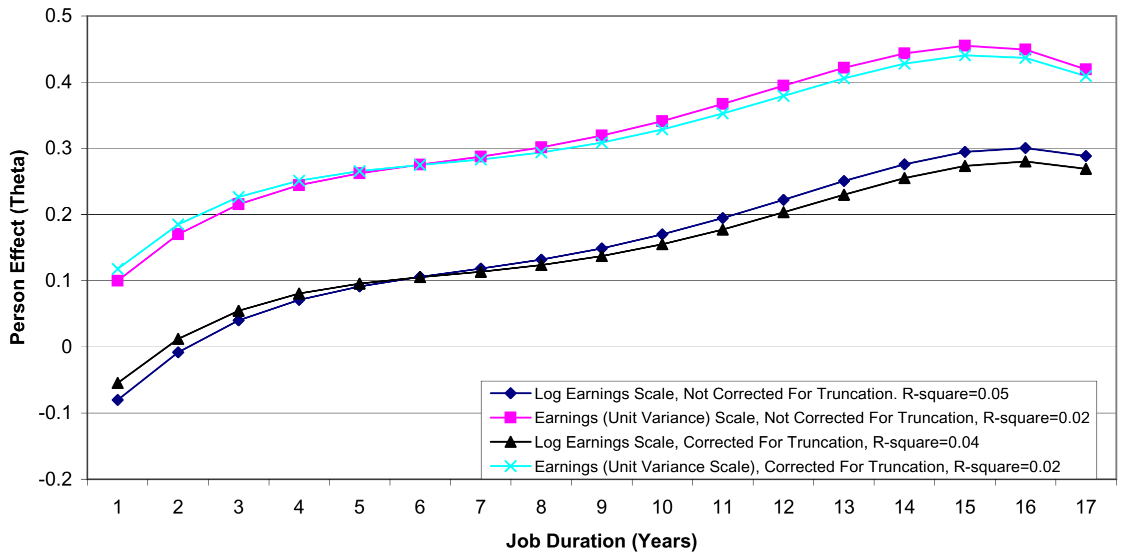


Figure 6
Estimated Relationship Between Firm Effect and Job Duration
Model With Unrestricted Within-Match Residual Covariance
Right-Censored Spells Excluded From Regression (N=54,661)

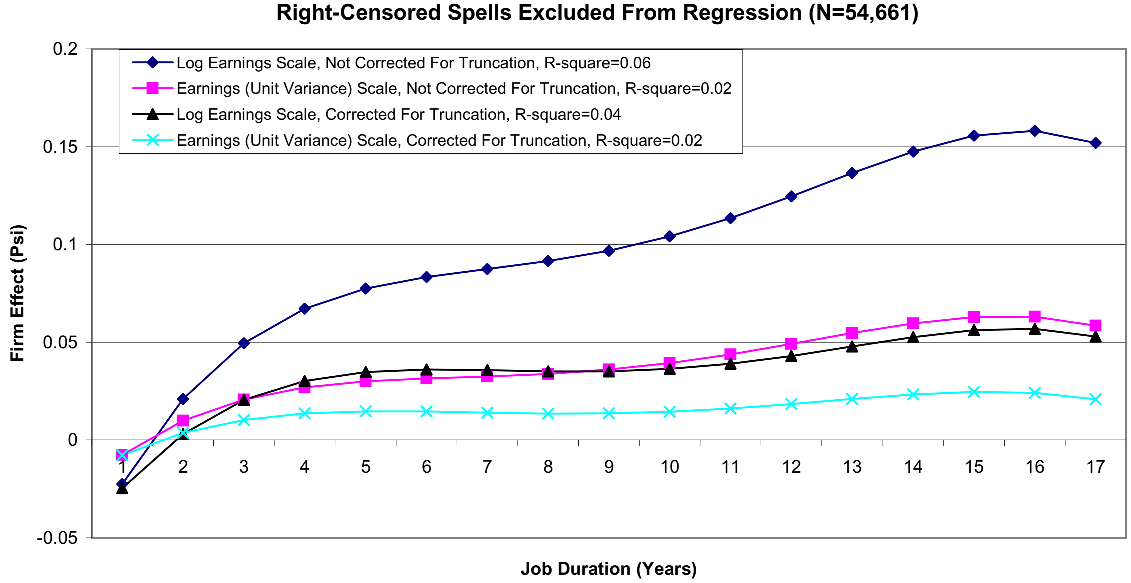


Figure 7
Estimated Relationship Between Firm Effect and Log(1997 Employment)
Model With Unrestricted Within-Match Residual Covariance (N=27,421)

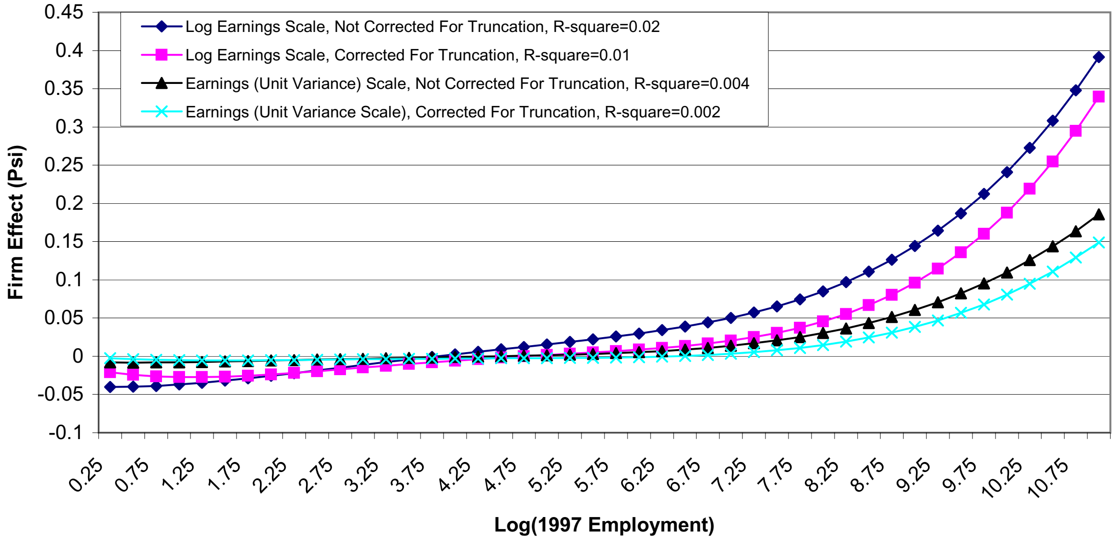


Figure 8
An Illustration of the Firm-Specific Sampling Rate