ENDOGENOUS MOBILITY

John M. Abowd
Department of Economics
Cornell University
John.Abowd@cornell.edu

Ian M. Schmutte
Department of Economics
Terry College of Business
schmutte@uga.edu

December 2013

PRELIMINARY DRAFT – PLEASE DO NOT CITE

Abstract

We correct for endogeneity bias in fixed-effects estimates of worker- and firm-specific earnings heterogeneity using longitudinal linked employer-employee data. Realized job assignments may be determined by unobservable components of earnings. We exploit the relational structure of the data to model the evolution of the matched data as an evolving bipartite graph in a Bayesian latent class framework. We estimate the model using data from the LEHD infrastructure file system of the U.S. Census Bureau. Our results suggest that the correction for endogeneity is meaningful but does not overturn the qualitative findings from previous analyses that assumed mobility to be exogenous.

JEL Codes: J64, J30, C11
Keywords: Earnings heterogeneity, Mobility bias, Search, Matching

Abowd acknowledges direct support from NSF Grants SES-0339191, CNS-0627680, SES-0922005, TC-1012593, and SES-1131848. Some of the research for this paper was conducted using the resources of the Social Science Gateway, which is partially supported by NSF grant SES-0922005. This paper was written while the first author was Distinguished Senior Research Fellow and the second author was RDC Administrator at the U.S. Census Bureau. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. This research uses data from the Census Bureau’s Longitudinal Employer-Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation.
1 Introduction

The objective of this paper is to explore the consequences of endogenous mobility for estimates of worker and firm effects in the decomposition of earnings in longitudinally-linked employer-employee data and to propose methods for correcting those estimates. Abowd et al. (1999) pioneered the identification, computation, and inference for the fixed-effects estimator of the decomposition of log earnings into components associated with unobserved worker and employer heterogeneity. A major factor in the interpretation of their statistical model is that it requires the assumption that the assignment of workers to firms is random conditional on all observable characteristics and the design of the time-invariant unobservables. This assumption is at odds with many models of job search and assignment, in particular, those in which workers sort into jobs according to their comparative advantage. Since structural interpretations of the measured heterogeneity have major consequences for our understanding of the labor market, it is important to relax the assumption that job mobility and assignments are exogenous to earnings.

The central problem of this paper is one manifestation of a fundamental challenge of empirical social science: separating the influence of correlated unobservables and sorting from the direct effect of group membership. Our approach is analogous to estimating treatment effects in the presence of selection on unobservables. In our work, the complication is that the number of possible treatments, that is employers, is in the millions. We construct an instrument for the actual assignment of workers to firms that exploits the relational structure of our data. The key insight is that the work histories of one’s coworkers and previous employers are informative of one’s own employment history, while being plausibly unrelated to whatever idiosyncratic wage innovations drive assignment at the margin.

Correcting estimated firm effects on wages for endogeneity bias is useful in a number of applications. First, this paper contributes to the ongoing debate as to whether estimates of employer-specific wage premia constitute evidence in contradiction of the law of one wage. If the bias correction does not affect the overall
contribution of firm effects in the presence of worker effects, it suggests that firms really do play an important direct role in wage determination, consistent with sociological evidence, but contrary to the neoclassical model of a competitive labor market. Furthermore, it helps to resolve some of the debates spawned by the early empirical results based on the assumption of exogenous mobility. These early results show little correlation between estimated employer wage premia and worker-specific earnings. In other words, high wage workers do not systematically appear in high-wage firms. This has often been cited as evidence against theories of assortative matching, and in favor of models of frictional search, which predict this lack of assortativity. The original empirical results spawned a theoretical literature attempting to construct frictional search models with assortative matching in which estimated employer and worker wage components would misrepresent the true assignment structure in the economy (Abowd et al. 2012; Shimer 2005; Lentz 2010).

Estimates of individual and firm effects are also being increasingly used in downstream applications. Iranzo et al. (2008) and Abowd et al. (2003) use estimates of person effects to measure the human capital distribution within firms. Combes et al. (2008) use a similar decomposition to estimate the contribution of neighborhoods to spatial earnings dispersion. Schmutte (2012) relies on consistent estimates of firm effects to infer the role of local job referral networks on earnings outcomes. Our estimates should be of interest in all such applications, as the specter of endogenous mobility clouds the interpretation of empirical results that rely on the consistency of estimated individual or firm effects. Recently, Card et al. (2012) have used longitudinally-linked German employer-employee data estimates produced under the assumption of exogenous mobility to conclude that employer effects figure prominently in the evolution of wage inequality.

There is also a parallel literature in the economics of education that uses value-added models to estimate the contributions of teachers and classrooms to student achievement. Endogenous assignment is just as much a problem for those models as it is here. Indeed, several recent studies have shown that the assumption of exogenous assignment in value-added models is rejected by the data (Roth-
stein 2010; Koedel and Betts 2010). Nevertheless, validation studies have shown that estimates from the value-added models are significantly correlated with independent assessments of teacher productivity. Our techniques provide a direct method of assessing whether correcting the endogeneity bias in value-added models would substantially change their results. Our method can be implemented as long as one has data in which the realized network of connections between individuals and groups is sufficiently detailed to provide identifying variation.

We proceed by setting up the log earnings decomposition proposed by Abowd et al. (1999) so that we can clearly articulate the nature of the endogenous mobility problem. Then, we report the results of two tests of the exogenous mobility assumption conducted by Abowd et al. (2010). As we will see, the null of exogenous mobility is rejected, but the associated analysis of mobility patterns provides interesting information about the nature of the true model. Next, we present an illustrative theoretical model with endogenous job mobility and suggest an empirical approach based on the use of relational information derived from the network structure. Turning to implementation, we set up a formal statistical model with endogenous mobility, in which earnings and job mobility are determined by latent classifications of workers, firms, and matches. This is a mixture model in which the probability of forming a link between a given worker and firm depends on a latent classification. We estimate it using the Gibbs sampler.

2 Background and Motivation

2.1 Exogenous Mobility in the Abowd, Kramarz and Margolis Model

Abowd et al. (1999) originally proposed the linear decomposition of log wage rates as the least squares fit of the equation

\[
\ln w = X\beta + D\theta + F\psi + \varepsilon
\]  

(1)
where $\ln w$ is the $[N \times 1]$ stacked vector of log wage outcomes $\ln w_{it}$, $X$ is the $[N \times k]$ design matrix of observable individual and employer time-varying characteristics (the intercept is normally suppressed, with $y$ and $X$ measured as deviations from overall means); $D$ is the $[N \times I]$ design matrix for the individual effects; $F$ is the $[w \times J - 1]$ design matrix for the employer effects (non-employment is suppressed here). $\varepsilon$ is the $[N \times 1]$ vector of statistical errors whose properties will be elaborated below; $\begin{bmatrix} \beta^T & \theta^T & \psi^T \end{bmatrix}^T$ are the unknown effects with dimension $[k \times 1], [I \times 1], [J - 1 \times 1]$ associated with each of the design matrices.\(^1\)

The assumption of exogenous mobility appears in the assumption that

$$E[\varepsilon | X, D, F] = 0.$$  

As long as the matrix of data moments has full rank – a non-trivial assumption – this conditional moment restriction yields a consistent estimator for the full parameter vector, including the individual and employer effects. Exogenous mobility imposes that a worker’s employment history is completely independent of the idiosyncratic part of earnings captured in $\varepsilon$, which in the AKM model includes the “match effect” – the average amount by which log wages in the current $(i, J (i, t))$ match deviate from their expected value. The AKM model is, thus, equivalent to assuming that all assignments are pre-determined at birth given full knowledge of $X, D, F$ and $\begin{bmatrix} \beta^T & \theta^T & \psi^T \end{bmatrix}^T$. There is no room for features included in many models of job mobility and assignment to affect either the duration of matches or the assignment of workers to particular employers.

Identification of $\begin{bmatrix} \beta^T & \theta^T & \psi^T \end{bmatrix}$

in the statistical model also requires that $[X D F]^T [X D F]$ be of full rank. Abowd et al. (2002) showed that this condition is equivalent to connectedness of the realized mobility network constructed by connecting sub-graphs of all workers who

\(^1\)The AKM formulation is an analysis of covariance with two high-dimension factors (individuals and employers) whose levels are estimated by least squares.
share a common employer and all employers who share a common worker over
the entire longitudinal sample. The realized mobility network is a static bipartite
graph on worker and employer nodes. As we will see, our identification strat-
ogy also has an interpretation in terms of the realized mobility network. We use
information in the realized mobility network that predicts employer assignments
but that we assume is conditionally independent of earnings residuals, exclusive
of the match effects, to achieve identification. The identification conditions for
the least squares solution for the parameters in equation (1) are orthogonality of
each component of the design matrix with respect to the estimated residual vector,
implying that the estimated effects are also orthogonal to the estimated residuals.

\[
\hat{\beta}^T X^T \hat{\varepsilon} = 0, \quad \hat{\theta}^T D^T \hat{\varepsilon} = 0 \quad \text{and} \quad \hat{\psi}^T F^T \hat{\varepsilon} = 0
\]  

(2)

2.2 Tests of Exogenous Mobility

Abowd et al. (2010) develop two formal tests for the null hypothesis of exogenous
job mobility against two omnibus alternatives that encompass many forms of en-
dogenous mobility. They apply these tests to longitudinally integrated employer-
employee data from the Longitudinal Employer-Household Dynamics (LEHD)
Program of the U.S. Census Bureau. Here we survey the basic nature of the tests
and their results. Their tests exploit the implicit restriction that future assignments
of workers to firms are uninformative about current earnings residuals. Under the
null hypothesis of exogenous mobility, these future assignments have no predic-
tive power with respect to the residual. The first test, “Test 1,” checks whether a
worker’s future employers are independent of the average residual in the current
job. The second test, “Test 2,” checks whether the future employees of a particular
employer are predictive of the residuals on their current period wage payments.

Both tests reject the null of exogenous mobility. The test statistic for Test 1 is
\[X^2_{8,991} = 7,438,692\] with \(Pr\{X^2_{8,991}\} < 0.001\). The test statistic for Test 2 is \(X^2_{900} = 172,295\) with \(Pr\{X^2_{900}\} < 0.001\). These are consistent with the related tests in Roth-
stein (2010) that reject the exogenous mobility assumption in value-added models.
for longitudinally linked education data.

3 Theoretical Model

3.1 Workers and Firms

The labor market consists of two discrete and finite populations: a population of \( I \) workers, and a population of \( J \) firms. \( I \) and \( J \) are both very large. Firms and workers match to produce output for an external market. Time is discrete.

Workers differ in the skills they bring to the market, and firms differ in their ability to extract output from a given quantity of efficient labor. When a worker is matched to a specific firm, the match can vary in quality in a manner that also affects productivity and output. More specifically,

- Any worker is endowed with ability, \( \theta \), drawn from the distribution \( S_\theta(\theta) \) with support \([\theta_{\min}, \theta_{\max}]\);
- Any firm is endowed with productivity \( \psi \), drawn from the distribution \( S_\psi(\psi) \) with support \([\psi_{\min}, \psi_{\max}]\);
- Any worker-firm match has a quality, \( \mu \), drawn from the conditional distribution \( H(\mu|\theta, \psi) \) with support \([\mu_{\min}, \mu_{\max}]\), independent of the support limits for \( \theta \) and \( \psi \).
- \( s_\theta(\theta), s_\psi(\psi), h(\mu|\theta, \psi) \) are probability densities corresponding to the distributions defined above.

Log output in period \( t \), \( z_{it} = \ln Z_{it} \) in a firm-worker match is defined as

\[
z_{it} = \psi_{J(i,t)} + \theta_i + \mu_{i,J(i,t)} \tag{3}
\]

where \( i \) is the name of the worker, whose ability is \( \theta_i \). \( J(i, t) \) identifies \( i \)'s employer at \( t \), the productivity of which is \( \psi_{J(i,t)} \). Finally, the quality of the match between
the two is \( \mu_{i,j(i,t)} \). \( \phi \) is the probability that an active match becomes inactive at the end of the current period.

If a match becomes inactive, the worker is unemployed at the start of the next period. Following Cahuc et al. (2006), an unemployed worker of type \( \theta \) receives log earnings in unemployment given by

\[
\ln w_{i0} = b + \theta_i
\]  

which is equivalent to being employed by a firm with productivity \( b \) in a match with quality \( \mu = 0 \) and receiving all of the output as a wage.

### 3.2 Matching

Workers and firms come together pairwise through sequential, random, and costly search. In each period, an unemployed worker is matched with a firm for an offer with probability \( \lambda_0 \), and an employed worker is matched with a firm for an offer with probability \( \lambda_1 \). Given that a worker has an offer, the firm from which it originates is randomly selected from the population of active firms as follows. The probability that a firm of type \( \psi \) is sampled is

\[
\hat{f}_t(\psi) = \frac{f(\psi)}{\sum_{z:z=\psi_j \text{ for some } j \text{ active at } t} f(z)}
\]

for all \( \psi \) such that some active firm \( j \) is type \( \psi_j \) at time \( t \). \( \hat{f}_t(\psi) = 0 \) for all \( \psi \) that are not represented among active firms at time \( t \). \( f \) represents the search distribution by type if the full population of types are represented. This cumbersome notation is necessary to handle two-sided matching in a model with discrete and finite quantities of workers and firms.
3.3 Wages

Firms offer workers a constant piece rate share of the match surplus. The match surplus depends on the value of unemployment, which is constant across workers of the same ability type, and the output from the current job. The piece rate $\alpha$ is constant across employer types and match qualities. The log wage offered to a worker with ability $\theta$ from a firm with type $\psi$ into a match with quality $\mu$ is therefore the sum of the piece rate and the individual productivity in the match

$$\ln w_{it} = \ln (\alpha Z_{it}) = \alpha + \theta_i + \psi_{J(i,t)} + \mu_{iJ(i,t)}.$$  \hspace{1cm} (6)

This is a simplification of equilibrium wage-posting models. Firms can match outside offers leading to renegotiation of the wage within a match. By focusing on a constant piece rate, we illuminate the selection issues driving endogenous mobility. We do, however, retain the implication that a worker only changes jobs to move into a better match.

3.4 Timing

At the end of the period we have the following transition probabilities:

- Matches exogenously dissolve with probability $\phi$.
- An employed worker receives outside offer with probability $\lambda_1$.
- Assume $\phi + \lambda_1 \leq 1$ (Probability that nothing happens to a worker is $[1 - (\phi + \lambda_1)]$).
- Unemployed workers receives offers with probability $\lambda_0$.

3.5 Job Mobility

Workers move when, and only when, they get an offer from job offering a higher wage. Let the joint firm-match productivity be $\kappa = \psi + \mu$. Define the sampling
distribution of firm-match quality offers to be $G(\kappa|\theta)$, which can be derived from $F$ (search distribution of employer productivities) and $H$ (the sampling distribution of match quality conditional on $\theta$ and $\psi$), which are defined above. Let $\kappa_{it} = \psi_{j(i,t)} + \mu_{i,j(i,t)}$ be the overall quality of worker $i$’s job in period $t$. Let the outside job offer be given by $\kappa^* = \zeta^* + \mu^*$.

For an employed worker, given $\theta_i$, transition dynamics over $\kappa_{it}$ are simply

$$
\kappa_{it+1} = \begin{cases} 
    b & \text{with prob. } \phi; \\
    \kappa_{it} & \text{with prob. } 1 - \phi - \lambda_1 [1 - G(\kappa_{it}|\theta_i)]; \\
    \kappa & \text{with density } \lambda_1 g(\kappa|\theta_i).
\end{cases}
$$

We can rewrite these in terms of firm and match heterogeneity so that the transition dynamics for an employed worker are

$$
(\psi_{it+1}, \mu_{it+1}) = \begin{cases} 
    (b, 0) & \text{with prob. } \phi; \\
    (\psi_{it}, \mu_{it}) & \text{with prob. } 1 - \phi - \lambda_1 [1 - G(\psi_{it} + \mu_{it}|\theta_i)]; \\
    (\psi, \mu) & \text{with density } \lambda_1 h(\mu|\psi, \theta_i) f(\psi).
\end{cases}
$$

For a non-employed worker, given $\theta_i$, transition dynamics are

$$
\kappa_{it+1} = \begin{cases} 
    b & \text{with prob. } 1 - \lambda_0; \\
    \kappa & \text{with density } \lambda_0 g(\kappa|\theta_i).
\end{cases}
$$

In terms of firm and match heterogeneity

$$
(\psi_{it+1}, \mu_{it+1}) = \begin{cases} 
    (b, 0) & \text{with prob. } 1 - \lambda_0; \\
    (\psi, \mu) & \text{with density } \lambda_0 h(\mu|\psi, \theta_i) f(\psi).
\end{cases}
$$

We now have enough information to completely specify wage and mobility dynamics. The model generates endogenous mobility because the correlation between $\theta, \psi$ and $\mu$ induces differential mobility. Again, for two workers of the same ability in the same firm, one with a better match will stay longer and exit to a better match. Therefore, assignment to a new employer is not independent either of
the match on the “inside” job, or of the “outside” job.

Relative to our more general model, presented below, this model imposes a restriction on the data generating process. Namely, the model assumes that the way firms are sampled does not depend on worker quality or match quality, so all of the observed association between worker type, firm type, residual, and future employer type is generated by the Roy-style selection process, which is embedded here into the probability of job change: \( \Pr(c = 1|\theta, \psi, \mu) \).

### 3.6 Moments of interest

Define the variable \( c_{it}^J = 1 \) if worker \( i \) changes jobs at the end of period \( t \). \( c_{it}^J = 0 \) otherwise. For employed workers, (and suppressing subscripts)

\[
\Pr(c^J = 1|\theta, \psi, \mu) = \lambda_1 [1 - G(\psi + \mu|\theta)].
\]

(11)

For non-employed workers

\[
\Pr(c^J = 1|\theta) = \lambda_0.
\]

(12)

The sampling distribution for aggregate firm-match quality given \( \theta \) is

\[
g(k|\theta) = \Pr(\psi, \mu|\psi + \mu = k, \theta)
\]

\[
= \int_{-\infty}^{\infty} h(k - \psi|\theta, \psi)f(\psi)d\psi.
\]

(13)
Therefore, the probability of changing jobs for a worker of type $\theta$, conditional on receiving an offer, is

$$1 - G(\bar{\psi} + \bar{\mu}|\theta) = \int_{\bar{\psi} + \bar{\mu}}^{\infty} g(k|\theta)dk = \int_{\bar{\psi} + \bar{\mu}}^{\infty} \left[ \int h(k - \psi|\theta, \psi)f(\psi)d\psi \right]dk = \int \left[ \int_{\bar{\psi} + \bar{\mu}}^{\infty} h(k - \psi|\theta, \psi)dk \right] f(\psi)d\psi = \int \left[ 1 - H(\bar{\psi} + \bar{\mu} - \psi|\theta, \psi) \right] f(\psi)d\psi = 1 - \mathbb{E}_F \left[ H(\bar{\psi} + \bar{\mu} - \psi|\theta, \psi) \right]$$

where $\bar{\psi} + \bar{\mu}$ is the productivity of the reservation offer. To summarize, we find the probability of drawing a match quality sufficient to move conditional on employer type, and integrate that across the employer-type distribution.

### 3.7 Likelihood Function for the Correlated Matches Model

For any worker, $i$, the model generates the following vector of observed data:

$$y_{it} = [\ln w_{it}, c_{it}, m_{it}, i, J(i, t), J(i, t + 1)],$$

where $J(i, t)$ indicates the employer of $i$ at the beginning of period $t$, $J(i, t + 1)$ indicates the employer of $i$ at the end of period $t$ (beginning of period $t + 1$), and $c_{it}$ identifies whether any change of employment status took place during period $t$. Specifically, $c_{it} = \max (c_{it}^m, c_{it}^l)$, where each entry can either be equal to zero or exactly one of the following:

- $c_{it}^m = 1$, if and only if $i$ exits to non-employment because of match dissolution;
- $c_{it}^l = 1$, if and only if $i$ makes a direct job-to-job transition.
\( m_{it} \) is a binary indicator with \( m_{it} = 1 \) if \( i \) is employed during period \( t \) and 0 otherwise.

Our model also generates latent data that are observed by the market participants, but not the econometrician. The latent data are the heterogeneity classifications of each worker, firm, and match that appears in the data. The latent data vector is therefore

\[
Z = [\theta_1, \ldots, \theta_I, \psi_1, \ldots, \psi_J, \mu_{11}, \mu_{12}, \ldots, \mu_{1J}, \mu_{21}, \ldots, \mu_{IJ}] .
\] (16)

In addition to the heterogeneity already described, we assume earnings are afflicted by classical measurement error with \( \varepsilon \sim N(0, \sigma^2) \), so that

\[
\ln w_{it} = \alpha + \theta_i + \psi_{J(i,t)} + \mu_{iJ(i,t)} + \varepsilon_{it}.
\] (17)

The parameters to be estimated are

\[
\rho = [\alpha, \theta, \psi, \mu, \lambda_0, \lambda_1, \phi, h, f, s_\psi, s_\theta]
\]

The likelihood of the joint distribution of the observed and latent data is

\[
L (\rho | Y, Z) \propto \prod_{i=1}^{T} \left( \prod_{t_i=1}^{T_i} \left( \frac{1}{\sqrt{2\pi}\sigma^2} \exp \left[ -\frac{(\ln w_{it} - \alpha - \theta_i - \psi_{J(i,t)} - \mu_{iJ(i,t)})^2}{2\sigma^2} \right] \right) \right)^{m_{it}}
\]

\[
\times \prod_{i=1}^{I} \left\{ \prod_{t_i=1}^{T_i-1} \left[ 1 - \phi - \lambda_1 [1 - G(\kappa_{it}[\theta_i])] \right]^{1 - c_{it}} \left[ \lambda_1 [1 - G(\kappa_{it}[\theta_i])] \right]^{c_{it}} \right\}^{1 - m_{it}}
\]

\[
\times \prod_{i=1}^{T_i} \left\{ \prod_{t_i=1}^{T_i-1} \left[ 1 - (\lambda_0 (1 - G(b|\theta_i))) \right]^{1 - c_{it}} \left[ \lambda_0 (1 - G(b|\theta_i)) \right]^{c_{it}} \right\}^{c_{it}}
\]

\[
\times \prod_{i=1}^{T_i-1} \left[ f(\psi_{J(i,t+1)}) \right]^{c_{it}}
\]

\[
\times \prod_{i=1}^{I} \prod_{j=1}^{J} h(\mu_{ij}|\theta_i, \psi_j) s_\psi(\psi_j) s_\theta(\theta_i)
\] (18)
The bracketed term contains the likelihood contribution of each individual’s work history. The first row in the bracketed term contains the probability of the wage, when one is observed. The second and third rows give the probability of different kinds of transition for a worker who is currently employed. The fourth row is the probability of exiting or remaining in non-employment for a non-employed worker. The fifth row expresses the probability of moving into a job with an employer of a particular productivity and a match of a particular quality. The final row in the likelihood function contains the probability of the latent heterogeneity, on which the rest of the likelihood function is conditioned.

4 Empirical Model

We use a latent class model to identify workers and firms with similar mobility and earnings patterns and then to estimate the effect on earnings related to membership in these classes. We conduct Bayesian inference using an adaptation of the Gibbs sampler algorithm for finite mixture models (Tanner 1996; Diebolt and Robert 1994) to our case allowing for multiple overlapping levels of correlation across observations. Our application and proposed procedure are related to stochastic blockmodels and other methods for the detection of “communities” of nodes in social networks. Our main innovation is the use of both node and edge characteristics in predicting the matches (Hoff et al. 2002; Newman and Leicht 2007; Neville and Jensen 2005).

4.1 Model Setup

Agents of the model are workers, indexed \( i \in \{1 \ldots I\} \equiv I \) and firms, indexed \( j \in \{0 \ldots J\} \equiv J \), where \( j = 0 \) is “not employed.” Each worker has a latent ability class denoted \( a_i \in A \) and each firm, except \( j = 0 \), has a latent productivity class denoted \( b_j \in B \). In addition, each worker-firm match has an associated heterogeneity component that affects both wages and mobility: \( k_{ij} \in K \). \( A, B \) and \( K \) are discrete with cardinality \( L, M + 1 \) and \( Q \). The “not-employed firm” is
a single entity in its heterogeneity class, so the class $b_0$ has no employer heterogeneity. To make the subsequent formulas easier to interpret, assume that the elements of $A, B$ and $K$ are rows from the identity matrices $I_L, I_M$ and $I_Q$, respectively. For instance, if $L = 2$, we have $A = \{(1, 0), (0, 1)\}$. The assignments of workers and firms to ability and productivity classes are independent multinomial random variables with parameters $\pi_a, \pi_b$. We allow for endogeneity in the match quality by letting the probability of $k$ depend on ability and productivity. So $\Pr(k_{ij} = k|a_i = a, b_j = b) = \pi_{k|ab}$.

The log of earnings, when actually employed in any match, is given by

$$\ln w_{ijt} = \alpha + X_{it}\beta + a_i\theta + b_j\psi + k_{ij}\mu + \varepsilon_{it}$$

(19)

where $X$ is a vector of observable time-varying characteristics and $\theta, \psi, \mu$ are vectors of parameters describing the effect on the level of log earnings associated with membership in the various heterogeneity classes. We take $\varepsilon$ to be normal with mean 0 and variance $\sigma^2$, independent and identically distributed across individuals and over time. When not employed, the individual earns a reservation log wage of

$$\ln w_{i0t} = \alpha + X_{it}\beta + a_i\theta + b_0\psi_0 + \varepsilon_{it}$$

(20)

where $\psi_0$ is the fixed productivity in the non-market sector, which is constrained to equal the minimum estimated $\psi$ in the employer population.

We formalize endogenous mobility by allowing those matches and employment durations that are observed to depend on ability, productivity and match quality. Let $J(i, t)$ be the index function that returns the identifier of the firm in which $i$ is employed in period $t$. Define the variable $s_{it} = 1$ if $i$ separates from his current job at the end of period $t$ and $s_{it} = 0$ otherwise. We let the probability of separation depend on the match quality by specifying

$$\Pr(s_{it} = 1|k_{i, J(i,t)}) = f_{se} (a_i, b_{J(i,t)}, k_{i, J(i,t)}; \gamma) \equiv \gamma_{abk}$$

(21)

where $0 \leq \gamma_{abk} \leq 1$. Conditional on separation, the productivity class of the next
employer depends on the productivity of the current employer, the ability of the worker, and the quality of the current match

$$\Pr(b_{j(i,t+1)} | a_i, b_{j(i,t)}, k_{i,j(i,t)}) = f_{tr}(a_i, b_{j(i,t)}, k_{i,j(i,t)}; \delta) \equiv \delta_{abk} \in \Delta^{M+1} \quad (22)$$

where $\delta_{abk} \equiv [\delta_{0|abk}, ..., \delta_{M|abk}]$ is a $1 \times (M + 1)$ vector of transition probabilities, $\Delta^{M+1}$ is the unit simplex, and $J(i,0) = 0$ for all $i$. The transition probabilities are indexed by all of the latent heterogeneity in the model. Within a heterogeneity class, the identity of the precise employer selected is completely random, as is the identity of an individual within an ability class.

### 4.2 Likelihood functions

We begin by developing the likelihood function for the observed and latent data. The observed data, $y_{it}$, consist of wage rates, observable time-varying characteristics, separations, accessions, and identifier information:

$$y_{it} = [\ln w_{i,j(i,t)}, X_{it}, s_{it}, i, J(i,t), J(i,t+1)] \text{ for } i = 1, ..., I \text{ and } t = 1, ..., T. \quad (23)$$

The latent data vector, $Z$, consists of the heterogeneity classifications:

$$Z = [a_1, ..., a_I, b_1, ..., b_J, k_{i11}, k_{i12}, ..., k_{i1}, k_{i21}, ..., k_{iJ}]. \quad (24)$$

In practice, we only use or update the heterogeneity classifications for the matches that actually occur, the number of which is bounded above by $T \times I$. That is, we only care about $k_{ij}$ where $i, j$ is such that $j = J(i,t)$ for some $t$. Finally, the complete parameter vector is

$$\rho^T = [\alpha, \beta^T, \theta^T, \psi^T, \mu^T, \sigma, \gamma, \delta, \pi_a, \pi_b, \pi_{k|ab}], \rho \in \Theta \quad (25)$$

We assume that workers and firms are infinitely-lived. The complete process starts at $t = 1$, with continuous sampling continuing to date $T$. We model initial
conditions by assuming that everyone enters the labor force at \( t = 1 \) and is assigned an employer completely at random. In other words, we assume that the matches initially observed are exogenous. The observed data matrix for this time interval is denoted \( Y \). The likelihood function for the joint distribution \((Y, Z)\) is given by

\[
\mathcal{L}(\rho | Y, Z) \propto \prod_{i=1}^{I} \left[ \prod_{t=1}^{T-1} \left[ 1 - \gamma(a_i) \gamma(b_{j(i,t)}) \gamma(k_{j(i,t)}) \right] \right]^{s_{it}} \times \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{\ell=1}^{L} \prod_{m=1}^{M} \prod_{q=1}^{Q} \left[ \left( \pi_{a\ell} \right)^{a_{i\ell}} \left( \pi_{b\ell} \right)^{b_{jm}} \left( \pi_{q\ell} \right)^{k_{ijq}} \right]
\]

where the notation \( \pi_{a\ell} \) denotes the \( \ell^{th} \) element of \( \pi_a \) (similarly for \( \pi_{b\ell}, b_{jm}, \) etc.) and \( \langle x \rangle \) means the index of the non-zero element of the vector \( x \).

Note that the likelihood function in Equation (26) is directly related to the likelihood function derived from the theoretical model in Equation (18). Specifically, we discretized the latent worker, firm, and match heterogeneity distributions to avoid making parametric restrictions on those distributions. Note that we still allow for arbitrary correlation between the match effects and the worker and firm effects. Furthermore, the statistical likelihood (26) generalizes (18) by allowing the type of the incumbent firm and match to have an independent effect on the sampling distribution of firm heterogeneity. In the theoretical model, any correlation between incumbent match quality and employer type is driven by selection and not unobserved heterogeneity.

4.3 Link Between Structure and AKM

To formalize the connection between the model and the realized mobility network, it is useful to think about the steady-state distribution of employment spells across
matches of different types. Under certain regularity conditions, the Markov process defined by the model, which articulates how workers move from one type of match to the next, has a unique stationary distribution.

- Let \( \lambda_{\ell,m,q} \) be the measure of type \((\ell, m, q)\) matches observed in the steady-state.
- Define the diagonal matrix
  \[
  \Lambda = \text{diag}(\lambda_{111}, \lambda_{112}, \ldots, \lambda_{LMQ})^T.
  \]

Note that \( \Lambda \) does not account for transitions to non-employment. In the \( 2 \times 2 \times 2 \) case, \( \Lambda \) is an \( 8 \times 8 \) matrix.

- In steady-state, observed log earnings data \( \ln w \) are drawn from a discrete distribution proportional to \( \Lambda \). \( \ln w \) is \( LMQ \times 1 \) vector with
  \[
  \ln w_{\ell,m,q} = \alpha + \theta_{\ell} + \psi_b + \mu_q.
  \]

### 4.4 Network Interpretation of Endogenous Mobility Models

Continuing with the notation developed above, define a set of indicator matrices analogous to the person, employer, and match design matrices. For the \( 2 \times 2 \times 2 \) model, this matrix is simply

\[
\begin{bmatrix}
D & F & G
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 
\end{bmatrix}.
\]
The abuse of notation associated with using \( D, F, \) and \( G \) in this context helps highlight the connection between this development and the conventional AKM model.

Observe that the full cross product matrix is just

\[
\begin{bmatrix}
D & F & G
\end{bmatrix}^T \Lambda \begin{bmatrix}
D & F & G
\end{bmatrix} =
\begin{bmatrix}
D^T \Lambda D & D^T \Lambda F & D^T \Lambda G \\
F^T \Lambda D & F^T \Lambda F & F^T \Lambda G \\
G^T \Lambda D & G^T \Lambda F & G^T \Lambda G
\end{bmatrix}.
\]

Notice that the upper-left block of the cross-product matrix in (30) is a model for the Laplacian of the realized mobility network, which is random noise around this steady-state distribution.

**4.5 Estimation Procedure**

The likelihood function factors into a part due to the observed data conditioned on the latent data, and the latent data conditioned on the parameters. The observed-data likelihood conditional on the latent data factors further into separate contributions from the earnings and the mobility processes. The mobility process is Markov, and conditionally independent of the earnings realizations once we know the latent classifications of the workers, firms and matches. The power of the model comes from the predictive equation for \( Z \) given the observed data and the parameters, which we can compute as the complete-data likelihood divided by the observed-data likelihood. The observed-data likelihood is calculated by integrating out the latent data.

We start with initial values for the parameter vector and latent data, \( \rho^{(0)}, Z^{(0)} \).

We defined above the distributions of the parameters given the observed and latent data. To complete the specification, we define the distributions for the latent variables conditional on the observed data and the parameters. For instance, to update the ability classifications for the workers, we need to sample from a multi-
nominal with probability of the $\ell^{th}$ class equal to

$$p(a_i = \ell | a_{-i}, b, k, Y, \rho) = \frac{p(a_{-i}, b, k, Y | \rho, a_i = \ell) p(a_i = \ell)}{p(a_{-i}, b, k, Y | \rho)}$$

$$= \frac{\pi_{a\ell} p(a_{-i}, b, k, Y | \rho, a_i = \ell) p(a_i = \ell)}{\sum_{\ell' = 1}^{L} [\pi_{a\ell'} p(a_{-i}, b, k, Y | \rho, a_i = \ell') p(a_i = \ell')]}.$$  \hspace{1cm} (31)

This requires computing the likelihood function under each assignment of $i$ to an ability classification, requiring roughly $L$ evaluations of the likelihood per individual. Since the posterior probability of $a_i$ is independent of $a_i$, conditional on the rest of the data (latent and observed),

$$p(a_i = \ell | a_{-i}, b, k, Y, \rho) = p(a_i = l | b, k, Y, \rho).$$ \hspace{1cm} (32)

This fact allows us to speed computation by updating the latent classifications of each worker in parallel. The proof is a straightforward consequence of the conditional independence in the likelihood of the augmented data.

We can show in similar fashion that

$$p(k_s = q | a, b, k_{-s}, Y, \rho) = p(k_s = q | a, b, Y, \rho),$$ \hspace{1cm} (33)

which likewise follows from the conditional independence assumptions in the model. Hence, for a given classification of workers and employers, $(a, b)$, the latent quality of each match is conditionally independent from the others. We exploit the conditional independence by parallelizing these updates as well.

The posterior distribution for employer types exhibits a conditional dependence that is not present for workers or matches. This is because, when a worker changes jobs, the type of firm in the successor job is a function of the employer type in the current job. Therefore, the posterior probability that a firm is of a particular type depends directly on the types of firms it is connected to through the realized mobility network. The following result arises from the same logic as the
formulas above
\[ p(b_j = m|a, b_{-j}, k, Y, \rho) = p(b_j = m|a, b_{N(j)}, k, Y, \rho) \] (34)
where \( b_{N(j)} \) denotes the latent classifications of the employers in \( N(j) \), the set of neighbors of \( j \) (in the employer projection of the realized mobility network).

To understand our updating algorithm, consider a “coloring” of the employers, \( \varphi : J \rightarrow \{1, \ldots, N_C\} \) such that \( c(j) = c(j') \) if and only if \( N(j) \cap N(j') = \emptyset \). Then, the conditional posterior probability of the vector of employer types is given by
\[ p(b|a, k, Y, \rho) = \prod_{c=1}^{N_C} \left[ \prod_{j: c(j) = c} p\left( b_j = m|a, b_{N(j)}, k, Y, \rho \right) \right]. \] (35)
By construction, the inner terms are conditionally independent, hence they can be computed independently (in parallel).

With the posterior distributions as defined in the previous section, the Gibbs sampler can be implemented as follows:

\[ \sigma^{(1)} \sim p\left( \sigma | \alpha^{(0)}, \theta^{(0)T}, \psi^{(0)T}, \mu^{(0)T}, \gamma^{(0)}, \delta^{(0)}, \pi_a^{(0)}, \pi_b^{(0)}, \pi_{k|ab}^{(0)}, Z^{(0)}, Y \right) \] (36)

\[ \begin{bmatrix} \alpha \\ \beta \\ \theta \\ \psi \\ \mu \end{bmatrix}^{(1)} \sim p\left( \begin{bmatrix} \alpha \\ \beta \\ \theta \\ \psi \\ \mu \end{bmatrix} | \gamma^{(0)}, \delta^{(0)}, \pi_a^{(0)}, \pi_b^{(0)}, \pi_{k|ab}^{(0)}, Z^{(0)}, \sigma^{(1)}, Y \right) \] (37)

\[ \gamma^{(1)} \sim p\left( \gamma | \delta^{(0)}, \pi_a^{(0)}, \pi_b^{(0)}, \pi_{k|ab}^{(0)}, Z^{(0)}, \alpha^{(1)}, \theta^{(1)T}, \psi^{(1)T}, \mu^{(1)T}, Y \right) \] (38)

\[ \vdots \] (39)

\[ k_{IJ}^{(1)} \sim p\left( k_{IJ} | \rho^{(1)}, a_1^{(1)}, \ldots, a_I^{(1)}, b_1^{(1)}, \ldots, b_J^{(1)}, k_{11}^{(1)}, \ldots, k_{I-1}^{(1)}, Y \right) \] (40)
**Graph Coloring** For a general graph, the problem of finding the minimum number of colors is intractable. For our task, it is sufficient to come up with a coloring that yields a small number of partitions. To that end, we use the greedy sequential coloring algorithm described in Gebremedhin et al. (2005). The algorithm works by sorting network nodes from highest to lowest degree (that is, sorting employers in descending order by the number of job-to-job separations). The first node is assigned a color at random. For every other node, we assign the least frequent color that has not already been applied to one of its neighbors. If there is no such color, we add a new color to the list. In our data, the application of this algorithm yields a coloring that partitions the employers into 24 subsets, which is well below the algorithmic worst-case guarantee\(^2\).

### 4.6 Calculation of Monte Carlo Standard Errors

We report Monte Carlo standard errors (MCSE) in place of the posterior standard deviation. The MCSE is a measure of uncertainty about the location of the posterior distribution associated with using a finite set of potentially serially correlated draws from the Markov Chain Monte Carlo simulation. In practice, there is a substantial amount of autocorrelation in samples drawn under our model. The MCSE uses time-series methods to provide a measure of uncertainty about the location of the posterior distribution that reflects the autocorrelation in the simulated values. This method lets us to take full advantage of the information in the simulated values relative to more conventional ad hoc approaches like thinning the sample.

Several approaches to calculating the MCSE are possible, and the interested reader is referred to the survey chapter by Geyer (2011) for an overview of the topic. We implement the multivariate extension due to Kosorok (2000) of “initial sequence methods” originally proposed by Geyer (1992).\(^3\) The initial sequence

\(^2\)The worst-case would be a coloring with as many colors as the highest-degree node in the graph, which in our case is on the order of thousands

\(^3\)We prepared our own code to compute the initial sequence estimates of the MCSE. Univariate initial sequence estimation is available through the *MCMC* package in R. In practice, we compute the univariate MCSE for each parameter due to numerical instability in the autocovariance matrices.
methods use reversibility of the Markov Chain to find the largest lag that should be included in the autocorrelation calculation. There are three initial sequence estimators of the asymptotic variance of the posterior expectation:

- initial positive sequence
- initial monotone sequence
- initial convex sequence

The key theoretical result is that these estimators provide asymptotic overestimates of the true variance, with the initial convex sequence estimator giving the smallest overestimate, and the initial positive sequence estimator giving the largest. We calculate the MCSE under all three methods, and find the results to be practically indistinguishable. The values reported in our tables are estimates from the initial positive sequence, which are the most conservative.

5 Results

We implement the model empirically using matched employer-employee data from the LEHD program of the U.S. Census Bureau. The basic structure of these data is described in Abowd et al. (2009). The data processing and estimation procedures used to create the AKM decomposition for these data are described in Abowd et al. (2003).

5.1 Data Used for Structural Estimation

We use two estimation samples for the empirical results described below. The first of these samples was used to recompute the AKM decomposition for the entire U.S. universe. It is described in detail in other papers.\(^4\)

The estimation sample for the structural model in this paper began with the universe of persons employed in the states of Illinois, Indiana, and Wisconsin in

\(^4\)Abowd, McKinney, Schmutte and Zhao (in preparation).
all sectors except federal employees (who are not yet available in these LEHD micro-data) in the years 1999-2003. For these individuals, we selected all jobs, regardless of the state in which they occurred, over the years 1990-2010. Thus, the complete 21-year work history was used. In each year, the dominant job, defined as the one that produced the greatest labor market earnings, was retained. If there was no labor force income from any job, that year was marked as “not employed.”

There are 16.9 million persons in these data. We used the estimated AKM decomposition for their jobs, taken directly from the universe statistics described above, for comparison to our structural estimation. The selected individuals and the control regressors are identical for both the universe AKM decomposition and the structural estimation.

To create starting values for our models, the estimated values of the AKM parameters $\theta$ and $\psi$ from the universe of Illinois, Indiana and Wisconsin jobs over the entire 1990-2010 period were discretized into ten categories based on the deciles for this of persons and employers. The AKM residual was averaged over all periods in which a match was present to obtain the continuous version of the AKM $\mu$, which was then discretized into ten categories using the deciles of the match population.

5.2 Estimation Sample and Details

We fit the likelihood function in equation (26) using a 0.5% simple random sample of individuals, keeping all employers and match-years associated with those individuals. The 0.5% simple random sample of persons has 84,690 Persons, 181,592

---

5Over the period from 1990 to 2001, the data for many states in the LEHD infrastructure are not available due to the start date for that state’s participation in the program and the availability of data in the state’s archival repositories. The start dates for the data in each state are exogenous to any statistical model. They depended entirely on the condition of the state’s information technology and the willingness to join the Local Employment Dynamics federal/state partnership that coordinates these activities for the Census Bureau and the individual states. Thus, early data in this system are missing completely at random. However, the pseudo-employer in our model that represents “not employed” includes individuals employed in establishments whose data are not available for the reason just given.
Firms, 389,718 Matches, and 1,778,490 Person-years (including years spent in non-
employment).

The Gibbs sampler described in Equations (36) through (40), was used to fit the structural model to the sample LEHD data. We estimate the model using 10 support points for each distribution of latent heterogeneity. We used a model selection procedure designed to bring the variance of the structural residual close to the variance of the AKM residual after removing the match effect from the AKM residual. That is, we sought a model with granularity sufficient to have as much explanatory power in the wage equation as the unrestricted AKM model. As we see in the results, we were able to get a good match on residual variation. In practice, our model may include more support points for the employer heterogeneity than needed. Only three estimated points of support in the firm type distribution have non-negligible mass. In this case, the lack of parsimony should not result in meaningful problems with our results, since we can just collapse redundant classes in our over-parameterized model.

Our reported results are based on 2,000 draws from the Gibbs sampler. The draws were taken in approximately equal proportion from three parallel runs of the sampler. We initialized the sampler using starting values based on the AKM estimates. The sampler appears to converge after around 300 iterations, but displays extensive within-chain autocorrelation. Our Monte Carlo Standard Error procedure, described in Section 4.6 indicates that as a practical matter, the autocorrelation within threads does not much affect our results. procedures. Because of the complexity of the model and the limitations of the server that was used for estimation, we do not have the luxury of selecting every 1,000th sample to analyze. Consequently, we use a Monte Carlo standard error procedure, described in Section 4.6 indicates that properly accounts for the serial correlation in the samples we used.
5.3 Wage and Mobility Results

Figure 1 depicts the posterior distribution of the structural wage equation parameters. All of these posterior distributions are tight around the modal value. The figures plot the posterior mean together with the 5th and 95th percentile of the posterior distribution (left column) and the posterior mean \( \pm 2 \times MCSE \). For consistency with the conventional AKM decomposition, we express the wage parameters as deviations from the grand mean.

Regarding the wage parameters, note that there is variation in the estimated earnings parameters on all three heterogeneity dimensions. The dispersion on the match effects is much greater than that of the person effects \((\theta)\) or employer effects \((\psi)\). Also, there is very little variation in the estimated employer effects between classes 3 and 8. As we will see, our model only detects four distinct employer classes.

Figures 2–4 reports the probability for worker type, \( \pi_A \), the probability for employer type, \( \pi_B \), and the marginal probability for match type, \( \pi_K \). The latter probability is computed by integrating the conditional probability, \( \pi_k|ab \) over worker and employer types. All worker types occur in the population with positive probability, however, workers are more likely to be of the highest and lowest type in the population. The case is even more extreme for employers. The distribution of employer types has only three points of support with non-negligible mass. The distribution of employer types is thus very coarse. By contrast, the distribution of match classes is the most granular, though non-uniform. The marginal distribution of match types is skewed toward higher match types.

Table 1 compares the wage decomposition parameters estimated by least squares (labeled AKM) and our endogenous mobility model (labeled Gibbs). In computing this table, we used the continous values from the AKM decomposition but, of necessity, only the discrete values from our structural model. The structural person effect, \( \theta_{Gibbs} \), and firm effect, \( \psi_{Gibbs} \), explain slightly less of the variance of the dependent variable (labeled \( y \)) than their AKM counterparts. The structural match component, \( \mu_{Gibbs} \), explains less of the variation in the dependent variable.
than its AKM counterpart. In the AKM estimates, the person and firm effect have a positive correlation of 0.1665. In the structural estimation the correlation is much weaker, at 0.0312. The AKM match effect, because it is estimated from the least squares residual must be essentially uncorrelated with the person and firm effects, as is the case in our subsample. However, the structural person and firm effects are weakly (person) and strongly (firm) negatively correlated with the structural match effect.

The bottom-left panel of Table 1 presents the correlation between the inconsistent AKM parameters and the structural parameters. Three key facts appear in the table. First, there is a positive correlation between each structural parameter and its AKM counterpart. Second, the structural worker and firm effects apparently combine information from the AKM person, firm, and match effects. Third, the structural match effect has a small negative correlation with the AKM firm effect, but a positive correlation with the AKM worker and match effects.

Table 2 displays the same information in the form of a regression of the structural estimates of the wage decomposition components on the AKM estimates of all components. These regressions, therefore, compute the conditional expectation of the structure given the AKM estimates. They can be use to compute endogenous mobility-corrected estimates of the wage components from data for which only the AKM estimates are available. Of course, more specification testing should be performed to confirm that the other samples share the same mobility parameters as the one we used to estimate this table.

Next, we discuss our estimates of the mobility model. We summarize the mobility model by presenting the stationary distribution of employment spells across firm-match types. We obtain the steady-state distribution by computing the kernel of the Markov transition matrix implied by the mobility model from the estimated parameters, $\gamma, \delta, \pi_A, \pi_B$ and $\pi_K|AB$. This gives us, for each worker type, the steady-state probability of observing a worker of that type in a particular type of match. We report the steady-state distribution implied by our theoretical model and contrast it with the steady-state distribution implied by the inconsistent AKM estimates. These summaries are depicted in Figures 5 and 6.
Figures 5 and 6 have a common, but unusual presentation. Each line in the plot represents the steady-state distribution across firm-match pairs for a given worker type class. Hence, there are 10 lines, corresponding to worker type $\ell = 1$ to $\ell = 10$. The height of each line corresponds to the probability that, conditional on sampling an employment spell, we will sample a match involving a worker of type $\ell$ in a firm of type $m$ on a match of type $q$. The firm-match pair types are arranged so that position 1–10 on the x-axis correspond to firm type 1 and match class 1–10, positions 11–20 correspond to firm type 2 and match class 1–10, and so on.

In Figure 5, we observe that both low and high wage workers are more likely to be in low wage firms in the implied steady-state, as measured by the AKM wage components. Workers in the middle of the $\theta$ distribution are slightly more likely to be in high-wage firms. Match quality has modes at two mass points, the positions of which are common across worker and firm types. That is, there is no systematic relationship between the average AKM residual and worker type or firm type in the steady-state, consistent with the evidence from Table 1.

Figure 6 tells a different story. As we already saw in Figure 3, there are effectively no firms in the middle of the firm type distribution. Unlike the AKM steady-state, we see here more of the mobility segmentation in the data. There are three firm types, and high wage workers are somewhat more likely to be observed in high wage firms. Furthermore, while it is hard to see the way the graph is drawn, in this graph, steady-state match quality is decreasing in employer type, and in worker type. That is, match quality is negatively associated with worker and employer type, just as Table 1 suggests.

Figure 8 shows the non-employment probabilities in the steady-state, conditional on the worker’s type, for AKM and our structural estimates. Both lines indicate high non-employment rates, perhaps due to the way we sampled from only three states allowing some measured “non-employment” to actually be “employment elsewhere in the U.S.” Nevertheless, in the structural estimates, it is clear that low $\theta$ worker types are more likely to experience non-employment than high $\theta$ workers. This effect is much more pronounced in the structural estimates.
than in the AKM values.

6 Conclusion

Could we randomly assign workers to jobs without changing the manner in which their wages were determined? Either their employers know of the random assignment, and would, presumably, compensate them differently than workers they hired, or they would not, in which case those workers would be non-randomly selected from the pool of potential applicants. It is easy to imagine randomizing applications but not realized assignments. We do not have an ideal experiment that identifies the effect of assignments of workers to firms. It is difficult to think of what an ideal experiment would be.

The central problem of this paper is one manifestation of a fundamental challenge of empirical social science: separating the influence of correlated unobservables and sorting from the direct effect of group membership. Exploiting the wealth of information about labor market behavior locked in the relational structure of matched data holds great potential to address these problems. We use matched data to construct a complete model for the actual assignment of workers to firms that exploits the dynamic network structure of our data.

Bibliography


Figures and Tables

Figure 1: Posterior Distribution of Wage Equation Parameters
Figure 2: Posterior Distribution of Worker Population Heterogeneity
Figure 3: Posterior Distribution of Employer Population Heterogeneity
Figure 4: Posterior Marginal Distribution of Match Population Heterogeneity
Figure 5: Steady-State Distribution Across Firm and Match Types: AKM Starting Values
Figure 6: Steady-State Distribution Across Firm and Match Types: Structural Estimates
Figure 7: Steady-State Probability of Non-Employment Conditional on Worker Class: Structural less AKM
Figure 8: Steady-State Probability of Non-Employment Conditional on Worker Class: Structural and AKM
Table 1: Correlation Matrix of Wage Equation Parameters: LEHD Data 0.5% Sample, (10, 10, 10) Model

<table>
<thead>
<tr>
<th></th>
<th>$y$</th>
<th>$X\beta_{AKM}$</th>
<th>$\theta_{AKM}$</th>
<th>$\psi_{AKM}$</th>
<th>$\mu_{AKM}$</th>
<th>$\varepsilon_{AKM}$</th>
<th>$X\beta_{Gibs}$</th>
<th>$\theta_{Gibs}$</th>
<th>$\psi_{Gibs}$</th>
<th>$\mu_{Gibs}$</th>
<th>$\varepsilon_{Gibs}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>1</td>
<td>0.4419</td>
<td>0.3866</td>
<td>0.4981</td>
<td>0.3362</td>
<td>0.2003</td>
<td>0.7759</td>
<td>0.4841</td>
<td>0.2736</td>
<td>0.0948</td>
<td>0.2686</td>
</tr>
<tr>
<td>$X\beta_{AKM}$</td>
<td></td>
<td>1</td>
<td>-0.4865</td>
<td>0.0684</td>
<td>0.0257</td>
<td>-0.0171</td>
<td>0.5498</td>
<td>0.1483</td>
<td>0.0162</td>
<td>0.0485</td>
<td>0.0022</td>
</tr>
<tr>
<td>$\theta_{AKM}$</td>
<td></td>
<td></td>
<td>1</td>
<td>0.1665</td>
<td>-0.0000</td>
<td>0.0002</td>
<td>-0.0000</td>
<td>0.0362</td>
<td>0.0162</td>
<td>-0.0362</td>
<td>0.0032</td>
</tr>
<tr>
<td>$\psi_{AKM}$</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>-0.0000</td>
<td>0.0002</td>
<td></td>
<td>0.0000</td>
<td>-0.0000</td>
<td>0.0000</td>
<td>-0.0000</td>
</tr>
<tr>
<td>$\mu_{AKM}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_{AKM}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table entries are means of the correlation between the indicated variables across 2,000 draws from the Gibbs sampler described in the text.
Table 2: Regression of Structural Wage Decomposition Components on AKM Estimates of Wage Decomposition Components: LEHD Data 0.5% Sample, (10, 10, 10) Model

<table>
<thead>
<tr>
<th></th>
<th>$X\beta_{\text{Gibbs}}$</th>
<th>$\theta_{\text{Gibbs}}$</th>
<th>$\psi_{\text{Gibbs}}$</th>
<th>$\mu_{\text{Gibbs}}$</th>
<th>$\varepsilon_{\text{Gibbs}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X\beta_{\text{AKM}}$</td>
<td>0.4101</td>
<td>0.0352</td>
<td>0.0050</td>
<td>0.4334</td>
<td>0.0047</td>
</tr>
<tr>
<td></td>
<td>(.0004)</td>
<td>(.0018)</td>
<td>(.0010)</td>
<td>(.0110)</td>
<td>(.0000)</td>
</tr>
<tr>
<td>$\theta_{\text{AKM}}$</td>
<td>0.3372</td>
<td>0.1897</td>
<td>0.0237</td>
<td>0.1551</td>
<td>0.0067</td>
</tr>
<tr>
<td></td>
<td>(.0016)</td>
<td>(.0086)</td>
<td>(.0012)</td>
<td>(.0086)</td>
<td>(.0001)</td>
</tr>
<tr>
<td>$\psi_{\text{AKM}}$</td>
<td>0.2416</td>
<td>0.2214</td>
<td>0.8496</td>
<td>.2792</td>
<td>0.0608</td>
</tr>
<tr>
<td></td>
<td>(.0006)</td>
<td>(.0097)</td>
<td>(.0277)</td>
<td>(.0280)</td>
<td>(.0006)</td>
</tr>
<tr>
<td>$\mu_{\text{AKM}}$</td>
<td>0.0384</td>
<td>-.0022</td>
<td>0.2191</td>
<td>0.6024</td>
<td>0.1423</td>
</tr>
<tr>
<td></td>
<td>(.0000)</td>
<td>(.0001)</td>
<td>(.0102)</td>
<td>(.0111)</td>
<td>(.0006)</td>
</tr>
<tr>
<td>$\varepsilon_{\text{AKM}}$</td>
<td>-.0242</td>
<td>0.0071</td>
<td>0.0032</td>
<td>0.0327</td>
<td>0.9576</td>
</tr>
<tr>
<td></td>
<td>(.0000)</td>
<td>(.0004)</td>
<td>(.0005)</td>
<td>(.0013)</td>
<td>(.0001)</td>
</tr>
<tr>
<td>Constant</td>
<td>5.9515</td>
<td>-.1436</td>
<td>0.0100</td>
<td>-3.1987</td>
<td>-.0318</td>
</tr>
<tr>
<td></td>
<td>(.0140)</td>
<td>(.0073)</td>
<td>(.0004)</td>
<td>(.0760)</td>
<td>(.0002)</td>
</tr>
</tbody>
</table>

Results from running a regression of the wage components estimated under the endogenous mobility model on wage components estimated using the AKM decomposition. The reported values are the mean parameter estimate and the correlated-draw Monte Carlo standard errors across 2,000 draws from the Gibbs sampler.