DISCUSSION PAPER SERIES

IZA DP No. 1521

Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.

Moshe Buchinsky
Denis Fougère
Francis Kramarz
Rusty Tchernis

March 2005

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.

Moshe Buchinsky
UCLA,CREST-INSEE
and NBER

Denis Fougère
CNRS, CREST-INSEE, CEPR
and IZA Bonn

Francis Kramarz
CREST-INSEE, CEPR
and IZA Bonn

Rusty Tchernis
Indiana University

Discussion Paper No. 1521
March 2005

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 1521
March 2005

ABSTRACT

Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.*

Much of the research in labor economics during the 1980s and the early 1990s was devoted to the analysis of changes in the wage structure across many of the world's economies. Only recently, has research turned to the analysis of mobility in its various guises. From the life cycle perspective, decreased wage mobility and increased job instability, makes the phenomenon of increasing wage inequality more severe than it appears to be at first sight. In general, workers' wages may change through two channels: (a) return to their firm-specific human capital (seniority); or (b) inter-firm wage mobility. Our theoretical model gives rise to three equations: (1) a participation equation; (2) a wage equation; and (3) an interfirm mobility equation. In this model the wage equation is estimated simultaneously with the two decision equations. We use the Panel Study of Income Dynamics (PSID) to estimate the model for three education groups. Our main finding is that returns to seniority are quite high for all education groups. On the other hand, the returns to experience appear to be similar to those previously found in the literature.

JEL Classification: C11, C15, J31, J63

Keywords: wage mobility, interfirm mobility, returns to seniority, panel data, Markov Chain Monte Carlo methods

Corresponding author:

Moshe Buchinsky
University of California, Los Angeles
Department of Economics
CA 90095-1477
USA
Email: buchinsky@econ.ucla.edu

* We thank Gary Chamberlain, Andrew Foster, Oded Galor, Hank Farber, Richard Freeman, Jinyong Hahn, Joe Hotz, Mark Machina, Magali Ponçon-Beffy, Jean-Marc Robin, John Rust, Bernard Salanie, Duncan Thomas, and Robert Topel for their comments and discussions. Participants of seminars and conferences in Amsterdam, Brown, Crest (Malinvaud seminar), ESEM 2001 in Lausanne, Harvard, Maryland, Michigan, Pennsylvania, Tel-Aviv, Toulouse, UCLA, UCSD, Wisconsin, and Yale, provided very helpful comments. Remaining errors are ours.

1 Introduction

Much of the research in labor economics during the 1980s and the early 1990s was devoted to the analysis of changes in the wage structure across many of the world's economies. In particular, wage inequality has been one of the prime topics of investigation. It is only recently that research has turned to the analysis of mobility in its various guises. A large share of this recent effort has linked mobility with instability (see Farber (1999) for a detailed assessment), while a smaller fraction has been devoted to the analysis of mobility of individuals through the wage distribution (see Buchinsky and Hunt (1999)). The shift in focus is not surprising, and stems from the fact that measures of inequality alone are not sufficient to assess changes in the wage determination process.

While wage inequality increased in the United States during most of the 1980s, in European countries it was generally stable. Nevertheless, during the same period almost all countries witnessed a sharp decrease in wage mobility.[1] Furthermore, workers are more likely to be in a worse situation if there is an increase in the instability of jobs, as has been documented in the United States for males (see, again, Farber (1999), for a discussion of the evidence). From the life cycle perspective, decreased wage mobility and increased job instability, makes the phenomenon of increasing wage inequality more severe than it appears to be at first sight. Specifically, increased inequality (as in the U.S.), or a high unemployment rate (as in France) raises a lot more concerns on part of the policy makers than if mobility through the distribution were relatively high.

In general, workers' wages may change through two channels. Workers can stay in the same firm for some years and collect the return to their firm-specific human capital (seniority). Alternatively, they can switch to a different employer if their outside wage offer exceeds that of their current employer or when they become unemployed. These two possibilities can be empirically investigated. If Topel (1991) is right, then the first scenario provides a more plausible explanation for understanding wage increases. However, if Altonji and Williams (1997) are correct, then interfirm mobility is necessary for wage increases to occur. Some comparative results (which do not take selection biases into account) seem to show that interfirm mobility is associated with larger absolute changes in wages (e.g. Abowd, Finer, Kramarz, and Roux (1997)), but there are also considerable variations in the returns to seniority across firms (e.g. Abowd, Finer, and

[1] See, for example, Buchinsky and Hunt (1999) for the U.S. and Buchinsky, Fougère, and Kramarz (1998) for France.

Kramarz (1999) for the U.S., and Abowd, Kramarz, and Margolis (1999) for France).

The analysis of these two channels constitutes the prime motivation for this study, which lies at the intersection of two classical fields of labor economics: (a) the analysis of interfirm and wage mobility; and (b) the analysis of returns to seniority. Our theoretical model gives rise to three equations: (1) a participation equation; (2) a wage equation; and (3) an interfirm mobility equation. In this model the wage equation is estimated simultaneously with the two decision equations, namely the decision to participate in the labor force and the decision to move to a new firm. Each equation includes a person-specific effect and an idiosyncratic component. The participation and mobility equations also include lagged decisions as explanatory variables.

We use the Panel Study of Income Dynamics (PSID) to estimate the model for three education groups: (1) high school dropouts; (2) high school graduates with some post-high school education; and (3) college graduates. We adopt a Bayesian approach and employ methods of Markov Chain Monte Carlo (MCMC) for estimating the joint posterior distribution of the model's parameters.

Our main finding is that returns to seniority are quite high for all education groups. The returns to experience appear to be at approximately the level as previously found in the literature. While we use a somewhat different sample than the one used by Topel (1991), the results we obtain for the returns to seniority are, qualitatively, similar to Topel's results. Consequently, our estimate of total within-job growth are somewhat higher than Topel's estimate, and other analyses reported in the literature (e.g. Altonji and Shakotko (1987), Abraham and Farber (1987), and Altonji and Williams (1997)). Our statistical assumptions—by endogenizing both mobility and experience—incorporate elements on both sides, i.e., Topel (1991) and Altonji and Williams (1997), in order to include their main intuitions in a unified setting. Other close papers in the literature are Dustmann and Meghir (2005) and Neal (1995) who analyzed similar questions but took different routes.

In contrast to all previous studies, we explicitly model the participation and the mobility decisions. We consequently show that the differences between our estimates and those obtained in previous studies are direct consequences of two important factors. The first factor stems from insufficient modelling of the unobserved heterogeneity in the various decisions. This factor has some important implications on the implied choices of career paths.

The other factor stems from the direct modeling of possible discontinuous jumps in individuals' wages when they change jobs. We explicitly allow these jumps to depend on the level of seniority and labor market experience at the point in time when an individual changes jobs.[2]

The remainder of the paper is organized as follows. In Section 2, we outline the model. Section 3 presents the econometric specifications. Here we also introduce the likelihood function, which makes it clear why the usual ("frequentist") maximization routines are virtually impossible to implement. In Section 4 we provide a brief description of the numerical techniques used for computing the posterior distribution of the model's parameters. In Section 5 we provide a review of the literature on the return to seniority, and highlight the differences among the various studies. A brief discussion of the data extract used in this study is provided in Section 6. Section 7 presents the empirical results of the estimation procedure. Section 8 follows with a conclusion. There are three appendices. In Appendix A and Appendix B we provide some details and proofs that relate to the model being introduced in Section 2. Appendix C provides details about sampling from the posterior distribution of the parameters.

2 The Model

In this section, we consider a theoretical model that generates first-order state dependence for participation and mobility processes under some sufficient conditions that are stated explicitly below. This model is an extension of Hyslop's (1999) dynamic programming model of search behavior, in that we allow a worker to move directly from one job to another. Specifically, a worker who is currently employed in a firm receives, without searching, two wage offers in each period: one from his/her current employer, while the other comes from an outside firm.[3] If interfirm mobility occurs at the end of period $t-1$ the worker incurs the cost c_M, paid at the beginning of period t.

At any given period t a non-participant may search for a job at a cost γ_1 per period, paid at the beginning of the next period. This search cost is assumed to be strictly lower than the mobility cost c_M. We assume that hours of work are constant across jobs, so that we can concentrate only on the extensive margin of the participation process y_t, which takes the value 1 if the individual participates in period t, and takes the value 0 otherwise.

Each individual maximizes the discounted present value of the infinite lifetime intertempo-

[2] Below we describe this structure in details and explain the iclusion of what we refere to as the J^W function, when describing the specification of the wage function.

[3] While one can also incorporate cost assoiciated with search for an outside wage offer, say γ_2, the main results presented here do not change. To simplify the presentation we assume, therefore, that γ_2=0.

rally separable utility function given by

$$U_t = \sum_{s=0}^{\infty} \beta^s E_t \left[u \left(C_{t+s}, y_{t+s}; X_{t+s} \right) \right], \tag{1}$$

where $u(\cdot)$, the current period flow utility, is defined over consumption C_t and leisure $l_t = 1 - y_t$., conditional on a vector exogenous (observed and unobserved) individual characteristics X_t. The term β is simply the discount factor. The expectation E_t is taken conditional on the information available at time t. Assuming neither borrowing nor lending, the utility in (1) is maximized subject to the period-by-period budget constraint given by

$$C_t = z_t + w_t y_t - \gamma_1 \left(1 - y_{t-1} \right) - c_M \left(y_{t-1} m_{t-1} \right), \tag{2}$$

where the price of consumption in each period is normalized to 1, z_t is non-labor income, w_t is the individual's wage, m_{t-1} is a dummy variable that takes the value 1 if the individual moved to a new job at the end of period $t-1$, and takes the value 0 otherwise. It is important to note that, in this framework, first-order state dependence of participation and mobility processes relies crucially on the assumption that the cost of moving at the end of period $t-1$ is paid at the beginning of period t.

By virtue of Bellman's optimality principle, the value function at the beginning of period t given past participation y_{t-1}, and past mobility m_{t-1}, is given by

$$V_t \left(y_{t-1}, m_{t-1}; X_t \right) = \max_{y_t, m_t} \left\{ u \left(C_t, y_t; X_t \right) + \beta E_t \left[V_{t+1} \left(y_t, m_t; X_{t+1} \right) \right] \right\}. \tag{3}$$

If the individual does not participate in period $t-1$, namely if $y_{t-1} = 0$ (and obviously $m_{t-1} = 0$), the value function is given by

$$\begin{aligned}
V_t \left(0, 0; X_t \right) &= \max_{\{y_t, m_t\}} \left[V_t^0 \left(0, 0; X_t \right), V_t^1 \left(0, 0; X_t \right), V_t^2 \left(0, 0; X_t \right) \right], \quad \text{where} \qquad (4) \\
V_t^0 \left(0, 0; X_t \right) &= u \left(z_t - \gamma_1, 0; X_t \right) + \beta E_t \left[V_{t+1} \left(0, 0; X_{t+1} \right) \right], \\
V_t^1 \left(0, 0; X_t \right) &= u \left(z_t + w_t - \gamma_1, 1; X_t \right) + \beta E_t \left[V_{t+1} \left(1, 0; X_{t+1} \right) \right], \quad \text{and} \\
V_t^2 \left(0, 0; X_t \right) &= u \left(z_t + w_t - \gamma_1, 1; X_t \right) + \beta E_t \left[V_{t+1} \left(1, 1; X_{t+1} \right) \right],
\end{aligned}$$

where $V_t^0 \left(0, 0; X_t \right)$ denotes the value of non-participation in period t, $V_t^1 \left(0, 0; X_t \right)$ denotes the value of participating without moving at the end of period t, and $V_t^2 \left(0, 0; X_t \right)$ denotes the value of participating and moving at the end of period t.

For a "stayer", i.e., a participant who stays in his/her job at the end of period $t-1$, the value function at the beginning of period t is

$$\begin{aligned} V_t(1,0;X_t) &= \max_{y_t,m_t}\left\{V_t^0(1,0;X_t), V_t^1(1,0;X_t), V_t^2(1,0;X_t)\right\}, \quad \text{where} \qquad (5)\\ V_t^0(1,0;X_t) &= u(z_t,0;X_t) + \beta E_t V_{t+1}[(0,0;X_{t+1})],\\ V_t^1(1,0;X_t) &= u(z_t+w_t,1;X_t) + \beta E_t[V_{t+1}(1,0;X_{t+1})], \quad \text{and}\\ V_t^2(1,0;X_t) &= u(z_t+w_t,1;X_t) + \beta E_t[V_{t+1}(1,1;X_{t+1})]. \end{aligned}$$

Similarly, for a "mover", i.e., a participant who moves to another firm at the end of period $t-1$, the value function at the beginning of period t is

$$\begin{aligned} V_t(1,1;X_t) &= \max_{y_t,m_t}\left\{V_t^0(1,1;X_t), V_t^1(1,1;X_t), V_t^2(1,1;X_t)\right\}, \quad \text{where} \qquad (6)\\ V_t^0(1,1;X_t) &= u(z_t-c_M,0;X_t) + \beta E_t[V_{t+1}(0,0;X_{t+1})],\\ V_t^1(1,1;X_t) &= u(z_t+w_t-c_M,1;X_t) + \beta E_t[V_{t+1}(1,0;X_{t+1})], \quad \text{and}\\ V_t^2(1,1;X_t) &= u(z_t+w_t-c_M,1;X_t) + \beta E_t[V_{t+1}(1,1;X_{t+1})]. \end{aligned}$$

Transitions to a non-participation state will occur if the wage offer in period t is less than the minimum of two reservation wages, namely the wage levels that equate the value function of non-participation state with that in the participation states, with and without interfirm mobility. These two reservation wages, denoted by $w^*_{01,t}$ and $w^*_{02,t}$, respectively, are defined implicitly by

$$V_t^0(0,0;X_t) = V_t^1(0,0;X_t \mid w^*_{01}(t)) = V_t^2(0,0;X_t \mid w^*_{02}(t)). \qquad (7)$$

It follows immediately from (7) that

$$w^*_{02,t} \lesseqgtr w^*_{01,t} \quad \Leftrightarrow \quad E_t[V_{t+1}(1,0;X_{t+1})] \lesseqgtr E_t[V_{t+1}(1,1;X_{t+1})].$$

Note that if $w^*_{02,t} < w^*_{01,t}$, then the decision rule for a non-participant is to accept any wage offer greater than $w^*_{02,t}$, and to move to another firm in the next period. If $w^*_{01,t} < w^*_{02,t}$, then the optimal strategy for a non-participant is to accept any wage offer greater than $w^*_{01,t}$, and to stay in the firm for at least one more period.[4]

[4] This result is similar to the main finding obtained by Burdett (1978). In particular, see (Burdett, 1978, Propositions 1 and 2, p. 215).

The reservation wages for a stayer, denoted by $w_{11,t}^{*}$ and $w_{12,t}^{*}$, respectively, are defined implicitly by

$$V_t^0\left(1,0;X_t\right)=V_t^1\left(1,0;X_t\mid w_{11}^{*}\left(t\right)\right)=V_t^2\left(1,0;X_t\mid w_{12}^{*}\left(t\right)\right). \tag{8}$$

For a mover, these reservation wages, denoted $w_{21,t}^{*}$ and $w_{22,t}^{*}$, respectively, are defined by

$$V_t^0\left(1,1;X_t\right)=V_t^1\left(1,1;X_t\mid w_{21}^{*}\left(t\right)\right)=V_t^2\left(1,1;X_t\mid w_{22}^{*}\left(t\right)\right). \tag{9}$$

Mathematical relationships between these different reservation wages are explicited in Appendix A. We show in Appendix B that the participation and mobility equations exhibit first-order state dependence under the following sufficient conditions:

Condition 1: *When $w_{02,t}^{*}<w_{01,t}^{*}$ the participation and mobility equations exhibit first-order state dependence if, at any level of consumption, the marginal utility of consumption is either greater or lower when working, and if:*

$$c_M>\gamma_1\left[\frac{u'\left(z_t+w_{02,t}^{*},1;X_t\right)-u'\left(z_t,0;X_t\right)}{u'\left(z_t+w_{22,t}^{*},1;X_t\right)-u'\left(z_t,0;X_t\right)}\right] \tag{10}$$

where $u'(\cdot)$ denotes the marginal utility of consumption.

Condition 2: *When $w_{01,t}^{*}<w_{02,t}^{*}$ the participation equation exhibits first-order state dependence if, at any level of consumption, the marginal utility of consumption is either greater or lower when working, and if:*

$$c_M>\gamma_1\left[\frac{u'\left(z_t+w_{01,t}^{*},1;X_t\right)-u'\left(z_t,0;X_t\right)}{u'\left(z_t+w_{21,t}^{*},1;X_t\right)-u'\left(z_t,0;X_t\right)}\right]. \tag{11}$$

In that case, there is no inter-firm mobility: a non-participant moves to employment (respectively, stays in the non-participation state) at the end of period t if he or she is offered a wage greater (respectively, lower) than $w_{01,t}^{*}$. A participant becomes non-participant (respectively, remains employed) if he or she is offered a wage less (respectively, higher) than $w_{11,t}^{*}$.

We show in Appendix B that ratios appearing in expressions (10) and (11) are greater than 1. Hence, both conditions (10) and (11) are stronger assumptions than $c_M>\gamma_1$. This means that the mobility cost c_M, which is incurred at the next period after moving, has to be relatively high in order to generate first-order state dependence of both the participation and mobility processes. In the case where inter-firm mobility may happen, namely when $w_{02,t}^{*}<w_{01,t}^{*}$, we

also show in Appendix B that a mover at the end of period $t-1$ becomes non-participant at the end of period t if he/she is offered a wage less than $w^*_{22,t}$. A stayer becomes a non-participant if he/she is offered a wage less than $w^*_{12,t}$. Thus, the participation decision at period t can be characterized by

$$\begin{aligned} y_t &= \mathbf{1}\left[w_t > w^*_{02,t} - \gamma_{12}\ y_{t-1} + \gamma_{22}\ y_{t-1}\ m_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{02,t} + \gamma_{12}\ y_{t-1} - \gamma_{22}\ y_{t-1}\ m_{t-1} > 0\right], \end{aligned} \tag{12}$$

with

$$\begin{aligned} \gamma_{12} &= \gamma_1 \left[1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)}\right], \\ \gamma_{22} &= c_M \left[\frac{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)}\right], \end{aligned}$$

and where $\mathbf{1}(\cdot)$ denotes the usual indicator function.

A stayer (at $t-1$) may decide to move to another firm at the end of period t. This will happen if he/she is offered a wage w_t greater than $w^*_{12,t}$, but less than $w^*_{22,t}$. On the other hand, a mover (at $t-1$) will decide to move again only if the wage offer at t is at least $w^*_{22,t}$. Consequently, the mobility decision at period t can be characterized by

$$\begin{aligned} m_t &= \mathbf{1}\left[w_t > w^*_{12,t} + \gamma_{22}\ y_{t-1}\ m_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{12,t} - \gamma_{22}\ y_{t-1}\ m_{t-1} > 0\right] \end{aligned} \tag{13}$$

Substitution of the wage function, i.e. the equation for w_t into (12) and (13), gives the equations which are the basis for our econometric specification as detailed below.

3 The Econometric Specification

The econometric specification allows for elements that are key to labor markets decisions and wage setting outcomes. Motivated by the model of Section 2, we have three equations. The first two equations, namely the participation and mobility equations, were discussed in the previous section. The third equation is a log wage equation specifying the individuals' annual earnings function.[5]

[5] A similar model, but without the mobility equation, was also considered by Kyriazidou (1999).

In the first two equations we distinguish between periods for $t > 1$ and period $t = 1$, for which we need to specify some initial conditions as described below.

The participation equation for date t, $t > 1$, is given by

$$\begin{aligned} y_{it} &= \mathbf{1}(y_{it}^* \geq 0), \\ y_{it}^* &= x_{yit}' \beta_0 + \beta_y y_{i,t-1} + \beta_m m_{i,t-1} + \alpha_{yi} + u_{it}, \end{aligned} \tag{14}$$

where $\mathbf{1}(\cdot)$ is the usual indicator function, y_{it}^* denotes a latent variable that depends on: (a) the individual observable characteristics at time t, x_{yit} (which include, among other things, education and actual lagged labor market experience);[6] and (b) past realizations of the participation and the mobility decisions. The term α_{yi} is a person specific random effect, while u_{it} is a contemporaneous error term.[7]

The interfirm mobility equation at any date t, $t > 1$, is given by

$$\begin{aligned} m_{it} &= \mathbf{1}(m_{it}^* \geq 0) \times \mathbf{1}(y_{i,t-1} = 1, y_{it} = 1), \\ m_{it}^* &= x_{mit}' \lambda_0 + \lambda_m m_{i,t-1} + \alpha_{mi} + v_{it}, \end{aligned} \tag{15}$$

where m_{it}^* denotes a latent variable that depends on x_{mit}, the observable characteristics for the ith individual at time t. Among other things x_{mit} (which need not be the same as x_{yit} in (14)) includes education, lagged labor market experience (and its square), and lagged seniority (or tenure) in the firm where he/she is employed (and its square).[8] The term α_{mi} is a person specific random effect, while v_{it} is a contemporaneous error term. An obvious implication of the above specifications, in (14) and (15), is that, by definition, one cannot move at date t unless he/she participated at both $t - 1$ and t.

The (log) wage equation for individual i at all dates t, is specified as follows:

$$\begin{aligned} w_{it} &= w_{it}^* \times \mathbf{1}(y_{it} = 1), \\ w_{it}^* &= J_{it}^W + x_{wit}' \delta_0 + \alpha_{wi} + \xi_{it}, \end{aligned} \tag{16}$$

where w_{it}^* denotes a latent variable that depends on observable characteristics x_{wit}. Among other things x_{wit} includes education, labor market experience (and its square), seniority (or tenure) in the firm where he/she is employed (and its square). The term α_{wi} is a person

[6] Note that the labor market experience is the sum of the individual sequence of y_{it}.

[7] As is common in the literature, we make no distinction in this specification between unemployment and non-participation in the labor force.

[8] Note that seniority is the sum of the individual sequence of m_{it} with his/her current employer.

specific random effect, while ξ_{it} is a contemporaneous error term. Finally, the term J_{it}^W denotes the sum of all wage changes that resulted from the moves until date t. We include this term to allow for a discontinuous jump in one's wage when he/she changes jobs. The jumps are allowed to differ depending on the level of seniority and total labor market experience at the point in time when the individual changes jobs. Specifically,

$$J_{it}^W = (\phi_0^s + \phi_0^e e_{i0})\, d_{i1} + \sum_{l=1}^{M_{it}} \left[\sum_{j=1}^{4} \left(\phi_{j0} + \phi_j^s s_{t_l - 1} + \phi_j^e e_{t_l - 1} \right) d_{jit_l} \right]. \tag{17}$$

Suppressing the i subscript, the variable d_{1t_l} equals 1 if the lth job lasted less than a year, and equals 0 otherwise. Similarly, $d_{2t_l} = 1$ if the lth job lasted between 2 and 5 years, and equals 0 otherwise, $d_{3t_l} = 1$ if the lth job lasted between 6 and 10 years, and equals 0 otherwise, $d_{4t_l} = 1$ if the lth job lasted more than 10 years and equals 0 otherwise. The quantity M_{it} denotes the number of job changes by the ith individual, up to time t (not including the individual's first sample year). If an individual changed jobs in his/her first sample then $d_{i1} = 1$, and $d_{i1} = 0$ otherwise. The quantities e_t and s_t denote the experience and seniority in year t, respectively.[9]

Note that from the previous section, it follows that all the variables that determine the wage function also affect the participation and mobility decisions, and all the variables that affect participation also affect the mobility decision. Consequently, both the participation and mobility equations need to include the J^W function. However, to simplify matters we assume that the J^W function is a linear combination of the (observed and unobserved) variables included in the participation equation. Hence, the J^W function appears only in the the wage equation.

It is important to note that our model comprises a unique outcome equation, i.e., the same wage equation holds for those who move and those who do not move. The only difference in the observed wage come from differences in observables, and specifically the seniority in the firm. Using the conceptual framework of the evaluation literature, one cannot compute the expected wage of the treated (e.g., those who move) had they not been treated (i.e., had they stayed with their previous employer). Indeed, had we had two equations, as in Robinson (1989) for the union and non-union sector, the construction of a counterfactual wage would have been

[9] This specification for the term J_{it}^W produces thirteen different regressors in the wage equation (16). These regressors are: a dummy for job change in year 1, experience in year 0, the numbers of switches of jobs that lasted less up to one year, between 2 and 5 years, between 6 and 10 years, or more than 10 years, seniority at last job change that lasted between 2 and 5 years, between 6 and 10 years, or more than 10 years, and experience at last job change that lasted less than one year, between 2 and 5 years, between 6 and 10 years, or more than 10 years.

possible. However, the specification of two wage equations, one for the origin firm and one for the destination firm, would be logically inconsistent, since an origin firm for one worker is the destination firm of others. Hence, one must have a unique distribution from which shocks should be drawn, i.e., a unique idiosyncratic random term in the wage equation.[10]

Initial Conditions:

The likelihood function for the ith individual, conditional on the individual's specific effects $\alpha_i = (\alpha_{yi}, \alpha_{mi}, \alpha_{wi})'$ and the exogenous variables $x_{it} = (x'_{yit}, x'_{mit}, x'_{wit})'$ is given by

$$l\left\{(y_{it}, m_{it}, w_{it})_{t=1,\ldots,T} \, |\alpha_i, x_{it}\right\} = \prod_{t=2}^{T} l\left\{(y_{it}, m_{it}, w_{it}) \, |\alpha_i, x_{it}, y_{it-1}, m_{it-1}, J_{it}^{W}\right\} \times l\left\{w_{i1}|\alpha_i, x_{i1}, y_{i1}, m_{i1}\right\} l\left\{y_{i1}, m_{i1}\right\}, \tag{18}$$

where the last term of the right hand side of (18) is the likelihood of (y_{i1}, m_{i1}) for the initial state at time $t = 1$. Following Heckman (1981), we approximate this part of the likelihood by a probit specification given by

$$\begin{aligned} y_{i1} &= 1(y_{i1}^{*} \geq 0), \quad \text{where } y_{i1}^{*} = a x_{yi1} + \delta_y \alpha_{yi} + u_{i1}, \text{ and} \\ m_{i1} &= 1(m_{i1}^{*} \geq 0) \times 1(y_{i1} = 1), \quad \text{where } m_{i1}^{*} = b x_{mi1} + \delta_m \alpha_{mi} + v_{i1}, \end{aligned} \tag{19}$$

and α_{yi} and α_{mi} are the individual specific effects in the participation and mobility equations defined in (14) and (15), respectively. That is, δ_y and δ_m are allowed to differ from one.

Thus the individual likelihood function, integrated over the individual specific effects is given by

$$\begin{aligned} l\left\{(y_{it}, m_{it}, w_{it})_{t=1,\ldots T} \, |x_{it}\right\} &= \int \left[\prod_{t=2}^{T} l\left\{(y_{it}, m_{it}, w_{it}) \, |\alpha_i, x_{it}, y_{it-1}, m_{i,t-1}, J_{i,t-1}^{W}\right\}\right] \\ &\quad \times l\left\{w_{i1}|y_{i1}, m_{i1}, \alpha_i, x_{it},\right\} \times l\left\{y_{i1}, m_{i1}\right\} f_{\alpha}\left(\alpha_i\right) d\alpha_i, \end{aligned}$$

where $f_{\alpha}\left(\cdot|x_{it}\right)$ denotes the pdf of α_i, conditional on x_{it}.

In the analysis reported below, we adopt a Bayesian approach whereby we computed the conditional posterior distribution of the parameters, conditional on the data, using Markov Chain Monte Carlo (MCMC) methods as explained below.[11]

[10] An alternative solution would be to draw one shock for each firm. But, in the data set used here this is infeasible, since individual firms cannot be identified.

[11] One can also use an alternative ("frequentist") approach such as Simulated Maximum Likelihood (SML) method (see, for example, Gouriéroux and Monfort (1996), McFadden (1989), and Pakes and Pollard (1989) for an excellent presentation of this type of methodology). However, the maximization is rather complicated and highly time consuming. For comparison we estimated the model using the SML method only for one group (the smallest one).

Stochastic Assumptions:

In this subsection, we specify the stochastic structure of the random terms in equations (14)–(16) and provide the distributional assumptions for the random terms.

First, the individual specific effects are stochastically independent of the idiosyncratic shocks, that is $\alpha_i \perp (u_{it}, v_{it}, \xi_{it})$. Furthermore, we assume that α_i is a vector correlated individual specific effects, with

$$\begin{aligned} \alpha_i &\sim N(0, \Gamma_i), \quad \text{where} && (20) \\ \Gamma_i &= D_i \Delta_\rho D_i, \\ D_i &= \operatorname{diag}(\sigma_{yi}, \sigma_{mi}, \sigma_{wi}), \quad \text{and} \\ \{\Delta_\rho\}_{j,l} &= \rho_{\alpha_{ji}\alpha_{li}}, \quad \text{for } j, l = y, m, w. \end{aligned}$$

Note that the matrix Γ_i is indexed by i, since we allow σ_{ji} to be heteroskedastic, i.e., to depend on x_{yit}, x_{mit}, and x_{wit}, respectively. That is,

$$\sigma_{ji}^2 = \exp\left(h_j\left(\gamma_j, x_{i1}, ..., x_{iT}\right)\right), \quad \text{for } j = y, m, w, \tag{21}$$

where the $h_j(\cdot)$'s are some real valued functions. Furthermore, we simplify the variances in (21) to be only a function of the average of the regressors over the sample years.[12] In generic form we have

$$h_j(\gamma_j, x_{ji1}, ..., x_{jiT}) = \bar{x}_{ji}'\gamma_j, \quad \text{for } j = y_0, m_0, y, m, w, \tag{22}$$

where $\bar{x}_{ji} = (\sum_{i=1}^{T} x_{jit})/T$, for $j = y, m, w$. For convenience we also define $\gamma = (\gamma_y', \gamma_m', \gamma_w')'$.

Finally, the idiosyncratic error components from (14), (15), and (16), i.e., $\tau_{it} = (u_{it}, v_{it}, \xi_{it})'$, are assumed to be contemporaneously correlated white noise, with

$$\tau_{it} \sim N(0, \Sigma), \quad \text{where} \tag{23}$$

$$\Sigma = \begin{pmatrix} 1 & \rho_{uv} & \rho_{u\xi}\sigma_\xi \\ \rho_{uv} & 1 & \rho_{v\xi}\sigma_\xi \\ \rho_{u\xi}\sigma_\xi & \rho_{v\xi}\sigma_\xi & \sigma_\xi^2 \end{pmatrix}, \tag{24}$$

and for identification purposes, we have normalized the variance of u_{it} and v_{it} to be $\sigma_u^2 = \sigma_v^2 = 1$.

[12] Even though it applies to the variance, this simplification is reminiscent of Mundlak (1971) where the mean of fixed effect was modelled.

It is worthwhile noting that the specification of the joint distribution of the person specific effects has direct implications for the correlation between the regressors and these random effects. To see this, consider an individual with seniority level $s_{it} = s$. Note that s_{it} can be written as

$$\begin{aligned} s_{it} &= (s_{it-1} + 1)\mathbf{1}(m_{it} = 0, y_{it} = 1), \\ &= s_{it-s}\left[\prod_{l=0}^{s}\mathbf{1}(m_{it-l} = 0, y_{it-l} = 1)\right] + \sum_{l=0}^{s}\left[\prod_{k=0}^{l}\mathbf{1}(m_{it-k} = 0, y_{it-k} = 1)\right], \end{aligned}$$

where the second equality is obtained by recursion. Since the seniority level for those who are currently employed depends on the sequence of past participation and mobility dummy variables, it must also be correlated with the person-specific effect of the wage equation α_{wi}. This, in turn, is correlated with α_{mi} and α_{yi}, the person-specific effects in the mobility and participation equations, respectively. Similarly, experience and the J^W function are also correlated with α_{wi}. Given that lagged values of participation and mobility, as well as the seniority level appear in both the participation and mobility equations, it follows that the regressors in these two equations are also correlated, albeit in a complex fashion, with the corresponding person-specific effects, namely α_{yi} and α_{mi}, respectively.

4 The Posterior Distribution

Since it is analytically impossible to compute the exact posterior distribution of the model's parameter, conditional on the observed data, our goal here is to summarize the posterior distribution of the parameters of the model using a Markov Chain Monte Carlo (MCMC) algorithm.

Let the prior density of the model's parameters be denoted by $\pi(\theta)$, where θ contains all the parameters of the model, i.e., $\theta = \{\beta, \delta_y, \delta_m, \alpha, \Sigma, \Delta_\rho, \gamma\}$, as defined in detail below. The posterior distribution of the parameters would then be:

$$\pi(\theta \mid z) \propto Pr(z \mid \theta)\pi(\theta),$$

where z denotes the observed data. This posterior density cannot be easily simulated due to the intractability of $\Pr(z \mid \theta)$. Hence, we follow Chib and Greenberg (1998), and augment the parameter space to include the vector of latent variables, $z^*_{it} = (y^*_{it}, m^*_{it}, w^*_{it})$, where y^*_{it}, m^*_{it}, and w^*_{it} are defined in (14), (15), and (16), respectively.

With this addition it is easier to implement the Gibbs sampler. The Gibbs sampler iterates through the set of the conditional distributions of z^* (conditional on θ) and θ (conditional on z^*).[13] In Appendix C we provide all relevant details on the implementation of the Gibbs sampler.

A key element for computing the posterior distribution of the parameters is the choice of the prior distributions for the various elements of the parameter space. In this study we use conjugate, but very diffused priors on all the parameters of the model, reflecting our lack of knowledge about the possible values of the parameters. In all cases we use proper priors (although very dispersed) to ensure that the posterior distribution is a proper distribution.

A comprehensive sensitivity analysis that we carried out shows that the choice of the particular prior distribution hardly affects the posterior distribution of the parameters. In fact, we re-estimated the model starting at very different points for the mean of the prior distributions, particularly for the main parameter of interests, namely the coefficients on the various experience and seniority variables. In all cases we converge to virtually the same posterior distribution of the parameters. This clearly indicates that the chosen prior distributions are not dogmatic, in the sense that they have virtually no effect on the resulting posterior distributions. In fact, while all the prior distributions for the parameters are centered around zero (except for σ_w^2, which is centered around 4), with a very large variance, the posterior distributions (as is also clear from the results provided below) are centered away from zero, and have relatively small variance. This last result stems largely from the fact that the data set used is rather large.

In all the estimation reported below we employed 25,000 repetitions after the initial number of 5,000, which were discarded.

5 Where Does the Literature Stand?

The debate on returns to seniority really started with Topel (1991) whose views stood in stark contrast with previous results, mainly those of Altonji and Shakotko (1987) and of Abraham and Farber (1987).

Starting from the evidence that the costs of displacement are strongly related to prior job tenure, Topel (1991) singles out two potential explanations. The first explanation is that wages rise with seniority, while the other explanation is that tenure merely acts as a proxy for the

[13] Recent presentation of the theory and practice of Gibbs sampling and Markov Chain Monte Carlo methods may be found in the book written by Robert and Casella (1999), and in the survey by Chib (2001). In econometrics, recent applications to panel data include the papers by Geweke and Keane (2000), Chib and Hamilton (2002) and Fougère and Kamionka (2003).

quality of the job.[14] His base equation is given by

$$y_{ijt} = X_{ijt}\beta_1 + T_{ijt}\beta_2 + \epsilon_{ijt} \tag{25}$$

where y_{ijt} is the logarithm of the wage of individual i in job j at period t, X_{ijt} denotes experience, T_{ijt} denotes seniority, and ϵ_{ijt} is a residual term. This last term can be decomposed into three components:

$$\epsilon_{ijt} = \phi_{ijt} + \mu_i + v_{ijt}. \tag{26}$$

The first component in (26) is specific to the individual-work (or job) pair, the second is tied to the individual's ability, while the last component represents market-wide random shocks and measurement error. An endogeneity problem may arise if ϕ_{ijt} is correlated with experience or tenure, that is,

$$\phi_{ijt} = X_{ijt}b_1 + T_{ijt}b_2 + u_{ijt}. \tag{27}$$

With the specification of the error term in (26) and (27), the expectation of the OLS estimate of β_2 in (25) is $E(\widehat{\beta}_2) = \beta_2 + b_2$. Topel makes a convincing argument that if $\beta_2 > 0$, that is, individuals with an extra year of tenure earn a wage which is higher, on average, by β_2, then the composition of the data is such that it implies that $b_2 < 0$. This is because the workers who stay with their current employer have offers that do not compensate for the loss of tenure, while those who move get higher than average offers, because that is what led them to move. Consequently inclusion of these marginal workers in the sample reduces the average wage of "stayers" and increases the average wage of "movers". These two effects induce a negative b_2. Furthermore, the presence of mobility costs would make b_2 even smaller. It is crucial to note that in Topel's approach, experience at the entry level is exogenous and uncorrelated with the error terms.

Using a two-step estimation procedure, Topel finds that $\widehat{\beta}_2 = .0545$ and hence 10 years of seniority imply a 28% wage increase. He also estimates $\widehat{b}_1 + \widehat{b}_2 = .0020$. Topel argues that the above results provide a lower bound for the average returns to job seniority. But, there are potential biases in the estimates of $\beta_1 + \beta_2$. First, his sample may include jobs with unusually high wage growth and the jobs that are more likely to last are jobs with a large firm-specific

[14] There is an additional possibility that was not explored by Topel, that workers with long tenures are simply more able.

growth component (potentially known to the workers). However, when he restricts the data extract to include only jobs that last longer, the resulting estimate for $\beta_1 + \beta_2$ change very little.

A second potential bias may come from the fact that more able, or more productive, workers may be less mobile. He therefore examines the maintained hypothesis that the individual's ability, say μ_i, is unrelated to tenure. If, in fact, the best people move less, then the potential bias can be seen from

$$E\widehat{\beta}_2 = \beta_2 - b_1 - \gamma_{X_0 T} \times (b_1 + b_2) - \gamma_{X_0 \mu}, \tag{28}$$

where $\gamma_{X_0 T}$ and $\gamma_{X_0 \mu}$ are the coefficients from the regressions of tenure and the individual's ability on the initial experience, respectively.

To evaluate this potential bias Topel suggests using total experience X, since it is correlated with X_0, but not with μ_i, if it is assumed that the distribution of ability is unrelated to experience. Under this assumption he obtains an instrumental variable estimator whose expectation is given by

$$E\left(\widehat{\beta}_2^{IV}\right) = \beta_2 - b_1 - \frac{\gamma_{XT}}{1-\gamma_{XT}}(b_1 + b_2).$$

Since the coefficient of the regression of T on X is equal to 0.5, the resulting estimates of the returns to seniority (the lower bound) fall to .052.

In contrast, Altonji and Shakotko (1987) (AS, hereafter) use an instrumental variables approach assuming that $\phi_{ijt} = \phi_{ij}$, i.e., the individual job term is time-invariant. Hence, since

$$\sum\nolimits_t (T_{ijt} - T_{ij})(\phi_{ij} + \mu_i) = 0,$$

the deviation of seniority from the average seniority in that job is a valid instrument in the level equation, such as (25). Topel (1991) shows that this method is, in fact, a variant of his two-step approach. The instrumental variable estimate for $\beta_1 + \beta_2$ is similar to that obtained by Topel (taking within-job differences instead of first differences). Consequently, the resulting estimate of β_1 becomes

$$E\widehat{\beta}_1^{IV} = E\widehat{\beta}_1^{2S} + \left[\frac{\gamma_{XT}}{1-\gamma_{XT}} - \gamma_{X_0 T}\right](b_1 + b_2).$$

Empirically, the coefficient from the regression of T on X_0 is $-.25$. Hence, the procedure used by AS appears to induce a downward bias. Another reason for downward bias in the estimates is severe measurement error in the tenure data used by AS.

The differences in treatment of time in the regression is yet another factor contributing to the differences in estimates due to these two approaches. While Topel uses a specific index for the aggregate changes in real wages based on the CPS, AS simply use a time trend. If there is growth of quality of jobs because of better matches across time, then this growth causes an additional downward bias in the estimated returns to tenure.

Abraham and Farber (1987) (AF, hereafter) have a different set of assumptions. They use completed tenure to proxy unobserved dimensions of persons or jobs quality. The problem is that many workers have censored spells. Moreover, they also use a log wage equation that is quadratic in experience and linear in tenure.

Recently Altonji and Williams (1997) (AW, hereafter) have set up their model following Topel's (1991) model. They thoroughly investigate the consequences of various modelling assumptions on the estimated returns to seniority. First, they specify the person specific effect μ_i as

$$\mu_i = X_{ijt}c_1 + T_{ijt}c_2 + \omega_{ijt},$$

and let

$$T_{ijt} = \mu_i d_1 + X_{ijt}d_2 + \nu_{ijt}.$$

As in both Topel (1991) and AS, they assume that $\mathrm{Cov}(\mu_i, X_{ijt}) = 0$, so that d_2 is merely the coefficient from the regression of T on X , denoted by γ_{XT}. It follows then that

$$c_1 = -\gamma_{XT}c_2 < 0, \quad \text{and } c_2 = d_1 \frac{\mathrm{Var}(\mu)}{d_1^2\mathrm{Var}(\mu) + \mathrm{Var}(\nu)} > 0.$$

Note also that in (28) the bias term is given by $\gamma_{X_0\mu} = c_2\mathrm{Var}(T \mid X)/\mathrm{Var}(T)$. Since in the data used by both Topel and AS, $\mathrm{Var}(T \mid X)/\mathrm{Var}(T) = .676$, it implies that the bias due to individual heterogeneity (i.e., μ_i) in Topel's case is about 2/3 the size of the bias in the OLS regression, from the same source. In this sense, our approach is closer in spirit to that of AW, in that we model experience and allow it to be correlated with wage through its various components, especially the person-specific effect.

The two papers differ considerably in other aspects too. Topel claims that time trends are problematic because they can be correlated with ϕ when workers have had more time to locate jobs with higher ϕ. Moreover, time may be correlated with μ because of changes in the sample. In contrast, AW claim that the covariance between t and ϕ is 0, conditional on X, or conditional on T and X, or conditional on T and X_0. While AW do not appear to be convinced that the specification of the time trend may bring about a serious problem, they do conclude that a way to circumvent the above problems is the use of t in deviation from the person's mean.[15]

AW eventually reach the conclusion that all estimators are biased upwards by job match heterogeneity (i.e., due to the correlation between T and ϕ). Accounting for this issue and correcting for the timing problem in Topel's paper reduces the estimate of the returns to seniority from .223 to .126 (when evaluated at 10 years of seniority). A crucial assumption that seems to account for a significant fraction of the differences between Topel's and AW's estimates, is the assumption about the exogeneity of X_0, reducing Topel's estimate further when instrumenting for X_0. One important conclusion is that individual heterogeneity is important in the wage growth process. Hence, a large part of the reduction in the upward bias of the estimate found by Topel's is due to a reduction in the bias from individual heterogeneity.

In a recent paper, Dustmann and Meghir (2005) (DM hereafter) allow for three different sources of returns due to accumulation of specific human capital, namely experience, sector-specific seniority, and firm-specific seniority. To estimate returns to experience they focus on the new jobs of displaced workers, assuming that such workers cannot predict closure of an establishment (more than a year in advance). Furthermore, they assume that displaced workers have preferences for work similar to those that govern the sector choice. Hence, controlling for the endogeneity of experience also controls for the endogeneity of sector tenure. Then, DM estimate two reduced-form equations, one for experience and the other for participation. The residuals from these two regressions are then plugged into the wage regression based on the first wage records after displacement. These residuals are also interacted with potential experience, general experience, and sectorial tenure. DM employ a similar procedure for estimating the returns to seniority, accounting for the selection of the individuals into staying with their current employer. Using data from Germany (IAB) and the US (NLSY79), DM find that the returns to tenure for skilled workers are large (on the order of 2% a year for the first five years), and only slightly smaller for unskilled workers. They also find evidence of heterogeneous returns to tenure. While the returns to sector tenure are small, they are statistically significant. They

[15] AW also discuss the dating of the earnings measure used in Topel (1991). The latter uses earnings and tenure at date t whereas AS and AF use tenure at t and earnings at $t-1$, since employer tenure reported in the PSID refers to date t, whereas the wage measure is annual earnings (divided by annual hours) in the previous calendar year. Hence, whenever there is a job change, the measured earnings are mixtures of the old and the new job compensation. They provide some evidence that this difference alone can significantly reduce the estimated return to seniority.

also find that the returns to experience for skilled workers are larger than those for tenure (4% a year for at least 5 years), and are somewhat smaller for the unskilled workers. Finally, they find that in both countries the between-job wage growth is larger than within-job wage growth.

All the papers discussed so far focus on returns to tenure, but they do not explicitly model the mobility decision that generates the observed seniority levels. Farber (1999) notes that this process has some specific features that must be modelled. For instance, he shows that in the first few months of a job, there is an increase in the probability of separation, while this probability decreases steadily thereafter. To fully understand this process, person heterogeneity and duration dependence must be distinguished. If only pure heterogeneity prevails, the number of jobs previously held by an individual should be a sufficient statistics for the probability of a move. Conditional on experience and the number of previous job changes, the mobility should be independent of tenure in this situation. Farber gives evidence that contradicts the simple model of pure heterogeneity, and seems to suggest that one should estimate a full structural model since completed job duration is surely jointly determined with wages.

Farber (1999) also surveys some aspects of the displaced workers. If specific capital matters, then theory implies that firms will choose to lay off less senior workers, and Farber finds strong support for this. Furthermore, he finds that job losses result in substantial permanent earning losses. Moreover, the empirical findings indicate that individuals with longer tenure lose more. As is shown by Gibbons and Katz (1991) selection of the laid off workers does not seem to provide the right sample for investigating issues related to the returns to tenure. One potential explanation is that the average tenure will be higher on those jobs with high wages because of the reduced probability of job quits. If more stable workers are also more productive, then long tenure in the past jobs commands high re-employment wage. The relation between tenure and the re-employment wage is indeed pervasive in tenure literature, but has not yet been modeled explicitly.

The J^W function we introduce in the wage equation allows us to tackle this issue, and particularly the empirical findings in the literature on displaced workers (e.g. Addison and Portugal (1989), Jacobson, LaLonde, and Sullivan (1993), and Farber (1999)). The inclusion of job market experience in previous jobs as a determinant of the earnings change at the time of a job change allows us to distinguish between displaced workers, who went through a period of non-employment after displacement, from workers who move directly from one job to another. Furthermore, the inclusion of the seniority level at the end of each of the past jobs allows us to control for the quality of the previous job matches.

Neal (1995) and Parent (1999, 2000) focus their investigation on the importance of sector- and firm-specific human capital. In the current study we do not directly address this question, largely because it is computationally infeasible at this point. This is because in order to model the first choice, an additional *sector mobility* equation should be added to our system of equations. As for the choice of firm, estimation of this model is beyond reach, even when matched employer-employee data are available.

Finally, even though they do not focus their study on the returns to seniority, nor on firm-to-firm mobility, it is worthwhile emphasizing the recent work by Geweke and Keane (2000). Their model includes two equations, one for earnings and the other for marital status. Earnings are modelled as an autoregressive process, with a person specific effect and an AR(1) process for the idiosyncratic term. The marital status depends on a latent variable that is a function of the marital status at $t-1$, as well as earnings at $t-1$. Using the estimates for this model they construct simulated data for individuals with the observed characteristics of the sampled individuals. These data are used then in examining earnings mobility (between quintiles) and the present value of lifetime earnings. Their estimation technique shares some common features with ours, and it uses the same data set.

This brief review of the literature makes clear the distinctive features of our approach. We estimate a structural model for the joint decisions of participation, firm-to-firm mobility, and earnings. This model allows for time-dependence and correlated unobserved heterogeneity in the various decisions and outcomes.

6 The Data

We follow Topel (1991) and Altonji and Williams (1997) in using the Panel Study of Income Dynamics (PSID), even though our analysis period is slightly different. The PSID is a longitudinal study of a representative sample of individuals in the U.S. and the family units in which they reside. The survey, begun in 1968, emphasizes the dynamic aspects of economic and demographic behavior, but its content is broad, including sociological and psychological measures.

Two key features give the PSID its unique analytic power: (i) individuals are followed over very long time periods in the context of their family setting; and (ii) families are tracked across generations, where interviews of multiple generations within the same families are often conducted simultaneously. Starting with a national sample of 5,000 U.S. households in 1968, the PSID has re-interviewed individuals from those households every year, whether or not they are living in the same dwelling or with the same people. While there is some attrition rate, the PSID has had significant success in its recontact efforts. Consequently, the sample size has grown somewhat in recent years.[16]

The data used in this study come from 18 waves of the PSID from 1975 to 1992. The sample is restricted to all heads of households who were interviewed for at least three years during the period from 1975 to 1992 and who were between the ages of 18 and 65 in these survey dates. We include in the analysis all the individuals, even if they reported themselves as self-employed. We also carried out some sensitivity analysis, excluding the self-employed from our sample, but the results remained virtually the same. We excluded from the extract all the observations which came from the poverty sub-sample of the PSID.

In the analysis reported below, the experience and tenure variables play a major role. Nevertheless, there are some crucial difficulties with these variables, especially with the tenure variable, that one needs to carefully address. As noted by Topel (1991), tenure on a job is often recorded in wide intervals, and a large number of observations are lost because tenure is missing. Moreover, there are a large number of inconsistencies in the data. For example, between two years of a single job, tenure falls (or rises) sometime by much more than one year. The are many years with missing tenure followed by years in which a respondent reports more than 20 years of seniority. In short there is tremendous spurious year-to-year variance in reported tenure on a given job.

Since the errors can basically determine the outcome of the analysis, we reconstructed the tenure and experience variables using the same procedure suggested by Topel (1991). Specifically, for jobs that begin in the panel, tenure is started at zero and is incremented by one for each additional year in which the person works for the same employer. This procedure seems consistent. For those jobs that started before the first year a person was in the sample a different procedure was followed. The starting tenure was inferred according to the longest sequence of consistent observations. If there was no such sequence then we started from the maximum tenure on the job, provided that the maximum was less then the age of the person minus his/her

[16] There is a large number of studies that use this survey for many different research questions. For a more detailed description of the PSID see Hill (1992).

education minus 6. If this was not the case then we started from the second largest value of recorded tenure. Once the starting point was determined, tenure was incremented by one for each additional year with the same employer. The initial experience was computed according to similar principles. Once the starting point was computed, experience was incremented by one for each year in which the person worked. Using this procedure we managed to reduce the number of inconsistencies to a minimum.

In addition to this procedure we also took some other cautionary measures. For example, we checked to see that: (i) the reported unemployment matches the change in the seniority level; and (ii) there are no peculiar changes in the reported state of residence and region of residence, etc.[17]

Summary statistics of some of the key variables in the extract used here are reported in Table 1. By the nature of the PSID data collection strategy, the average age of the sample individuals does not increase much over time. We do note that the average education rises by over half a year between 1975 and 1992, since the new individuals entering the sample have, on average, higher education that those in the older cohorts. Also, experience and seniority tend to increase over the sample from 20.8 and 5.1 years in 1975 to 24.2 and 7.3 in 1992, respectively.

AW provide a summary statistics for the samples used in the various studies (see Table 1 of AW. The experience variable used here seems to be very much in line with that reported in other studies. However, there is a substantial difference between the tenure variable use here and that used by Topel, and especially AW. The main reason for this discrepancy is the fact that we restrict the sample for individuals for whom we can obtain consistent data for three consecutive years. Since individuals with lower seniority are also those that tend to leave the PSID sample more frequently. It is important to emphasize that the most relevant comparison of the samples use in the various studies is with respect to the results they yield when using similar estimation procedure. We provide below evidence that the differences in the results are not due to selection of inherently different extract, but rather that in our approach we are able to control for the endogeneity of experience as well as that of mobility. This, in turn, leads to results that are quite different. However, the differences can be explained as is demonstrated below.

The mobility variable indicates that in each of the sample years somewhere between 6.5%

[17] The resulting program, written in Matlab, contains a few thousand of lines of code. The programs are available from the corresponding author upon request.

and 14.1% of the individuals changed jobs. The mobility is very large in the first year of the sample, mainly due to measurement error; because of the difficulty in identifying whether or not a person actually moved. This illustrates the need for treating initial conditions as separate equations.

Consistent with other data sources, the average real wage increased slightly over the sample years. This is mainly because the individuals entering the sample have wages that are increasing over time in real terms, while those who leave the sample have wages that decrease somewhat over the sample years. More importantly, note that the wage dispersion increases across years. Also consistent with other data sources, the participation rate of individuals in the sample (largely men) decreases steadily over the sample years from about 92% in 1975 to 86% in 1992. Note also that a significant fraction of the sample is non-whites. This is due to oversampling of that population in the PSID. However, since the results changes very little when we use a representative sample, we keep all the individuals satisfying the conditions specified above. Approximately 20% of the sample have children who are two years old and below, although this fraction decreases somewhat over the sample years, as does the fraction of the sample that have children who are between the ages of 3 and 5. The total number of children also declines somewhat over the sample period. This is consistent with the general finding in the literature about the decline in the size of the typical American family.

It is worthwhile noting (although for brevity we do not report here) that, as has already been documented elsewhere in the literature, the distribution of individuals across the various industries changes over the sample period. The most dramatic change was the decline in the fraction of individuals employed in the manufacturing sector, which went from 24.5% to 20.1% between 1975 and 1992.

Looking at the changes in the distribution of cohorts in our sample period, we observe that the fraction of people in the youngest cohort increases steadily, in particular between 1988 and 1990, dates between which the number of observations in our sample increases quite strongly together with the fraction of Hispanics.[18] In the meantime, the fraction of people in the oldest cohorts decreases over the sample years. This simply illustrates the importance of controlling for the cohort composition of the sample in the regression analysis, which we do in the results

[18] In fact, those in charge of the PSID made a special effort to collect information for those who left the sample in the previous years. The changes in the age and race structure are due to strong geographic mobility of these young workers.

reported below.

7 The Results

The estimation is carried out for three separate education groups. The first group includes all the individuals with less than 12 years of education, i.e., those who, at some point, dropped out of high school. The second group consists of those who are high school graduates, who may have acquired some college education or who may have earned a degree higher than high school diploma, but have not completed a four-year college. Finally, the third group consists of those that are college graduates, i.e., those who have at least 16 years of education. We refer to these three education groups as the *high school (HS) dropouts*, *high school graduates*, and *college graduates*, respectively. Below we present the results from the simultaneous estimation of participation, mobility, and wage equations, together with the initial conditions, for each group. For brevity we do not report the estimates for the initial conditions' equations.

Specification:

The participation equation includes the following variables: a constant, education, quartic in lagged labor market experience, a set of three regional dummy variables (for North East, North Central, and South, excluding the West), a dummy variable for residence in an SMSA, other family unearned income, a dummy variable for being an African American and a dummy variable for being Hispanic, county of residence unemployment rate, number of children in the family, number of children less than 2 years old, number of children between the ages of 2 and 5, a dummy variable for being married, a set of four dummy variables for the cohort of birth, namely being of age 15 or less in 1975, being of age 16 to 25, age 26 to 35, and 36 to 45. The excluded dummy variable is for those who were over 45 years old in 1975. Finally we include a full set of year dummy variables.

The mobility equation includes all the variables that are included in the participation equation, and it also includes: quartic in lagged seniority on the current job, and a set of nine industry dummy variables.

The (log) wage equation includes the following variables: a constant, education, quartic in experience, a quartic in seniority on the current job, a set of variables and dummy variables giving rise to possible discrete jumps in the wage as a result of a job mobility as explained in equation (17) above, a set of three regional dummy variables, a dummy variable for residence in an SMSA, a dummy variable for being an African American and a dummy variable for being Hispanic, a set of nine industry dummy variables, county of residence unemployment rate, a set of four cohort dummy variables, and a full set of year dummy variables. The dependent variable in this equation is the log of deflated annual wage.

Recall that the covariance matrix for the individual random effects α_i is given in (20). In order to estimate Γ_i for all individuals one needs to obtain estimates for both the elements of Δ_ρ and the coefficient vectors γ_j in (22). As explained in Appendix B below, the numerical computation of the posterior distribution of the γ_j's is difficult to obtain, especially when the γ_j's are of a high dimension. Hence, instead of using $\bar{x}_{ji}$, we only use the first three principle components of $\bar{x}_{ji}$, as well as a constant term.[19]

For brevity we omit all the time effect dummy variables from all the tables reporting the results. In all tables we report the means of the posterior marginal distributions of the relevant parameters. The medians of the posterior marginal distributions are almost indistinguishable from the means. In addition, the marginal distributions of all the parameters of the model, and all the marginal and cumulative returns to experience and seniority are very close to that of normal random variables. Therefore, it is sufficient to report the mean and the standard deviation for all the relevant posterior distributions. For that reason we also do not provide any figures for the posterior marginal distributions.

7.1 Participation and Mobility

As explained above in Section 3, in both the participation and the mobility equations we allow for both duration dependence and unobserved heterogeneity. In the participation equation we have lagged participation and mobility, while in the mobility equation we have lagged mobility. In addition, both equations include individual-specific terms that are allowed to be correlated with each other, as well as with the person-specific term in the wage equation.

Participation:

The results for the participation equation for the three education groups are presented in Table 2. It is apparent from Table 2 that the education level is almost irrelevant for the high school dropout group. In sharp contrast, the level of education is a very important factor in the participation decision for the high school graduate group, and even more so for the college

[19] The first three principle components account for over 98% of the total variance of $\bar{x}_{ji}$, so that there is almost no loss of information by doing so. On the other hand, this significantly reduces the computation time.

graduate group. These results are in complete accordance with the classical human capital theory. The high school dropouts have little general human capital, and the human capital they do have does not matter much for the type of jobs they typically obtain. In contrast, the level of education matters a lot more for those who have valuable additional general human capital, as appears to be the case for the high school graduates, and even more so for college graduates.

For labor market experience we get similar qualitative results. Experience seems to be more important for the more educated individuals. The additional on-the-job accumulation of human capital is a more important factor for the more educated individuals than the less educated individuals. One interpretation of this empirical finding is that there is a limited amount of general human capital that is accumulated by the least educated individuals. The more educated an individual, the more general human capital he/she accumulates on the job, and, in turn, the larger the loss due to forgone wage for not participating.

The results for experience are conditional on controlling for lagged participation. As predicted by the human capital theory, participation in the period preceding the current period has a positive effect on the participation probability for all three education groups, as does mobility in the preceding period. That is, individuals who participate at $t-1$ are a lot more likely to participate in period t as well. Moreover, individuals who moved from one job to another in the period $t-1$ are more likely to participate in period t. As also predicated by the human capital theory, the latter effect is stronger for the least educated individuals, for whom the switching cost tends to be the largest and the longer, on average, it takes to find a new job.

For the family variables we note that for all three groups, family unearned income has negligible effect on the participation, and the effect is not statistically significant for any of the three groups. Being married has, as expected, a positive effect on participation. Nevertheless, the effect is more significant for the lowest two educational groups. The reason for this finding is likely to be that college graduates have stronger labor market attachment, independent of their marital status. Hence, one's marital status is a good predictor for participation for the less educated, but is a poor predictor for the highly educated individuals. On the other hand, in general, the number of children in the family, at any age, seems to have little effect in all education groups.

A striking effect is the racial effect (which is measured relative to white individuals). While African-Americans are less likely to participate at all education levels, the negative effect seems to be particularly large for the high school dropouts and high school graduate groups, especially in comparison with the Hispanics in these groups. In fact, Hispanics in the lower two educational are as likely to participate as their white counterparts. For the college educated group, both African-American and Hispanics are less likely to participate than whites. However, for this particular group the effect seems to be stronger for Hispanic than for African-Americans.

Note that the cohort effects seem to be important, especially for the high school dropouts and high school graduates. The younger cohorts are somewhat more likely to participate than the older cohorts. This is especially true for the high school graduate group. For the more highly educated individuals there are very little differences among the various cohorts. This is consistent with similar finding in the immigration literature.

County unemployment rate has a negative and significant effect on the participation probability of high school dropouts and high school graduate groups, but essentially no effect on the most highly educated workers. This finding is consistent with the general finding in the literature that labor markets are defined more locally for those who are less educated individuals.

Mobility:

The results for the mobility equation for the three education groups are presented in Table 3. For brevity we omit from the table not only the dummy variables for time effects, but also the industry dummy variables.

Note first that education is completely irrelevant for the mobility decision, for all three educational groups. This is in accordance with the human capital theory, since education represents the most general form of human capital that can be carried by the workers when moving from one job to another, and hence should not have any affect on the mobility decision. Also in accordance with the human capital theory is the relative magnitude of the experience and seniority effects. The seniority effect is three to four times as large as that of experience at comparable levels of experience and seniority. While experience represents a general form of human capital, which can be moved from one job to another, seniority in the firm represents a firm-specific human capital that is lost once a person leaves the firm.

The negative coefficients on lagged mobility in all three groups indicates that individuals are not likely to move in two consecutive years. Also, this effect is stronger for the more highly educated individuals.

The average seniority in the high school dropout group is about 5.7 years. Overall, the probability of a move, when evaluated at the mean level of the regressors, and at the average level of seniority, is only 0.0463. If seniority increases by 5 years, this probability decreases to 0.0324, i.e., a decline of 30%. The probability of a move for those with 15.7 years of seniority (i.e., 10 years above average) is only 0.0278. At the entry level, the probability of a move is .1079, that is, while a worker is unlikely to move two years in a row, he/she still has significant probability of moving in the first year, which declines dramatically as tenure increases.

For the high school graduates the average seniority in the sample is only 4.9 years. The probability of a job switch for workers with high school diploma is 0.0638, when it is evaluated at the mean level of the regressors and the mean level of seniority. The probability of a move for a person with 9.9 years of seniority is reduced to only 0.0373, and for a person with 14.9 years of seniority, it decreases further to 0.0308. For the college graduate group the average seniority is 5.2. The corresponding probabilities for the college graduate group are 0.0549, 0.0371, and 0.0320, respectively. These results are also consistent with the general theory on human capital accumulation. At the same level of seniority, the less educated individuals are more likely to stay with their current employer, specifically because a larger share of their human capital is firm specific. The higher the education level the larger the part of human capital that is general, and hence can be transferred from one job to another. Consequently, we observe that the least educated individuals have on average longer tenure on their job.

An interesting finding is that except for the family unearned income, no other family variables seem to affect the likelihood of a move. And, even the income variable, although statistically significant, affects the probability of a move by only a few percentage points. It is worthwhile noting that at the same level of unearned income the least educated are affected the most. Nevertheless, since, on average, the unearned income for the less educated is lower than that for the more educated, when evaluated at the mean level of unearned income, the marginal addition to the probability of a move is about the same for all groups.

Race also does not seem to play a major role in moving decisions. The Hispanics are not statistically different from the white groups. For the African-Americans, the only noticeable difference is for the most highly educated individuals. African-American in the latter group are more likely to move than their white counterparts.

Finally, note that the there are no cohort effects on the probability of a move, that is, recent generation are not different from the older generation in their moving decisions. This lack of systematic pattern seems to hold at all education levels.

7.2 The Returns to Human Capital

The main results of this study, namely the returns to the various types of human capital, are presented in Table 4, which provides the results for the wage regression for the key variables, and most importantly for education, experience, and seniority. In Tables 5 and 6 we provide cumulative and marginal returns which are based on the joint posterior distribution of the parameters reported in Table 4. In Table 5 we provide the results for the return to experience, while in Table 6 we provide similar results for the return to seniority.

The returns to education:

First, it is worth noting that there is substantial variation in the level of education in each of the three education groups we consider here. This is what allows us to separately identify the returns to education for the three groups.

As is clearly apparent from the estimates of the returns to education reported on line (2) of Table 4, the estimates are consistent with the human capital theory, in that the mean returns to education are larger for the high school graduates than for the high school dropouts, namely 4.8% and 2.5%, respectively. A declining marginal return to education also lead to lower returns to education for the college graduates than for the high school graduates, namely 4%. However, the cumulative return for college graduates far exceeds that of the high school graduates.

The results provided here are consistent with recent findings in the literature that the returns to education are somewhat smaller when estimated within education groups, rather then over all educational groups. Yet, the estimates obtained here are even smaller than those reported in a recent survey by Card (2001). It appears to be that once one controls for the selection effects due to the participation and mobility decisions, the return to education is substantially reduced. We will return to this general issue when discussing the results regarding the parameter estimates of the stochastic elements of the model.

The returns to experience:

The results for the returns to experience are provided in Table 5, using the estimates of the experience components reported in Table 4. The results reported in Table 4 are based on the quartic model, that is, they are based on a model that includes polynomials in experience and seniority up to the fourth order. In addition to the cumulative and marginal returns that are computed based on the quartic model we also provide in Table 5 cumulative and marginal returns which are based on the quadratic model, that is, they are based on a model that includes polynomials in experience and seniority only up to the second order. It is worthwhile reminding the reader that the experience variable used here is of actual labor market experience computed based on the information provided in the data (see Section 6 for more details).

The results in Panel A of the table are largely consistent with those previously obtained in the literature, particularly those obtained by Topel (1991). They are also within the range of estimates reported by AS and AW. The return to experience implied by Topel's estimates, i.e., Topel (1991, Tables 1 and 2), of the cumulative return to experience at level of 10 years of experience is .354. The estimates of AS range between .372 and .442, while those of AW range between .310 and .374. Our estimates range between .362 for the high school dropouts and .661 for the college graduates. Since the larger fraction of the sample is comprised of high school graduates, it is reasonable to state that the overall estimate for the entire sample would be close to .402, the estimate of the cumulative return to experience for that group.

These results are quite surprising, since we use here a completely different estimation method, applied to a somewhat different time period. Nevertheless, we do find substantial differences across the various educational groups at all levels of experience. While the order of the cumulative returns is not surprising, the magnitude of the difference is quite large. Even at the entry level of 5 years of experience the cumulative return for the college graduates is almost twice as large as that for the other two groups and somewhat larger for the high school graduates than for the high school dropouts.

The results reported in Panel B of Table 5 are quite different from those obtained for the quartic model. In particular, the difference in the cumulative return between the lowest and the highest educational groups, at 10 years of experience, is much smaller than for the quartic model. Nevertheless, the magnitude of the returns are at a similar order of magnitude to that previously obtained in the literature.

In Panels C and D of Table 5 we report the results of the returns to experience when we omit the J^W function defined in (17). We do so because this is one of the most important differences between our specification of the wage function and that adopted by Topel, AS, and AW. This change seems to have tremendous effects on the cumulative returns for all educational groups. There are a number of possible reasons leading to this difference. First, it might be the case that our data extract is inherently different from those previously used in the literature. However, we show below that this is clearly not the case.

A second possibility for the differences is because of the difference in our treatment of the experience variable. Essentially AS, Topel, and AW all assume that experience is an exogenous variable. This is clearly not the case, as the participation rate is not 1, and it has deteriorated over time. Incorporating the selection into the labor force increases the estimate because we are able to correct for the fact that both the specific effects and the idiosyncratic errors from the participation equation are negatively correlated with the corresponding elements in the wage equation (see rows (2) and (5) of Table 7). Hence, not controlling for the selection into the labor force would induce a downward bias on the return to experience. Moreover, the larger the correlation between the individual specific elements (the dominating part of the stochastic term) the larger should be the bias, as is seen from comparisons of the returns to experience across the various educational groups in Panel C of Table 5.

Finally, unlike the three previous studies mentioned above we control here for the interactions among unobserved heterogeneity in the various equations . As os shown below, we found this to be an extremly important factor that one needs to control for. The correlation between the individual specific factor and the variables included in the wage regression imply that without controlling for unobserved heterogeneity the estimates of the return to experience would have a downward bias. Since this aspect has not been examined in previous studies on the return to seniority it is hard for us to verify this latter possibility. However, as we show below, using our sample with previous methods yield very similar results, indicating that if previous studies had controlled for unobserved heterogeneity, the results would have been closer to those obtained here.

The returns to seniority:

The results for the returns to seniority are provided in Table 6. As for the returns to experience, the cumulative and marginal returns to seniority are computed based on the coefficient on the various components of seniority reported in Table 4.

The results for the returns to seniority previously obtained by Topel (1991) are larger than those obtained by AS and AW, as well as other papers in the literature. The returns provided in Table 6 are even larger than those reported by Topel for all educational groups at all levels of seniority. As indicated above the mean return to seniority in our sample is somewhat lower than that of the extracts used by either Topel or AW. Therefore, one needs to be cautious in comparing the results from the two studies. Nevertheless, the results are consistent with the human capital theory. At the starting level of 2 years of seniority the returns are quite similar. As seniority increases, the cumulative return for the least educated rises by more than for the highly educated. This is because for the less educated individuals who stay in their job, and hence accumulate higher seniority, the main source of human capital is job specific. Controlling for the selection due to the (endogenous) moving decision induces larger estimates for the return to seniority.

As mentioned above, it is clear from the results that there are some differences between the returns to seniority across the three education groups, especially at different stages of their spell with a given firm. However, there is one clear consistent finding for all groups, that is, the returns are large and highly significant. In fact, we find that the entire supports of the marginal posterior distributions for the returns to seniority, are all in the positive segment of the real line.

In contrast with the quartic model, the quadratic model yields returns which are not always consistent with the theory of human capital (see Panel B). Specifically, the returns to seniority are lower for the least educated group than for the more highly educated groups. Moreover, the increase in the return to seniority at higher levels of seniority is much higher for the latter groups. Yet, the returns for the quadratic model are also significantly higher than those previously reported in the literature.

One key reason for the difference between our results and Topel's (1991) results is the treatment of experience. As mentioned above, Topel treats the experience level at new jobs as exogenous. There is strong evidence in the literature (e.g. Farber (1999)) indicating that this is not the case. In fact, the results presented above for the mobility equation indicate that experience plays a major role in the moving decision and hence is highly correlated with the level of experience in a new job. This also indicates the importance for having the J^W function in the wage equation, since it controls for all past decisions of the individuals, which happen at different levels of experience and seniority.

When we omit the first equation from our estimation and treat experience as exogenous the resulting returns to seniority are reduced somewhat, but they are still significantly higher than those obtained by Topel. If, in addition, we omit the J^W component from the wage equation the resulting estimates for the return to seniority become even lower.

In Panels C and D of Table 6 we report the returns to seniority when the J^W component is omitted from the wage equation, but we still control for the participation and mobility decisions. The resulting estimates, especially in Panel C, are remarkably close to those obtained by Topel. While the cumulative returns are lower for the more highly educated than for the less educated, the average over all educational groups is extremely close to those of Topel. However, when the J^W component is omitted from the wage equation the estimates for the returns to experience are much lower in our study than in Topel's case, as we previously discussed (see Panel C of Table 5).

7.3 Implications on Long-run Wage Growth

It is somewhat difficult to exactly assess the overall effects of continuity in the participation and mobility on wage growth. To address this issue and to be able to assess the overall contributions of experience and seniority to wage growth we follow the wage growth for workers with two alternative career paths. One group of workers works throughout the entire period in one firm, while the other works for the first four years, then changes to a new job, where the group spends the remainder of the sample period. Figure 1 depicts the results for the high school graduate group, in four graphs. In Figure 1a we provide the results for the new entrants—those individuals who started with 5 years of experience and 2 years of seniority—who did not experience any episode of unemployment. Figure 1b provides similar results for the experienced workers, namely those who started with 15 years of experience and 6 years of seniority. Figure 1c is for the same group as in Figure 1a, only that this group experiences a change of job in the fifth period. A similar graph for the experienced workers is provided in Figure 1d. In all figures we break down the overall wage growth due to accumulation of human capital into the part that comes from rising experience and the part that comes from rising seniority. In Figures 2a–2d we provide the results for the college graduate group. The pattern of the results for the high school dropouts is very similar to that of the high school graduates, so we omit them for brevity.

Note first that, consistent with the human capital theory, the wage growth earlier in their career is larger, and more significantly so, for the more highly educated individuals. For example, the wage growth for the new entrants high school graduates (Figure 1a) is 0.74 log points, while for the new entrants college graduates it is 0.94 log points, even though the college graduates start from a higher point. For the experienced workers the wage growth for the more educated is somewhat lower, 0.42 log points, compared with a growth of 0.44 log points for the high school graduates. This simply represents the fact that, as implied by the human capital theory, the wage profile for the new entrants is steeper than for the experienced workers.

An interesting finding is also the fact that the wage growth that stems from increased seniority is larger than the wage growth that stems from increased experience. For the new entrants of the college graduate group the growth due to seniority is 0.59, while the growth due to experience is only 0.35 log points. For the experienced workers of that group the wage growth due to seniority is 0.50, while the growth due to experience is actually negative, -0.08. The differences for the high school graduates are even larger. For the new entrants an increase of 0.59 log points is due to seniority, while only 0.16 is due to experience. That is, firm-specific human capital is a lot more important for the less educated individuals. For the experienced workers of that group the increase in wage due to seniority is 0.48 log points while there is a decrease of -0.04 log points due to increased experience.

Nevertheless, the general patterns of wage growth across the two groups are similar, and particularly the relationship between the new entrants and experienced workers in each of these groups.

Since the gain in wages due to seniority is larger than that due to experience, there is a substantial loss in the earnings capacity for individuals who change jobs, due to the loss of firm specific skills. However, most of this loss is compensated by an increase in wage when an individual switches jobs. The overall effect for the high school graduate is such that the wage due to seniority remain constant in the year an individual changes jobs (see Figures 1c and 1d). For the college graduate, on the other hand, there is a substantial increase in wage due to seniority, because individuals with higher seniority experience, on average, larger wage increases when changing jobs. This can be viewed as the compensation offered by the firm they move to, due to the loss of firm specific skills they incur when leaving their previous job.

For both education groups we see that the wage growth due to seniority is larger for the individuals who change jobs than for the individuals who stay in one firm for the whole period. For the high school graduate who stays in the same firm throughout the period, the growth in wage due to seniority is from .13 to .72, an increase of .59 log points, while the increase for those who change jobs is from .13 to .81, an increase of .68 log points. For the experienced workers of that groups the increases are, respectively, from .33 to .81, and from .33 to .92, even though the initial loss induced by the fact that individuals lose their firm-specific human capital is larger for the more experienced group.

The pattern of changes for the college graduate group is even more pronounced. The increases in wage due to seniority are .59 and .50 for the new entrants and experienced workers who do not change jobs. For those who do change jobs the increases are much more dramatic, namely .88 and .89, for the new entrants and experienced workers, respectively.

In both Figures 1 and 2 we also see that the change of wage due to experience is quite small. For the college graduates there is a small loss of wage for both the new entrants and experienced workers. However, for the high school graduate there is hardly any change for both the new entrants as well as the experienced workers.

Overall, the results presented in Figures 1 and 2 illustrate that not only is firm-specific human capital (i.e., seniority) an important factor in workers' wage growth over time, but also the timing of the accumulation of this human capital is crucial. Once we control for the discrete jumps (positive or negative) when an employee changes jobs, the return to seniority plays a more crucial role in wage growth than the return associated with general human capital accumulated on the job (i.e., experience).

7.4 Should J^W be Included in the Wage Equation?

The interpretation of the various parameters of the J^W function can be summarized as follows. When a worker moves between one job and another at the end of period t, the starting wage in the new firm, w_{t+1}, in comparison with w_t: (a) increases by one year of experience; (b) decreases by an amount equal to the cumulative return at the previous job; and (c) jumps by whatever difference there is between the old and the new wage offer. This is exactly what is captured by the J^W function. This function is comprised of a constant that depends on seniority at the end of the previous job, a "seniority at exit of the last job" trend (with slope that depends on seniority at the end of the previous job), an "experience at exit of the last job" trend (with slope that depends on seniority at the end of the previous job). These various different terms capture the components of wage growth that should not be attributed to rise in either experience or seniority. If a person chooses to move it is because he/she got a better offer, that is, a jump in their wage that makes up for the loss of seniority and more. The timing of a job change in one's career clearly should matter, as well as the length of the spell in the previous job. For example, a short spell may indicate a bad match, whereas a long spell may be indicative of a good match for which the worker will expect to obtain higher compensation at a new job. In general, the "optimal" length of a spell need not be the same at all points of time in one's career.

From the results presented above, it is clear that the return to seniority is quite high whether or not the J^W function is included in the wage regression. Nevertheless, when the J^W function is not included the returns to experience and seniority are dramatically affected, but in opposite directions. On one hand, all the returns to seniority, at all seniority levels and for all education groups, are significantly lower than when the the J^W function is included. On the other hand, the returns to experience, at all experience levels and for all education groups, are significantly higher than when the the J^W function is included. While we argue that one must control for the discrete jumps that occur when a person changes job, it remains largely an empirical question as to whether or not the model should include the J^W function.[20]

Theoretically, a person who changes jobs can experience one of two possible observed outcomes. If the person changes jobs because of a better offer, there will be a discrete positive jump

[20] Note that in the data we cannot distiguish between involuantary and voluntary quits. Hence, we consider all job changes as if they were based on decisions made by the individuals. This is not such a string assumption, because any firm would be willing to employ an individual for the right wage, even if it meant that the wage should be zero.

in his/her wage. The newly obtained wage represents the value of the person to the firm hiring him. This value can be attributed only to the person's stock of general human capital, since he/she has no specific human capital at the new firm. If we fail to control for this discrete jump we will attribute more to the general part of human capital, i.e., education and experience, and less for specific human capital, because at that point of time it is nonexistent by definition. What is important for examining the return to seniority is the trajectory of wages within the firm, net of the increase in experience.

The other possible outcome is that a person loses his/her job before acquiring a new job. The trajectory of such a person can be very different from a person who chose to switch jobs because of a higher wage offer. In any case, the negative discrete jump in wage that such a person may experience will lead to exactly the reverse effect on the two returns of interest, namely experience and seniority. The specification of the J^W function in (17) allows for both of these possibilities. Furthermore, it allows the size of the jump to be a function of both experience and seniority at the time a job change occur. Therefore, a person that switches from one job to another is allowed to have a different jump in their wage relative to a person who first went through a period of unemployment.

The estimates of the parameters of the J^W function defined in (17) are provided in rows (7) through (19) of Table 4. While not all of the individual parameters are statistically significant, we can easily reject the null hypothesis that all coefficients are zeros at any reasonable significance level, for all three educational groups. Previous studies, explicitly or implicitly, assume that at the start of a new job only the level of experience matters. The results obtained here indicate that the entire past career of an individual matters as well. In fact, the results indicate there is substantial amount of exchangeability among alternative career paths. That is, individuals can have the same wage levels at a given point in time, while arriving at that point through different career paths.

Note that while there is considerable similarity across the three education groups, it is worthwhile noting that the coefficient estimates are significantly different across the three education groups. In general, there is a larger positive increase due to a move. The increases are larger on average for the more highly educated individuals. Also, the longer one stays on any given job the larger the jump will be when switching to a new job. As one might expect, the magnitude of the effect of seniority diminishes over time as seniority increases, although the cumulative effects of seniority are substantial and more so for the more educated individuals.

We can summarized the results for the J^W function presented in Table 4 as follows. First, the timing of a move in a worker's career matters. High-school dropouts lose when they move after a long spell of employment, and more so if they move late in their career. College-educated workers lose after a short spell of employment, and more so late in their career. However, potential losses exist only with respect to experience. In contrast, seniority in the previous job always has a positive effect on the starting wage of the new job.

The estimation of the J^W function is possible only because we endogenize both experience and seniority, by directly modelling the participation and mobility decision. The results clearly show that a specification that allows for only direct effects of seniority and experience does not account for all components of wage changes that stem from previous jobs. These effects introduce a non-linear interaction between seniority and experience that is important for understanding wage growth.

For very similar arguments, that are articulated in detail above, Topel (1991) hypothesizes that the estimates he obtained provide lower bounds on the returns to seniority (see the literature review above). Once we control for the various potential biases, we are able to validate his claim, namely that his estimates are a lower bound on the returns to seniority, and most importantly, that the downward bias is quite substantial.

7.5 Estimate of the Stochastic Elements

In Table 7 we provide the estimate of all the stochastic elements of the model, that is the parameters of Σ in (24), the correlation parameters of Δ_ρ in (20) and δ_y and δ_m defined in (19). We omit for brevity the estimates of the γ_j's vectors in (22).

Note that the parameter $\rho_{u\xi}$ of Σ, i.e., the correlation between the idiosyncratic errors in the participation and wage equations is negative and significant for all education groups. The other correlations, particularly those between the idiosyncratic terms of the participation and mobility equation are not statistically significant and are quite small in magnitude. Also, consistent with the literature on the wage structure in the U.S., the estimated variance for the wage equation is larger for the group of less educated individuals.

Of particular significance are the estimates of the parameter of the Δ_ρ matrix defined in (20). First we see that except for two of the correlations for the high school dropouts, all other coefficients are highly significant and with the expected signs. Of particular importance is the positive correlation between the person-specific effects of the participation and wage equations, and the negative correlation between the participation and mobility equations. This structure has particular implications on the potential biases in the estimated returns to experience and seniority, when one does not control for the person-specific effects, and especially when one treats experience as an exogenous variable.

Specifically, since higher participation rates imply faster accumulation of labor market experience, the highly positive correlation between the person-specific effects in the wage equation and in the participation equation implies that, all other things equal, high-wage workers (i.e., workers with large person-specific effect in the wage equation) would tend to be more experienced workers. Omitting the person specific effect from the wage regression would therefore induce upward bias on the estimated return to experience. However, not controlling for the strong selection effect would induce a downward bias on the (positive) return to experience. Because of these two conflicting effects we cannot determine ex-ante the overall effect on the return to experience. Empirically, we find that the two effects actually cancel each other out. Hence, the overall estimate of the return to experience largely agrees with those found by Topel, AS, and AW.

Note however, that our wage equation includes the J^W function, which also contains some elements of past experience. The overall effect of these elements on wage changes when switching jobs is positive. Hence, omission of the J^W function would induce upward bias in the estimated returns to experience. We can verify that, in fact, this effect is very large, as is demonstrated by the unreasonably large returns reported in Panel C of Table 5. These effects are also larger for the more educated individuals, because the correlation coefficient between the person-specific effects are stronger for the more highly educated groups.

Similarly, the effect of the positive correlation between the person-specific effects in the participation and wage equation, the negative significant correlation between the person-specific effects in the mobility and wage induce similar biases as with the experience. To see that, note that this negative correlation implies that high-wage workers move less, that is, they have on average higher seniority than low-wage workers. The potential biases from not controlling for the person-specific effects and the selectivity due to mobility are identical to those described above for the experience variable. Empirically we find that the overall effect is a downward bias in the estimated returns to seniority that are corrected by the joint estimation of the mobility and the wage equations, and the inclusion of person-specific effect in both equations. As for the experience variable, omission of the J^W function induces a bias, only that the induced bias is a downward bias on the estimated returns to seniority, that is corrected by the introduction of the J^W function.

The last point worth emphasizing is the effect of the initial conditions on the estimation results obtained in this study. The last two rows of Table 7 provide the estimates of the coefficients on the individual-specific effects in the two initial condition equations specified in (19). The estimate for δ_y is positive, highly significant, and statistically indistinguishable from 1, for all education groups. This indicates that the initial condition for the participation is as important as those for the participation equation at any other year. The estimate for δ_m is also large, and statistically significant for the high school graduates and college graduates, but is insignificant for the high school dropouts. Hence, a failure to control for the initial conditions will results in downward biases in the estimation of the returns to experience and seniority as has already been indicated above.

7.6 Are the Results an Artifact of the Chosen Data Extract?

Obviously, the results presented in this paper, especially for the returns to seniority, are quite different from those previously reported in the literature. Above we investigated the differences that stem from the fact that we include the J^W function in the wage equation. We also showed the importance of a joint estimation of the participation and mobility decision along with the outcome variable, namely the wage equation. Finally, we demonstrated above that it is necessary to control for the individual-specific effects in all three equations.

A more basic question that one can ask is whether the mere choice of the data extract could have lead to the differences in the results reported here. Even though we use the PSID data, just like Topel (1991), AS, and AW, the selection criteria imposed here are somewhat different and the sample period has been extended to 1992. To examine whether or not the choice of the data extract lead to differences in the results we re-estimate the models specified by Topel (1991) and AW using the methods suggested in these two papers.

In Table 8 we report the results of this investigation. We organized the table similar to Table 2 in AW, so as to facilitate the comparison with previous literature. In the first four columns of the table we report the results from an ordinary least squares (OLS) estimation, for the whole sample, and then for the three education groups. In the next four columns we present the results based on the estimation suggested by AS, which had also been used by AW. This is the instrumental variable (IV) estimation that is referred to by AW as IV1. Finally, in the last four columns we report the results based on the two-step method suggested by Topel (1991). In all regressions we use the same specification as in AW, that is, they include quartic in tenure, quartic in experience, education, marital status, dummy variable for union membership, current disability status, SMSA status dummy variable, residence in a city with population over 500,000 and eight regional dummy variables according to the CPS definition. In addition when we use the AW method (i.e., IV1) we also include a linear time trend in the regressions.

Note first that for both the linear tenure coefficient and the linear experience coefficient we obtain almost the same results as in Table 2 of AW. For the linear tenure coefficients AW report .0529, .0343, and .0623, for the three methods described above, respectively. The estimates based on our data extract are .0493, .0398, and .0673 for the same three methods, respectively. Clearly, these estimates are very close, if not indistinguishable. The same happen for the linear experience coefficient. AW report .0479, .0458, and .0634, while the estimates based on our data extract are .0542, .0501, and .0643, for the same three methods, respectively. AW do not report the other coefficients that correspond to the higher order terms of tenure and experience, but they do report the implied cumulative returns to seniority and experience.

The results show that the estimates of the returns to seniority based on our sample extract are very close to those reported by AW. If anything, the returns obtained based on our extract are slightly smaller than those that are based on the data extract used by AW. The situation with the returns to experience is reversed in the sense that the estimates based on our data extract are slightly higher than those reported by AW. But this can be due to the fact that our sample extends to 1992, that is, it takes us over the period in which the return to experience increased for all parts of the population.

With these results in mind we feel comfortable in arguing that the differences in the results are not due to auxiliary decisions that have little to do with the methods employed. Rather, the differences stem from the specific modelling strategy adopted here as has already been discussed above.

8 Summary and Conclusions

The most fundamental prediction of the theory of human capital is that compensation, in the form of wage, rises with seniority in a firm. The existence of firm-specific capital explains the prevalence of long-term relationships between employees and employers. Nevertheless, there is much disagreement about the empirical evidence, as well as disagreement about the appropriateness of the methods used, to assess such theories. In a seminal paper, Topel (1991) concludes that there are significant returns to seniority and hence strong support for the theoretical literature on human capital. This finding is in stark contradiction with most previous studies in the literature that concluded that there is no evidence supporting the view that there are positive returns to seniority. One particular paper in the literature that criticized Topel's (1991) work is Altonji and Williams's (1997) study, which largely supports the earlier findings.

In this study we reinvestigate the interrelations between participation, mobility and wages, while examining several questions central to labor economics. Specifically, we model the joint decisions of participation and job mobility, while allowing for potential sample selection biases to exist when estimating the equation of interest, namely the wage function. This allows us to address, in a more satisfactory way, the question as to whether or not there is return to seniority in the United States. We provide new evidence on the returns to seniority and experience. To do that, we use data similar to that used by both Topel (1991) and Altonji and Williams (1997).

There are two main differences between the current study and earlier studies. Here we explicitly model a participation and a mobility equation along with the wage equation. Furthermore, we explicitly specify a model which allows for accumulation of returns to seniority within a job, as well as direct effects of past careers on the starting wage at a new job. The results clearly demonstrate the importance of this joint estimation of the wage equation along with the participation and mobility decisions. These two decisions have significant effects on the observed outcomes of annual earnings. We resort to a Bayesian analysis, which extensively uses Markov Chain Monte Carlo methods, allowing one to compute the posterior distribution of the model's parameters. Whenever possible we use uninformative prior distributions for the parameters and hence rely heavily on the data to determine the posterior distributions of these parameters.

We examine three educational groups. The first group consists of all those that acquired less than high school education. The second group consists of all those who acquired at least high school education, but have not completed four-year college. The third group is comprised of only college graduates. We find very strong evidence supporting Topel's (1991) claim, even though some aspects of our modelling strategy are closer to Altonji and Williams (1997). There are large, and statistically significant, returns to seniority for all groups considered, although some differences across groups do exist. Our estimates of the returns to experience are on the same order of magnitude as those previously obtained in the literature. Nevertheless, they are not uniform across education groups; they are much higher for the college graduates than for the other two education groups.

In addition, we are able to uncover the impact of job-to-job mobility on the pattern of wage growth, which differ markedly across education groups. In particular, the introduction of J_W, a non-linear summary of the career, has important consequences for all education groups. For instance, the timing of a move in a career matters. Early moves are most beneficial to college-educated workers whereas late moves can be very detrimental for workers with lower education. Short spells are sometimes good, sometimes bad. Mobility through the wage distribution is achieved through a combination of wage increases within the firm and across firms. The former is more important for wage growth of high school dropouts because of their lower returns to experience. The latter is more important for college graduates, because both returns to experience and seniority are quite large.

9 References

Abowd, J., Finer, H., and F. Kramarz (1999): "Individual and Firm Heterogeneity in Compensation: An Analysis of Matched Longitudinal Employer and Employee Data for the State of Washington," in *The Creation and Analysis of Employer-Employee Matched Data*, ed. by J. Haltiwanger, J. Lane, J. Spletzer, K. Troske. Amsterdam: North-Holland, 3–24.

Abowd, J., Finer, H., Kramarz, F., and S. Roux (1997): "Job and Wage Mobility: An Analysis of the Dynamics of Employment Durations Using Matched Employee and Employer Data from the U.S. and France," paper presented at the NBER Summer Labor Studies Group, Boston.

Abowd, J., Kramarz, F., and D. Margolis (1999): "High-Wage Workers and High-Wage Firms," *Econometrica,* 67, 251–333.

Abraham, K.G., and H.S. Farber (1987): "Job Duration, Seniority and Earnings," *American Economic Review*, 77, 278–97.

Addison, J.T., and P. Portugal (1989): "Job Displacement, Relative Wage Changes and Duration of Unemployment," *Journal of Labor Economics*, 7, 281–302.

Altonji, J., and R. Shakotko (1987): "Do Wages Rise with Job Seniority?" *Review of Economic Studies,* 54, 437–59.

Altonji, J., and N. Williams (1997): "Do Wages Rise with Job Seniority? A Reassessment," NBER Working Paper 6010.

Becker, G. (1964): *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education,* NBER, Cambridge, MA.

Buchinsky, M., Fougère, D., and F. Kramarz (1998): "La Mobilité des Salaires en France: 1967-1987," *Revue Economique,* 49, 879–90.

Buchinsky, M., and J. Hunt (1999): "Wage Mobility in the United States," *Review of Economics and Statistics*, 131, 351–68.

Burdett, K. (1978): "A Theory of Employee Job Search and Quit Rates," *American Economic Review*, 68, 212-220.

Burkhauser, R., Holtz-Eakin, D., and S. Rhody (1997): "Earnings Mobility and Inequality in the United States and Germany during the Growth Years of the 1980s," *International Economic Review*, 38, 775–94.

Card, D. (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometrics Problem," *Econometrica*, 69, 1127–1160.

Chib, S. (2001): "Markov Chain Monte Carlo Methods: Computation and Inference," Chapter 57 in *Handbook of Econometrics,* Volume 5, ed. by J. J. Heckman and E. Leamer. Amsterdam: North-Holland, 3569-3649.

Chib, S. and B.P. Carlin (1999): "On MCMC Sampling in Hierarchical Longitudinal Models," *Statistics and Computing*, 9, 17–26.

Chib, S. and E. Greenberg (1995): “Understanding the Metropolis-Hastings Algorithm,” *American Statistician*, 49, 327–35.

Chib, S. and E. Greenberg (1996): “Markov Chain Monte Carlo Simulation Methods in Econometrics,” *Econometric Theory*, 12, 409–31.

Chib, S. and E. Greenberg (1998): “Analysis of Multivariate Probit Models,” *Biometrika*, 85, 347–61.

Chib, S. and B. H. Hamilton (2002): “Semiparametric Bayes Analysis of Longitudinal Data Treatment Models,” *Journal of Econometrics*, 110, 67-89.

Dustmann, C. and C. Meghir (2005): “Wages, Experience and Seniority,” *Review of Economic Studies,* 72, 77–108.

Farber, H. (1999): “Mobility and Stability,” Chapter 37 in *Handbook of Labor Economics*, Volume 3B, ed. by O. Ashenfelter and D. Card. New-York: North-Holland, 2441–83.

Fougère, D. and T. Kamionka (2003): “Bayesian Inference for the Mover-Stayer Model in Continuous Time, with an Application to Labour Market Transition Data,” *Journal of Applied Econometrics*, 18, 697–723.

Geweke, J. and M. Keane (2000): “An Empirical Analysis of Earnings Dynamics Among Men in the PSID: 1968–1989,” *Journal of Econometrics*, 96, 293-356.

Gibbons, R. and L. Katz (1991): “Layoffs and Lemons,” *Journal of Labor Economics*, 9, 351-380.

Gouriéroux, C. and A. Monfort (1996): *Simulation-Based Econometric Methods*, CORE Lectures. Oxford and New York: Oxford University Press.

Heckman, J. (1981): “Heterogeneity and State Dependence,” in *Studies in Labor Markets,* ed. by S. Rosen. Chicago: University of Chicago Press.

Hill, M. S. (1992): *The Panel Study of Income Dynamics: A User's Guide.* Newbury Park, CA: Sage Publications.

Hyslop D. R. (1999): “State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women,” *Econometrica*, 67, 1255-1294.

Jacobson, L., LaLonde, L., and D. Sullivan (1993): “Earnings Losses of Displaced Workers,” *American Economic Review,* 83, 685–709.

Kyriazidou, E. (1999): “Estimation of Dynamic Panel Data Sample Selection Models,” Unpublished manuscript, Department of Economics, University of Chicago.

Lillard, L.A. and R.J. Willis (1978): “Dynamic Aspects of Earnings Mobility,” *Econometrica*, 46, 985–1012.

McFadden, D. (1989): “A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration,” *Econometrica*, 57, 995–1026.

Mincer, J. (1988): “Job Training, Wage Growth, and Labor Turnover,” National Bureau of Economic Research Working Paper: 2690.

Mundlak, Y. (1978): "On the Pooling of Time Series and Cross-Section Data," *Econometrica*, 46, 69–85.

Neal, D. (1995): "Industry-Specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 13, 653–677.

Pakes, A. and D. Pollard (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057.

Parent, D. (1999): "Wages and Mobility: The Impact of Employer-Provided Training," *Journal of Labor Economics*, 17, 298-317.

Parent, D. (2000): "Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics," *Journal of Labor Economics*, 18, 306-323.

Robert, C. and G. Casella (1999): *Monte Carlo Statistical Methods.* New-York: Springer-Verlag.

Robinson, C. (1989): "The Joint Determination of Union Status and Union Wage Effects: Some Tests of Alternative Models," *Journal of Political Economy*, 97, 639–667.

Topel, R.H. (1991): "Specific Capital, Mobility and Wages: Wages Rise with Job Seniority," *Journal of Political Economy*, 99, 145–76.

Topel, R.H. and M.P. Ward (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 439–479.

Appendix A—Mathematical Relationships between Reservation Wages

The reservation wages for non-participants and stayers, in (7) and (8), respectively, imply that

$$\begin{aligned} & u\left(z_t + w^*_{11,t}, 1; X_t\right) - u\left(z_t + w^*_{01,t} - \gamma_1, 1; X_t\right) \\ & = u\left(z_t, 0; X_t\right) - u\left(z_t - \gamma_1, 0; X_t\right), \\ & = u\left(z_t + w^*_{12,t}, 1; X_t\right) - u\left(z_t + w^*_{02,t} - \gamma_1, 1; X_t\right). \end{aligned} \tag{29}$$

First-order Taylor series expansions of the left and right hand sides of (29) around $z_t + w^*_{01,t}$, $z_t + w^*_{02,t}$, and z_t, respectively, gives

$$w^*_{1j,t} \approx w^*_{0j,t} - \gamma_1 \left[1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{0j,t}, 1; X_t\right)}\right] \approx w^*_{0j,t} - \gamma_{1j}, \tag{30}$$

for $j = 1, 2$, where $u'(\cdot)$ denotes the marginal utility of consumption, and

$$\gamma_{1j} = \gamma_1 \left[1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{0j,t}, 1; X_t\right)}\right].$$

If the utility function is concave with respect to consumption, then $u'\left(z_t + w^*_{0j,t}, 0; X_t\right) < u'\left(z_t, 0; X_t\right)$. However, if the marginal utility of consumption is greater when working, that is,

$$u'\left(z_t + w^*_{0j,t}, 0; X_t\right) < u'\left(z_t + w^*_{0j,t}, 1; X_t\right),$$

then γ_{1j} may be positive or negative, dependent on whether $u'\left(z_t + w^*_{0j,t}, 1; X_t\right)$ is greater or smaller than $u'\left(z_t, 0; X_t\right)$. But, if the marginal utility of consumption is lower when working, then γ_{1j} is always negative, because

$$u'\left(z_t + w^*_{0j,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right).$$

A similar inspection of the reservation wages for non-participants and movers, in (7) and (9), gives

$$\begin{aligned} & u\left(z_t + w^*_{21,t} - c_M, 1; X_t\right) - u\left(z_t + w^*_{01,t} - \gamma_1, 1; X_t\right) \\ & = u\left(z_t - c_M, 0; X_t\right) - u\left(z_t - \gamma_1, 0; X_t\right), \\ & = u\left(z_t + w^*_{22,t} - c_M, 1; X_t\right) - u\left(z_t + w^*_{02,t} - \gamma_1, 1; X_t\right). \end{aligned} \tag{31}$$

Again, first-order Taylor series expansions of the left and right hand sides of (31) around $z_t + w^*_{0j,t}$, $z_t + w^*_{2j,t}$, and z_t, respectively, give (for $j = 1, 2$):

$$\begin{aligned}
u\left(z_t + w^*_{2j,t} - c_M, 1; X_t\right) &\simeq u\left(z_t + w^*_{2j,t}, 1; X_t\right) - c_M \; u'\left(z_t + w^*_{2j,t}, 1; X_t\right), \\
u\left(z_t + w^*_{0j,t} - \gamma_1, 1; X_t\right) &\simeq u\left(z_t + w^*_{0j,t}, 1; X_t\right) - \gamma_1 \; u'\left(z_t + w^*_{0j,t}, 1; X_t\right), \\
u\left(z_t - c_M, 0; X_t\right) &\simeq u\left(z_t, 0; X_t\right) - c_M \; u'\left(z_t, 0; X_t\right), \\
u\left(z_t - \gamma_1, 0; X_t\right) &\simeq u\left(z_t, 0; X_t\right) - \gamma_1 \; u'\left(z_t, 0; X_t\right), \quad \text{and} \\
u\left(z_t + w^*_{2j,t}, 1; X_t\right) - u\left(z_t + w^*_{0j,t}, 1; X_t\right) &\simeq \left(w^*_{2j,t} - w^*_{0j,t}\right) \; u'\left(z_t + w^*_{0j,t}, 1; X_t\right).
\end{aligned} \tag{32}$$

Substitution of the expressions from (32) back into (31) gives (for $j = 1, 2$):

$$\begin{aligned}
w^*_{2j,t} &\approx w^*_{0j,t} - \gamma_{1j} + \gamma_{2j} \approx w^*_{1j,t} + \gamma_{2j}, \quad \text{where} \\
\gamma_{2j} &= c_M \left[\frac{u'\left(z_t + w^*_{2j,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{0j,t}, 1; X_t\right)}\right], \quad \text{and} \\
\gamma_{1j} &= \gamma_1 \left[1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{0j,t}, 1; X_t\right)}\right].
\end{aligned} \tag{33}$$

Note that if the marginal utility of consumption is greater when working, i.e.,

$$u'\left(z_t + w^*_{2j,t}, 0; X_t\right) < u'\left(z_t + w^*_{2j,t}, 1; X_t\right),$$

then γ_{2j} may be either positive or negative, that is,

$$\gamma_{2j} \gtrless 0 \quad \Leftrightarrow \quad u'\left(z_t + w^*_{2j,t}, 1; X_t\right) \gtrless u'\left(z_t, 0; X_t\right).$$

However, if the marginal utility of consumption is lower when working, then γ_{2j} is always negative, since

$$u'\left(z_t + w^*_{2j,t}, 1; X_t\right) < u'\left(z_t + w^*_{2j,t}, 0; X_t\right) < u'\left(z_t, 0; X_t\right).$$

Appendix B—Proof of Conditions 1 and 2

Throughout the proof, we assume that the mobility cost is strictly greater than the search cost, that is $c_M > \gamma_1$. We consider successively the four possible cases.

A. $w^*_{02,t} < w^*_{01,t}$ and the marginal utility of consumption is higher when working:

If utility is concave with respect to consumption, $w^*_{02,t} < w^*_{01,t}$ implies that

$$u'\left(z_t + w^*_{01,t}, 1; X_t\right) < u'\left(z_t + w^*_{02,t}, 1; X_t\right),$$

and thus

$$\frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} < \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)}. \tag{34}$$

A.1. If $w^*_{02,t} < w^*_{22,t}$:
The concavity of the utility function implies that

$$u'\left(z_t + w^*_{22,t}, 1; X_t\right) < u'\left(z_t + w^*_{02,t}, 1; X_t\right),$$

and thus

$$\frac{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)}. \tag{35}$$

1. If $u'\left(z_t, 0; X_t\right) > u'\left(z_t + w^*_{02,t}, 1; X_t\right)$ then

$$u'\left(z_t, 0; X_t\right) > u'\left(z_t + w^*_{01,t}, 1; X_t\right)$$

and equation (34) implies that $\gamma_{11} < \gamma_{12} < 0$. Consequently, (30) implies that $w^*_{02,t} < w^*_{12,t} < w^*_{11,t}$ and $w^*_{02,t} < w^*_{01,t} < w^*_{11,t}$. Moreover, since $w^*_{02,t} < w^*_{22,t}$, we have

$$u'\left(z_t, 0; X_t\right) > u'\left(z_t + w^*_{22,t}, 1; X_t\right).$$

Due to equation (35), this inequality implies that, if $c_M > \gamma_1$, $\gamma_{22} < \gamma_{12} < 0$. Consequently (31) implies then that $w^*_{22,t} < w^*_{02,t}$, which contradicts the initial assumption. Thus $w^*_{22,t}$ cannot be greater than $w^*_{02,t}$ when $c_M > \gamma_1$ and $u'\left(z_t, 0; X_t\right) > u'\left(z_t + w^*_{02,t}, 1; X_t\right)$.

2. If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{01,t}, 1; X_t\right)$ then

$$u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{02,t}, 1; X_t\right)$$

and $0 < \gamma_{11} < \gamma_{12}$. Thus, equation (30) implies that $w^*_{12,t} < w^*_{11,t} < w^*_{01,t}$ and $w^*_{12,t} < w^*_{02,t} < w^*_{01,t}$. Two cases must be considered then.
Case 1: If $u'\left(z_t + w^*_{22,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right)$, then $\gamma_{22} < 0$ and $\gamma_{12} > 0$. Thus $\gamma_{22} - \gamma_{12} < 0$, and (31) implies then that $w^*_{22,t} < w^*_{02,t}$, which contradicts the initial assumption.
Case 2: If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{22,t}, 1; X_t\right)$, then

$$0 < \frac{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)}.$$

Thus, $w^*_{22,t} > w^*_{02,t}$ if and only if $\gamma_{22} - \gamma_{12} > 0$, or alternatively, if and only if

$$c_M > \gamma_1 \left[\frac{u'\left(z_t + w^*_{02,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}\right].$$

The ratio between brackets being greater than 1, this last inequality is a stronger assumption than $c_M > \gamma_1$; nevertheless it is a sufficient condition to have $w^*_{22,t} > w^*_{02,t}$. Moreover, (32) implies that $w^*_{22,t} > w^*_{02,t} > w^*_{12,t}$. If $w^*_{21,t} > w^*_{22,t}$, which is satisfied if $\gamma_{21} > \gamma_{22}$, a mover (i.e. a worker who moved to another firm at the end of period $t-1$) becomes non-participant at the end of period t if he or she is offered a wage less than $w^*_{22,t}$. A stayer becomes non-participant if he or she is offered a wage less than $w^*_{12,t}$, because $w^*_{12,t} < w^*_{11,t}$. Thus, the participation decision at period t can be characterized by the equation

$$\begin{aligned} y_t &= \mathbf{1}\left[w_t > w^*_{02,t} - \gamma_{12}\ y_{t-1} + \gamma_{22}\ y_{t-1}\ m_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{02,t} + \gamma_{12}\ y_{t-1} - \gamma_{22}\ y_{t-1}\ m_{t-1} > 0\right] \end{aligned} \tag{36}$$

with $\gamma_{22} > \gamma_{12} > 0$. A stayer can accept to move to another firm at the end of period t if he or she is offered a wage w_t greater than $w^*_{12,t}$ but less than $w^*_{11,t}$, while a mover accepts only to move again if the wage offer in t is at least equal to $w^*_{22,t}$. Consequently, the mobility decision at period t can be characterized by the equation

$$\begin{aligned} m_t &= \mathbf{1}\left[w_t > w^*_{12,t} + \gamma_{22}\ y_{t-1}\ m_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{12,t} - \gamma_{22}\ y_{t-1}\ m_{t-1} > 0\right] \end{aligned} \tag{37}$$

with $\gamma_{22} > 0$.

3. If $u'\left(z_t + w^*_{01,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{02,t}, 1; X_t\right)$, then equation (34) implies that $\gamma_{11} < 0 < \gamma_{12}$. Thus, (30) implies that $w^*_{12,t} < w^*_{02,t} < w^*_{01,t} < w^*_{11,t}$. Note that two cases are possible.

Case 1: If $u'\left(z_t + w^*_{22,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right)$, then $\gamma_{22} < 0$. Thus $\gamma_{22} - \gamma_{12} < 0$ and equation (31) implies that $w^*_{22,t} < w^*_{02,t}$, which contradicts the initial assumption.

Case 2: If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{22,t}, 1; X_t\right)$, then

$$0 < \frac{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)}.$$

Thus, $w^*_{22,t} > w^*_{02,t}$ if and only if $\gamma_{22} - \gamma_{12} > 0$, or alternatively, if and only if

$$c_M > \gamma_1 \left[\frac{u'\left(z_t + w^*_{02,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}\right].$$

We find the same condition as previously. Under this condition (32) implies that $w^*_{22,t} > w^*_{02,t} > w^*_{12,t}$, and if $\gamma_{21} > \gamma_{22}$, then $w^*_{21,t} > w^*_{22,t}$ and (36) and (37) are still valid.

A.2. If $w^*_{02,t} > w^*_{22,t}$:

The concavity of the utility function implies that

$$u'\left(z_t + w^*_{22,t}, 1; X_t\right) > u'\left(z_t + w^*_{02,t}, 1; X_t\right).$$

Thus,

$$\frac{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} > 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)}. \tag{38}$$

Again, two cases are possible.

Case 1: If $u'\left(z_t + w^*_{22,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right)$, then equation (38) implies that $0 > \gamma_{22} > \gamma_{12}$ (if $c_M > \gamma_1$). Then we deduce from equation (31) that $w^*_{22,t} > w^*_{02,t}$, which contradicts the assumption.

Case 2: If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{22,t}, 1; X_t\right)$, then equation (38) implies that $\gamma_{22} > \gamma_{12} > 0$ (if $c_M > \gamma_1$). Then equation (31) implies that $w^*_{22,t} > w^*_{02,t}$, which contradicts the above assumption.

Consequently, the assumption that $w^*_{02,t} > w^*_{22,t}$ appears to be implausible if the mobility cost c_M is strictly greater than the search cost γ_1.

B. $w^*_{02,t} < w^*_{01,t}$ **and the marginal utility of consumption is lower when working:**

If the utility function is concave and if the marginal utility is lower when working, then

$$u'\left(z_t + w^*_{jl,t}, 1; X_t\right) < u'\left(z_t, 1; X_t\right) < u'\left(z_t, 0; X_t\right), \; j = 0, 1, 2, \; l = 1, 2.$$

Moreover, if $w^*_{02,t} < w^*_{01,t}$,

$$u'\left(z_t + w^*_{01,t}, 1; X_t\right) < u'\left(z_t + w^*_{02,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right),$$

which implies that $\gamma_{11} < \gamma_{12} < 0$. Then, (30) implies that

$$w^*_{02,t} < w^*_{01,t} < w^*_{11,t} \text{ and } w^*_{02,t} < w^*_{12,t} < w^*_{11,t}.$$

B.1. If $w^*_{02,t} < w^*_{22,t}$, then

$$u'\left(z_t + w^*_{22,t}, 1; X_t\right) < u'\left(z_t + w^*_{02,t}, 1; X_t\right)$$

and

$$\frac{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} < 0. \tag{39}$$

Thus, if $c_M > \gamma_1$, (39) implies that $\gamma_{22} < \gamma_{12} < 0$. From equation (32), we can deduce that $w^*_{22,t} < w^*_{02,t}$, which contradicts the assumption above.

B.2. If $w^*_{02,t} > w^*_{22,t}$, then

$$u'\left(z_t + w^*_{22,t}, 1; X_t\right) > u'\left(z_t + w^*_{02,t}, 1; X_t\right)$$

and

$$0 > \frac{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} > 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)}.$$

Thus, (31) implies that $w^*_{02,t} > w^*_{22,t}$ if and only if

$$c_M > \gamma_1 \left[\frac{u'\left(z_t + w^*_{02,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}\right].$$

Under this condition, (32) implies that $w^*_{12,t} > w^*_{02,t} > w^*_{22,t}$; if $0 > \gamma_{21} > \gamma_{22}$. Then, $w^*_{21,t} > w^*_{22,t}$ and (36) and (37) are verified.

C. $w^*_{01,t} < w^*_{02,t}$ and the marginal utility of consumption is higher when working:

If utility is concave with respect to consumption, $w^*_{01,t} < w^*_{02,t}$ implies that

$$\frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)}. \tag{40}$$

C.1. If $w^*_{01,t} < w^*_{21,t}$:

The concavity of the utility function implies that

$$u'\left(z_t + w^*_{21,t}, 1; X_t\right) < u'\left(z_t + w^*_{01,t}, 1; X_t\right).$$

Thus,

$$\frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)}. \tag{41}$$

1. If $u'\left(z_t, 0; X_t\right) > u'\left(z_t + w^*_{01,t}, 1; X_t\right)$, then

$$u'\left(z_t, 0; X_t\right) > u'\left(z_t + w^*_{02,t}, 1; X_t\right),$$

and (40) implies that $\gamma_{12} < \gamma_{11} < 0$. Then, (30) implies that $w^*_{01,t} < w^*_{11,t} < w^*_{12,t}$ and $w^*_{01,t} < w^*_{02,t} < w^*_{12,t}$. Moreover, since $w^*_{01,t} < w^*_{21,t}$, it follows that

$$u'\left(z_t, 0; X_t\right) > u'\left(z_t + w^*_{01,t}, 1; X_t\right) > u'\left(z_t + w^*_{21,t}, 1; X_t\right).$$

By (41), this inequality implies that, if $c_M > \gamma_1$, then $\gamma_{21} < \gamma_{11} < 0$. Thus, (31) implies then that $w^*_{21,t} < w^*_{01,t}$, which contradicts the initial assumption. Hence, $w^*_{21,t}$ cannot be greater than $w^*_{01,t}$ when $c_M > \gamma_1$ and $u'\left(z_t, 0; X_t\right) > u'\left(z_t + w^*_{01,t}, 1; X_t\right)$.

2. If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{02,t}, 1; X_t\right)$, then

$$u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{01,t}, 1; X_t\right)$$

and $0 < \gamma_{12} < \gamma_{11}$. Thus, (30) implies that $w^*_{11,t} < w^*_{12,t} < w^*_{02,t}$ and $w^*_{11,t} < w^*_{01,t} < w^*_{02,t}$. Two cases must then be considered.
Case 1: If $u'\left(z_t + w^*_{21,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right)$, then $\gamma_{21} < 0$ and $\gamma_{11} > 0$. Thus $\gamma_{21} - \gamma_{11} < 0$, and (31) implies that $w^*_{21,t} < w^*_{01,t}$, which contradicts the initial assumption.
Case 2: If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{21,t}, 1; X_t\right)$, then

$$0 < \frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)}.$$

Thus, $w^*_{21,t} > w^*_{01,t}$ if and only if $\gamma_{21} - \gamma_{11} > 0$, or alternatively, if and only if

$$c_M > \gamma_1 \left[\frac{u'\left(z_t + w^*_{01,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}\right].$$

The ratio between brackets being greater than 1, this last inequality is a stronger assumption than $c_M > \gamma_1$; nevertheless it is a sufficient condition to have $w^*_{21,t} > w^*_{01,t}$. Moreover, (32) implies that $w^*_{21,t} > w^*_{01,t} > w^*_{11,t}$. Now let us assume that $\gamma_{22} > \gamma_{21}$, which implies then that $w^*_{22,t} > w^*_{21,t}$. In that case, there is no inter-firm mobility: a non-participant moves to employment (respectively, stays in the non-participation state) at the end of period t if he or she is offered a wage greater (respectively, lower) than $w^*_{01,t}$. A participant becomes non-participant (respectively, remains employed) if he or she is offered a wage less than $w^*_{11,t}$. Thus, the participation decision at period t can be characterized by the equation

$$\begin{aligned} y_t &= \mathbf{1}\left[w_t > w^*_{01,t} - \gamma_{11}\ y_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{01,t} + \gamma_{11}\ y_{t-1} > 0\right], \end{aligned} \tag{42}$$

with $\gamma_{11} > 0$.

3. If $u'\left(z_t + w^*_{02,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{01,t}, 1; X_t\right)$, then (40) implies that $\gamma_{12} < 0 < \gamma_{11}$, and (30) implies that $w^*_{11,t} < w^*_{01,t} < w^*_{02,t} < w^*_{12,t}$. Again, there are two possible cases.
Case 1: If $u'\left(z_t + w^*_{21,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right)$, then $\gamma_{21} < 0$. Thus, $\gamma_{21} - \gamma_{11} < 0$, and (31) implies that $w^*_{21,t} < w^*_{01,t}$, which contradicts the initial assumption.
Case 2: If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{21,t}, 1; X_t\right)$, then

$$0 < \frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)}.$$

Thus, $w^*_{21,t} > w^*_{01,t}$ if and only if $\gamma_{21} - \gamma_{11} > 0$, or alternatively, if and only if

$$c_M > \gamma_1 \left[\frac{u'\left(z_t + w^*_{01,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}\right].$$

We find the same condition as previously. Under this condition, (32) implies that $w^*_{21,t} > w^*_{01,t} > w^*_{11,t}$, and if $\gamma_{22} > \gamma_{21}$, then $w^*_{22,t} > w^*_{21,t}$,and (42) is still valid.

C.2. If $w^*_{01,t} > w^*_{21,t}$:

The concavity of the utility function implies that

$$u'\left(z_t + w^*_{21,t}, 1; X_t\right) > u'\left(z_t + w^*_{01,t}, 1; X_t\right).$$

Thus,

$$\frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} > 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)}. \tag{43}$$

Note that there are two possible cases.
Case 1: If $u'\left(z_t + w^*_{21,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right)$, then (43) implies that $0 > \gamma_{21} > \gamma_{11}$ (if $c_M > \gamma_1$). Then, we can deduce from equation (31) that $w^*_{21,t} > w^*_{01,t}$, which contradicts the assumption above.
Case 2: If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{21,t}, 1; X_t\right)$, then (43) implies that $\gamma_{21} > \gamma_{11} > 0$ (if $c_M > \gamma_1$). Then, (31) implies that $w^*_{21,t} > w^*_{01,t}$, which still contradicts the assumption above.

Consequently, the assumption $w^*_{01,t} > w^*_{21,t}$ appears to be implausible if the mobility cost c_M is strictly greater than the search cost γ_1.

D. $w^*_{01,t} < w^*_{02,t}$ **and the marginal utility of consumption is lower when working:**

If the utility function is concave and if the marginal utility is lower when working, then

$$u'\left(z_t + w^*_{jl,t}, 1; X_t\right) < u'\left(z_t, 1; X_t\right) < u'\left(z_t, 0; X_t\right),$$

for $j = 0, 1, 2$, and $l = 1, 2$. Moreover, if $w^*_{01,t} < w^*_{02,t}$,

$$u'\left(z_t + w^*_{02,t}, 1; X_t\right) < u'\left(z_t + w^*_{01,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right)$$

which implies that $\gamma_{12} < \gamma_{11} < 0$. Then, (30) implies that $w^*_{01,t} < w^*_{02,t} < w^*_{12,t}$ and $w^*_{01,t} < w^*_{11,t} < w^*_{12,t}$.

D.1. If $w^*_{01,t} < w^*_{21,t}$, then

$$u'\left(z_t + w^*_{21,t}, 1; X_t\right) < u'\left(z_t + w^*_{01,t}, 1; X_t\right)$$

and

$$\frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < 0. \tag{44}$$

Thus, if $c_M > \gamma_1$, (44) implies that $\gamma_{21} < \gamma_{11} < 0$. From (32), we can then deduce that $w^*_{21,t} < w^*_{01,t}$, which contradicts the assumption above.

D.2. If $w^*_{01,t} > w^*_{21,t}$, then

$$u'\left(z_t + w^*_{21,t}, 1; X_t\right) > u'\left(z_t + w^*_{01,t}, 1; X_t\right)$$

and

$$0 > \frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} > 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)}.$$

Thus, (31) implies that $w^*_{01,t} > w^*_{21,t}$ if and only if

$$c_M > \gamma_1 \left[\frac{u'\left(z_t + w^*_{01,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}\right].$$

Under this condition, equation (32) implies that $w^*_{11,t} > w^*_{01,t} > w^*_{21,t}$. If $0 > \gamma_{22} > \gamma_{21}$, then $w^*_{22,t} > w^*_{21,t}$. Once again, the ranking of the reservation wages implies that there is no interfirm mobility, and the participation decision at period t is simply characterized by the equation (42).

Appendix C—Drawing from the Posterior Distribution

We use here the notation established in Section (4).

Note that in matrix form we can write the model in (14), (15), and (16) as

$$z_{it}^{*} = \tilde{x}_{it}\beta + L_t\alpha_i + \tau_{it}, \tag{45}$$

for $t = 1, ..., T$, where $\alpha_i \sim N(0, \Gamma_i)$, as is defined in (20), $\tau_{it} \sim N(0, \Sigma)$, as defined in (23).

For $t = 1$ we define

$$\tilde{x}_{i1} = \begin{pmatrix} x_{yi1} & 0 & 0 & 0 & 0 \\ 0 & x_{mi1} & 0 & 0 & 0 \\ 0 & 0 & x_{wi1} & 0 & 0 \end{pmatrix} \quad \text{and } L_1 = \begin{pmatrix} \delta_1 & 0 & 0 \\ 0 & \delta_2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

while for $t > 1$ we define

$$\tilde{x}_{it} = \begin{pmatrix} 0 & 0 & 0 & x_{yit} & 0 \\ 0 & 0 & 0 & 0 & x_{mit} \\ 0 & 0 & x_{wit} & 0 & 0 \end{pmatrix} \quad \text{and } L_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For clarity of presentation we define a few other quantities as follows. The parameter vector β consists of the regression coefficients in (14), (15), and (16), including the parameters from the function J_{it}^{W} defined in (17), and the parameter vectors from the initial condition equations in (19). The parameter vector γ consists of the coefficients in (21). Note that the covariance matrix for α_i, Γ_i, is constructed from γ and Δ_ρ defined in (20). Let the vector α contain all the individuals specific random effects, that is, $\alpha' = (\alpha_1', ..., \alpha_N')$. For convenience we use the notation $\Pr(t \mid \theta_{-t})$ to denote the distribution of t, conditional on all the elements in θ, not including t.[21] Below we explain the sampling of each of the parts in θ (augmented by z^*), conditional on all the other parts and the data.

Sampling the Latent Variables z^*:

There are three latent dependent variables: y_{it}^*, m_{it}^*, and w_{it}^*. While y_{it}^* and m_{it}^* are never directly observed, w_{it}^* is observed if the ith individual worked in year t. Conditional on θ, the distribution of the latent dependent variables is

$$z_{it}^*|\theta \sim N(\tilde{x}_{it}^*\beta + L_t\alpha_i, \Sigma).$$

From this joint distribution we can infer the conditional univariate distributions of interest, that is $\Pr(y_{it}^*|m_{it}^*, w_{it}^*, \theta)$ and $\Pr(m_{it}^*|y_{it}^*, w_{it}^*, \theta)$, which are truncated univariate normals, with truncation regions that depend on the values of y_{it} and m_{it}, respectively. Note that m_{it} and w_{it} are observed only if $y_{it} = 1$. Therefore, when $y_{it} = 1$ we sample m_{it}^* from the appropriate truncated distribution. In contrast, when $y_{it} = 0$, the distribution of m_{it}^* is not truncated. Similarly, we can infer the distribution of the unobserved (hypothetical) wages, $\Pr(w_{it}^*|y_{it}^*, m_{it}^*, \theta)$.

Sampling the Regression Coefficients β:

It can be easily shown (see Chib and Greenberg (1998) for details) that if the prior distribution of β is given by $\beta \sim N(\widetilde{\beta}_0, \widetilde{B}_0)$, then the posterior distribution of β, conditional on all

[21] For a similar hierarchical model see also Chib and Carlin (1999).

other parameters is

$$
\begin{aligned}
\beta|\theta_{-\beta} &\sim N(\hat{\beta}, B), \quad \text{where} \\
\hat{\beta} &= B\left(\widetilde{B}_0^{-1}\beta_0 + \sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{x}_{it}'\,\Sigma^{-1}(z_{it}^* - L_t\alpha_i)\right) \quad \text{and} \\
B &= \left(B_0^{-1} + \sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{x}_{it}'\Sigma^{-1}\tilde{x}_{it}\right)^{-1},
\end{aligned}
$$

where

$$
\begin{aligned}
\widetilde{\beta} &= (\beta', \delta_1, \delta_2)', \\
x_{it}^* &= (\widetilde{x}_{it}', D_{y1}, D_{m1}),
\end{aligned}
$$

$D_{y1} = 1$ and $D_{m1} = 1$ if the person is in his/her first year in the sample, and $D_{y1} = 0$ and $D_{m1} = 0$, otherwise.

Sampling the Individuals' Random Effects α_i:

The conditional likelihood of the random effects for individual i is

$$
l(\alpha_i) \propto \Sigma^{-T/2}\exp\left\{-.5\sum_{t=1}^{T}(z_{it}^* - \tilde{x}_{it}\beta - L_t\alpha_i)'\Sigma^{-1}(z_{it}^* - \tilde{x}_{it}\beta - L_t\alpha_i)\right\}.
$$

The prior distribution for the random effects is $N(0, \Gamma_i)$, so that the posterior distribution of α_i is

$$
\begin{aligned}
\alpha_i &\sim N(\alpha_i, V_{\alpha_i}), \quad \text{where} \\
V_{\alpha_i} &= \left(\Gamma_i^{-1} + \sum_{t=1}^{T} L_t'\Sigma^{-1}L_t\right)^{-1}, \quad \text{and} \\
\alpha_i &= V_{\alpha_i}\sum_{t=1}^{T} L_t'\Sigma^{-1}(z_t^* - \tilde{x}_{it}\beta).
\end{aligned}
$$

Sampling the Covariance Matrix Σ:

Recall that the covariance matrix of the idiosyncratic error terms, τ_{it}, is given in (24). Since the conditional distribution of Σ is not a standard, known distribution, it is impossible to sample from it directly. Instead, we sample the elements of Σ using the Metropolis-Hastings (M-H) algorithm (see Chib and Greenberg (1995)). The target distribution here is the conditional posterior of Σ, that is,

$$
p(\Sigma|\theta_{-\Sigma}) \propto l(\Sigma|\theta_{-\Sigma}, \alpha_i, z_{it}^*)p(\sigma_\xi^2)p(\rho).
$$

The likelihood component is given by

$$
l(\Sigma|\theta_{-\Sigma}, \alpha_i, z_{it}^*) = |\Sigma|^{-NT/2}\exp\left\{\sum_{i=1}^{N}\sum_{t=1}^{T} A_{it}'\,\Sigma^{-1}A_{it}\right\},
$$

where, $A_{it} = z_{it}^* - \tilde{x}_{it}\beta - L_t\alpha_i$.

The prior distributions for $\rho = (\rho_{uv}, \rho_{u\xi}, \rho_{v\xi})'$ and σ_ξ^2 are chosen to be the conjugate distributions, truncated over the relevant regions. For ρ we have $p(\rho) = N_{[-1,1]}(0, V_\rho)$, a truncated normal distribution between -1 and 1. For σ_ξ^2 we have $p(\sigma_\xi^2) = N_{(0,\infty)}(\mu_{\sigma_\xi}, V_{\sigma_\xi})$, a left truncated normal distribution truncated at 0. The candidate generating function is chosen to be of the autoregressive form, $q(x', x^*) = x^* + v_i$, where v_i is a random normal disturbance. The tuning parameter for ρ and σ_ξ^2 is the variance of v_i's.

Sampling Γ_i*,* Δ_ρ *and* γ:

Recall that the covariance matrix Γ_i has the form given by

$$\begin{aligned}
&\Gamma_i = \text{diag}(g_{i1}, g_{i2}, g_{i3})\Delta_\rho\text{diag}(g_{i1}, g_{i2}, g_{i3})', \quad \text{where} \\
&g_j = \big(\exp(\bar{x}'_{ij}\gamma_j)\big)^{1/2}
\end{aligned} \tag{46}$$

and Δ_ρ is the correlation matrix given in (20). As in the sampling of Σ, we have to use the M-H algorithm. The sampling mechanism is similar to the sampling of Σ. The only difference is that now we sample elements of γ and Δ_ρ, conditional on each other, and the rest of the elements of θ.

The part of the conditional likelihood that involves Γ_i is

$$l(\Gamma_i|\alpha_i) \propto \prod_{i=1}^{n} |\Gamma_i|^{-\frac{1}{2}} \times \exp\{\sum_{i=1}^{N} \alpha_i \Gamma_i^{-1} \alpha_i'\},$$

and the prior distributions of γ and elements of Δ_ρ are taken to be $N_K(0, V_\gamma)$ and $N_{[-1,1]}(0, V_\delta)$, respectively.

Table 1: Summary Statistics for the PSID Extract for Selected Years, 1975–1992

	Variable	Year								
		1975	1978	1980	1982	1984	1986	1988	1990	1992
Individual and Family Characteristics:										
	Observations	3,385	4,019	4,310	4,423	4,451	4,569	4,668	5,728	5,397
1.	Education	11.5894	11.7158	11.8872	12.0242	12.1550	12.4064	12.4820	12.0997	12.1077
		(3.3173)	(3.1556)	(3.0371)	(2.9673)	(2.9293)	(2.9643)	(2.8865)	(3.2887)	(3.3082)
2.	Experience	20.8083	20.7275	20.6316	21.0707	21.3302	21.5714	21.9206	22.5134	24.1688
		(14.3636)	(14.4777)	(14.4416)	(14.3433)	(14.1476)	(13.9300)	(13.7175)	(13.2659)	(12.8766)
3.	Seniority	5.1522	5.5258	5.4344	5.5998	5.9928	6.2579	6.5876	6.6993	7.2854
		(7.6857)	(7.5439)	(7.3769)	(7.3118)	(7.1423)	(7.2307)	(7.4338)	(7.3835)	(7.5628)
4.	Participation	0.9424	0.9211	0.9151	0.8965	0.8809	0.8796	0.8706	0.8785	0.8616
		(0.2330)	(0.2696)	(0.2788)	(0.3047)	(0.3239)	(0.3254)	(0.3357)	(0.3267)	(0.3454)
5.	Mobility	0.2355	0.1411	0.1406	0.0916	0.0865	0.0895	0.0848	0.0826	0.0650
		(0.4243)	(0.3481)	(0.3477)	(0.2884)	(0.2811)	(0.2855)	(0.2787)	(0.2753)	(0.2466)
6.	Log Wage	9.8694	9.8991	9.8874	9.8893	9.9023	9.9809	9.9615	9.9180	9.9237
		(0.7885)	(0.8191)	(0.8144)	(0.8908)	(0.8767)	(0.8103)	(0.8053)	(0.7879)	(0.8422)
7.	Black	0.3090	0.3269	0.3339	0.3242	0.3197	0.3193	0.3186	0.2627	0.2535
		(0.4622)	(0.4692)	(0.4717)	(0.4681)	(0.4664)	(0.4663)	(0.4660)	(0.4402)	(0.4350)
8.	Hispanic	0.0360	0.0351	0.0323	0.0335	0.0335	0.0320	0.0311	0.0690	0.0712
		(0.1864)	(0.1840)	(0.1767)	(0.1799)	(0.1799)	(0.1759)	(0.1735)	(0.2534)	(0.2571)
9.	Family other income	0.6709	0.9232	1.2189	1.7336	2.0422	2.3115	2.5908	2.6488	2.9604
		(1.9259)	(2.6099)	(3.7967)	(6.3152)	(7.0928)	(6.5999)	(7.2393)	(10.8527)	(13.4377)
10.	No. of children	1.3495	1.2070	1.1316	1.0757	1.0420	1.0230	1.0021	1.0918	1.1073
		(1.2246)	(1.1567)	(1.1158)	(1.0792)	(1.0434)	(1.0485)	(1.0457)	(1.1891)	(1.1664)
11.	Children 1 to 2	0.2160	0.2195	0.2311	0.2293	0.2096	0.1996	0.1877	0.1969	0.1731
		(0.3254)	(0.3477)	(0.3756)	(0.3769)	(0.3637)	(0.3622)	(0.3541)	(0.4012)	(0.3697)
12.	Children 3 to 5	0.2245	0.2152	0.2123	0.2037	0.2125	0.2020	0.1982	0.1903	0.1964
		(0.3330)	(0.3597)	(0.3609)	(0.3565)	(0.3675)	(0.3672)	(0.3664)	(0.3868)	(0.3855)
13.	Married	0.8541	0.8074	0.7947	0.7746	0.7890	0.7779	0.7652	0.7699	0.7771
		(0.4821)	(0.4923)	(0.4958)	(0.4958)	(0.4971)	(0.4977)	(0.4979)	(0.4964)	(0.4989)

Table 1: (Continued)

Variable	Year								
	1975	1978	1980	1982	1984	1986	1988	1990	1992
Cohort Effects:									
14. Age 15 or less in 1975	0	0.0926	0.1564	0.2044	0.2557	0.3171	0.3680	0.5120	0.5157
	(0)	(0.2899)	(0.3633)	(0.4033)	(0.4363)	(0.4654)	(0.4823)	(0.4999)	(0.4998)
15. Age 16 to 25 in 1975	0.2177	0.2613	0.2735	0.2745	0.2685	0.2519	0.2397	0.1861	0.1864
	(0.4128)	(0.4394)	(0.4458)	(0.4463)	(0.4432)	(0.4342)	(0.4270)	(0.3892)	(0.3895)
16. Age 26 to 35 in 1975	0.2960	0.2478	0.2204	0.2028	0.1867	0.1725	0.1590	0.1231	0.1236
	(0.4566)	(0.4318)	(0.4146)	(0.4021)	(0.3897)	(0.3778)	(0.3657)	(0.3286)	(0.3291)
17. Age 36 to 45 in 1975	0.1725	0.1421	0.1246	0.1139	0.1054	0.0974	0.0898	0.0697	0.0699
	(0.3779)	(0.3492)	(0.3303)	(0.3178)	(0.3071)	(0.2965)	(0.2859)	(0.2546)	(0.2549)

Table 2: Participation by Education Groups

	Variable	High School Dropouts				High School Graduates				College Graduates			
		Mean	Std	95% CI		Mean	Std	95% CI		Mean	Std	95% CI	
				Low	Up			Low	Up			Low	Up
1.	Constant	-0.5069	0.3465	-1.1918	0.1756	0.1846	0.2959	-0.4114	0.7463	-1.2454	0.6504	-2.5136	-0.0171
2.	Education	-0.0168	0.0148	-0.0466	0.0121	0.0462	0.0139	0.0182	0.0732	0.1302	0.0272	0.0765	0.1844
3.	Lagged Experience	-0.0291	0.0425	-0.1113	0.0553	-0.0890	0.0273	-0.1429	-0.0350	-0.0682	0.0612	-0.1871	0.0524
4.	Lagged Exp.^{2}/100	0.1795	0.2643	-0.3463	0.6912	0.7939	0.1913	0.4143	1.1725	0.8065	0.4131	-0.0218	1.6149
5.	Lagged Exp.^{3}/1000	-0.0332	0.0626	-0.1539	0.0905	-0.2510	0.0499	-0.3501	-0.1539	-0.2718	0.1054	-0.4796	-0.0644
6.	Lagged Exp.^{4}/10000	-0.0001	0.0050	-0.0100	0.0095	0.0217	0.0043	0.0132	0.0301	0.0233	0.0090	0.0057	0.0409
7.	Lagged Participation	1.7233	0.0605	1.6083	1.8456	1.6771	0.0506	1.5809	1.7805	1.8903	0.1037	1.6854	2.0899
8.	Lagged Mobility	0.6535	0.1322	0.4068	0.9213	0.4204	0.0764	0.2752	0.5685	0.3722	0.1563	0.0707	0.6923
	Family Variables:												
9.	Family other income	0.0081	0.0064	-0.0042	0.0208	-0.0035	0.0028	-0.0088	0.0019	-0.0025	0.0018	-0.0058	0.0012
10.	No. of Children	0.0695	0.0251	0.0215	0.1179	-0.0185	0.0222	-0.0617	0.0255	0.1929	0.0652	0.0666	0.3200
11.	Children 1 to 2	-0.0537	0.0645	-0.1790	0.0748	-0.1192	0.0481	-0.2137	-0.0239	-0.1575	0.1268	-0.4028	0.0967
12.	Children 3 to 5	-0.1218	0.0636	-0.2477	0.0039	-0.0084	0.0468	-0.1013	0.0830	-0.0618	0.1416	-0.3399	0.2153
13.	Married	0.1837	0.0643	0.0586	0.3081	0.3779	0.0580	0.2639	0.4919	0.1728	0.1196	-0.0618	0.4064
	Location:												
14.	Northeast	-0.2379	0.0838	-0.3972	-0.0688	0.0367	0.0528	-0.0639	0.1427	-0.0337	0.0879	-0.2086	0.1414
15.	North Central	0.0560	0.0690	-0.0839	0.1867	-0.1051	0.0419	-0.1905	-0.0236	0.0838	0.0870	-0.0824	0.2589
16.	South	0.2022	0.0599	0.0891	0.3200	0.0779	0.0408	-0.0044	0.1582	-0.0093	0.0810	-0.1631	0.1512
17.	Living in SMSA	-0.0735	0.0680	-0.2117	0.0610	-0.0698	0.0497	-0.1676	0.0272	-0.0320	0.1128	-0.2498	0.1897
18.	County unemp. rate	-0.0260	0.0088	-0.0436	-0.0091	-0.0309	0.0072	-0.0452	-0.0170	-0.0004	0.0186	-0.0378	0.0363
	Race:												
19.	Black	-0.4640	0.0940	-0.6457	-0.2782	-0.4336	0.0760	-0.5850	-0.2847	-0.2240	0.1725	-0.5496	0.1202
20.	Hispanic	-0.0259	0.1278	-0.2732	0.2276	-0.0506	0.1463	-0.3304	0.2383	-0.7383	0.3131	-1.3587	-0.1397
	Cohort Effects (as of 1975):												
21.	Age 15 or less	1.0897	0.1734	0.7680	1.4515	0.7170	0.1595	0.4116	1.0411	0.8660	0.3417	0.2107	1.5271
22.	Age 16 to 25	1.0594	0.1899	0.6817	1.4279	0.4970	0.1497	0.2018	0.8014	0.3264	0.3590	-0.3619	1.0411
23.	Age 26 to 35	1.1300	0.1884	0.7551	1.4903	0.4845	0.1427	0.2035	0.7786	0.6730	0.3232	0.0333	1.3224
24.	Age 36 to 45	0.7815	0.1507	0.4873	1.0750	0.4613	0.1216	0.2168	0.7014	0.5429	0.2580	0.0475	1.0625

Note: Omitted from the table are the coefficients on the year dummy variables.

Table 3: Mobility by Education Groups

	Variable	High School Dropouts				High School Graduates				College Graduates			
		Mean	Std	95% CI		Mean	Std	95% CI		Mean	Std	95% CI	
				Low	Up			Low	Up			Low	Up
1.	Constant	-1.2010	0.2733	-1.7425	-0.6589	-1.2925	0.1646	-1.6212	-0.9695	-1.2478	0.2664	-1.7682	-0.7267
2.	Education	-0.0004	0.0105	-0.0212	0.0204	0.0013	0.0077	-0.0137	0.0167	-0.0062	0.0104	-0.0266	0.0145
3.	Lagged Experience	0.0021	0.0301	-0.0567	0.0620	-0.0117	0.0142	-0.0397	0.0163	-0.0136	0.0265	-0.0650	0.0385
4.	Lagged Exp.^{2}/100	-0.1860	0.2069	-0.5975	0.2218	-0.1055	0.1147	-0.3326	0.1190	-0.2319	0.2112	-0.6509	0.1781
5.	Lagged Exp.^{3}/1000	0.0606	0.0549	-0.0472	0.1701	0.0482	0.0347	-0.0188	0.1170	0.1027	0.0643	-0.0214	0.2294
6.	Lagged Exp.^{4}/10000	-0.0060	0.0049	-0.0158	0.0035	-0.0052	0.0034	-0.0121	0.0013	-0.0112	0.0065	-0.0240	0.0013
7.	Lagged Seniority	-0.1131	0.0232	-0.1575	-0.0665	-0.1611	0.0123	-0.1849	-0.1367	-0.1494	0.0225	-0.1937	-0.1047
8.	Lagged Sen.^{2}/100	0.7289	0.3060	0.0929	1.3041	1.0345	0.1547	0.7275	1.3353	1.1328	0.3190	0.5021	1.7633
9.	Lagged Sen.^{3}/1000	-0.2151	0.1376	-0.4670	0.0739	-0.2906	0.0639	-0.4143	-0.1616	-0.3928	0.1529	-0.6918	-0.0893
10.	Lagged Sen.^{4}/10000	0.0219	0.0195	-0.0199	0.0566	0.0291	0.0080	0.0130	0.0444	0.0480	0.0227	0.0028	0.0919
11.	Lagged Mobility	-0.6787	0.0693	-0.8171	-0.5475	-0.8537	0.0369	-0.9257	-0.7814	-1.0171	0.0621	-1.1412	-0.8983
	Family Variables:												
12.	Family other income	-0.0343	0.0087	-0.0520	-0.0180	-0.0270	0.0033	-0.0336	-0.0206	-0.0099	0.0021	-0.0141	-0.0058
13.	No. of Children	-0.0062	0.0152	-0.0363	0.0230	0.0086	0.0107	-0.0123	0.0297	-0.0569	0.0200	-0.0966	-0.0187
14.	Children 1 to 2	0.0633	0.0393	-0.0157	0.1390	0.0363	0.0212	-0.0050	0.0776	0.0561	0.0351	-0.0126	0.1254
15.	Children 3 to 5	-0.0168	0.0395	-0.0945	0.0599	0.0039	0.0219	-0.0383	0.0471	-0.0173	0.0391	-0.0937	0.0591
16.	Married	-0.0060	0.0460	-0.0966	0.0838	-0.0913	0.0273	-0.1443	-0.0384	-0.0646	0.0405	-0.1442	0.0159
	Location:												
17.	Northeast	-0.0032	0.0441	-0.0891	0.0840	-0.0374	0.0192	-0.0752	0.0005	-0.0162	0.0258	-0.0670	0.0347
18.	North Central	-0.0354	0.0357	-0.1051	0.0345	-0.0027	0.0170	-0.0355	0.0309	-0.0221	0.0254	-0.0725	0.0273
19.	South	0.0337	0.0303	-0.0259	0.0926	0.0359	0.0153	0.0056	0.0657	0.0175	0.0238	-0.0295	0.0639
20.	Living in SMSA	-0.1303	0.0366	-0.2015	-0.0587	-0.0009	0.0206	-0.0418	0.0396	0.0529	0.0329	-0.0111	0.1184
21.	County unemp. rate	0.0021	0.0067	-0.0110	0.0152	0.0042	0.0037	-0.0030	0.0114	-0.0021	0.0065	-0.0149	0.0108
	Race:												
22.	Black	-0.0270	0.0398	-0.1044	0.0519	-0.0047	0.0226	-0.0491	0.0400	0.1057	0.0460	0.0159	0.1945
23.	Hispanic	0.0016	0.0697	-0.1358	0.1368	0.0180	0.0494	-0.0789	0.1141	-0.0187	0.1068	-0.2308	0.1878
	Cohort Effects (as of 1975):												
24.	Age 15 or less	0.0450	0.1245	-0.2034	0.2869	0.0334	0.0853	-0.1311	0.2015	0.0329	0.1278	-0.2150	0.2859
25.	Age 16 to 25	-0.0334	0.1159	-0.2635	0.1883	0.0196	0.0820	-0.1392	0.1816	0.0397	0.1248	-0.2007	0.2849
26.	Age 26 to 35	0.0232	0.1020	-0.1798	0.2178	0.0348	0.0771	-0.1139	0.1879	0.0425	0.1144	-0.1796	0.2704
27.	Age 36 to 45	-0.0148	0.0800	-0.1708	0.1391	0.0887	0.0668	-0.0398	0.2244	-0.0020	0.1006	-0.1969	0.1953

Note: Omitted from the table are the coefficients on the year dummy variables and the coefficients on the industry dummy variables.

Table 4: Wage Equation by Education Group

	Variable	High School Dropouts				High School Graduates				College Graduates			
		Mean	Std	95% CI		Mean	Std	95% CI		Mean	Std	95% CI	
				Low	Up			Low	Up			Low	Up
1.	Constant	7.6615	0.1393	7.3915	7.9325	7.8419	0.0821	7.6832	8.0057	8.1434	0.1373	7.8739	8.4115
2.	Education	0.0250	0.0058	0.0135	0.0364	0.0479	0.0036	0.0408	0.0549	0.0396	0.0052	0.0293	0.0499
3.	Experience	0.0641	0.0113	0.0416	0.0860	0.0759	0.0046	0.0669	0.0847	0.1118	0.0080	0.0960	0.1274
4.	Exp.^{2}/100	-0.3586	0.0684	-0.4944	-0.2243	-0.4672	0.0323	-0.5299	-0.4041	-0.5773	0.0539	-0.6834	-0.4716
4.	Exp.^{3}/1000	0.0883	0.0164	0.0558	0.1208	0.1231	0.0090	0.1057	0.1406	0.1335	0.0148	0.1044	0.1625
4.	Exp.^{4}/10000	-0.0083	0.0013	-0.0110	-0.0057	-0.0122	0.0008	-0.0138	-0.0106	-0.0130	0.0014	-0.0157	-0.0103
5.	Seniority	0.0722	0.0064	0.0597	0.0849	0.0690	0.0035	0.0622	0.0757	0.0753	0.0064	0.0625	0.0879
6.	Sen.^{2}/100	-0.3010	0.0747	-0.4466	-0.1538	-0.2853	0.0399	-0.3620	-0.2066	-0.3946	0.0785	-0.5484	-0.2412
5.	Sen.^{3}/1000	0.0972	0.0302	0.0370	0.1562	0.0773	0.0159	0.0463	0.1081	0.1340	0.0340	0.0670	0.2010
5.	Sen.^{4}/10000	-0.0112	0.0038	-0.0187	-0.0036	-0.0074	0.0020	-0.0112	-0.0035	-0.0156	0.0046	-0.0249	-0.0064
	Job switch in 1st sample year:												
7.	Job change in 1st year (ϕ_0^s)	-0.0057	0.0617	-0.1248	0.1167	0.0669	0.0270	0.0139	0.1196	0.1606	0.0483	0.0683	0.2588
8.	Lagged Experience (ϕ_0^e)	0.0126	0.0030	0.0066	0.0185	0.0120	0.0019	0.0082	0.0158	0.0077	0.0036	0.0007	0.0150
	Job switches that lasted:												
9.	Up to 1 year (ϕ_{10})	0.0486	0.0226	0.0043	0.0933	0.1092	0.0093	0.0909	0.1275	0.2134	0.0170	0.1793	0.2468
10.	2 to 5 years (ϕ_{20})	0.1366	0.0284	0.0805	0.1921	0.1074	0.0130	0.0815	0.1327	0.1566	0.0190	0.1194	0.1938
11.	6 to 10 years (ϕ_{30})	0.2147	0.0758	0.0661	0.3628	0.1900	0.0420	0.1078	0.2719	0.3265	0.0701	0.1889	0.4640
12.	Over 10 years (ϕ_{40})	0.0565	0.1058	-0.1488	0.2669	0.3730	0.0540	0.2675	0.4788	0.5171	0.0877	0.3452	0.6871
	Seniority of job that lasted:												
13.	2 to 5 years (ϕ_2^s)	0.0264	0.0116	0.0039	0.0488	0.0278	0.0049	0.0183	0.0374	0.0539	0.0071	0.0399	0.0678
14.	6 to 10 years (ϕ_3^s)	0.0110	0.0110	-0.0107	0.0325	0.0181	0.0057	0.0068	0.0293	0.0075	0.0099	-0.0119	0.0271
15.	Over 10 years (ϕ_4^s)	0.0264	0.0051	0.0164	0.0363	0.0346	0.0032	0.0284	0.0408	0.0019	0.0057	-0.0093	0.0130
	Experience of job that lasted:												
16.	Up to1 year (ϕ_1^e)	0.0031	0.0014	0.0003	0.0058	-0.0017	0.0009	-0.0035	0.0001	-0.0061	0.0016	-0.0093	-0.0030
17.	2 to 5 years (ϕ_2^e)	-0.0023	0.0016	-0.0055	0.0009	0.0001	0.0010	-0.0018	0.0021	-0.0051	0.0016	-0.0083	-0.0019
18.	6 to 10 years (ϕ_3^e)	0.0014	0.0023	-0.0030	0.0058	-0.0031	0.0015	-0.0060	-0.0002	-0.0012	0.0024	-0.0061	0.0035
19.	Over 10 years (ϕ_4^e)	0.0021	0.0032	-0.0043	0.0084	-0.0165	0.0021	-0.0206	-0.0124	-0.0040	0.0033	-0.0103	0.0024
	Location:												
20.	Northeast	0.0229	0.0285	-0.0340	0.0780	0.0460	0.0119	0.0220	0.0691	0.0496	0.0153	0.0195	0.0791
21.	North Central	0.0827	0.0243	0.0347	0.1297	-0.0060	0.0102	-0.0256	0.0137	-0.0493	0.0139	-0.0768	-0.0223
22.	South	-0.1010	0.0195	-0.1387	-0.0624	-0.0442	0.0090	-0.0617	-0.0264	-0.0124	0.0132	-0.0385	0.0140
23.	Living in SMSA	0.0804	0.0203	0.0403	0.1200	0.0631	0.0090	0.0453	0.0804	0.0396	0.0126	0.0145	0.0640
24.	County unemp. rate	-0.0017	0.0024	-0.0066	0.0030	-0.0045	0.0013	-0.0070	-0.0020	-0.0039	0.0020	-0.0078	0.0001

Table 4: (Continued)

	Variable	High School Dropouts				High School Graduates				College Graduates			
		Mean	Std	95% CI		Mean	Std	95% CI		Mean	Std	95% CI	
				Low	Up			Low	Up			Low	Up
	Race:												
25.	Black	-0.3120	0.0323	-0.3764	-0.2477	-0.2640	0.0171	-0.2967	-0.2304	-0.2239	0.0360	-0.2968	-0.1517
26.	Hispanic	0.0565	0.0408	-0.0236	0.1371	-0.0821	0.0295	-0.1396	-0.0247	0.0170	0.0727	-0.1246	0.1581
	Cohort effects (as of 1975):												
27.	Age 15 or less	0.5392	0.0716	0.4018	0.6858	0.3716	0.0484	0.2785	0.4689	-0.0102	0.0802	-0.1684	0.1459
28.	Age 16 to 25	0.3952	0.0742	0.2530	0.5447	0.2503	0.0477	0.1584	0.3458	-0.2324	0.0805	-0.3866	-0.0726
29.	Age 26 to 35	0.3719	0.0650	0.2462	0.5021	0.2346	0.0439	0.1512	0.3239	-0.1518	0.0711	-0.2913	-0.0143
30.	Age 36 to 45	0.2753	0.0601	0.1557	0.3976	0.2545	0.0423	0.1723	0.3373	0.0709	0.0688	-0.0623	0.2059

Note: Omitted from the table are the coefficients on the year dummy variables and the coefficients on the industry dummy variables. The coefficients in brackets in lines (7) through (19) are according to the definiton of the J^W function defined in the text

Table 5: Estimated Cumulative and Marginal Returns to Experience

Group	Cumulative Returns				Marginal Returns (in %)			
	Years of Experience				Years of Experience			
	5	10	15	20	5	10	15	20
Panel A: Quartic Model								
HS Dropouts	.241	.362	.410	.420	3.442	1.550	0.482	-0.012
	(.042)	(.062)	(.071)	(.073)	(.597)	(.307)	(.231)	(.232)
HS Graduates	.277	.402	.440	.438	3.776	1.447	0.232	-0.234
	(.017)	(.024)	(.028)	(.031)	(.233)	(.142)	(.135)	(.134)
College Graduates	.430	.661	.762	.786	6.339	3.116	1.117	-0.045
	(.029)	(.044)	(.052)	(.058)	(.421)	(.264)	(.244)	(.244)
Panel B: Quadratic Model								
HS Dropouts	0.101	0.246	0.472	0.678	2.661	1.959	1.256	0.554
	(.005)	(.012)	(.022)	(.028)	(.286)	(.248)	(.215)	(.191)
HS Graduates	0.146	0.253	0.320	0.349	2.526	1.744	0.961	0.179
	(.008)	(.015)	(.022)	(.027)	(.154)	(.135)	(.122)	(.116)
College Graduates	0.256	0.446	0.567	0.622	4.455	3.109	1.763	0.417
	(.015)	(.029)	(.040)	(.051)	(.285)	(.253)	(.231)	(.221)
Panel C: Quartic Model with no J^W Function								
HS Dropouts	0.312	0.489	0.583	0.634	4.684	2.553	1.344	0.774
	(.040)	(.058)	(.065)	(.067)	(.557)	(.281)	(.217)	(.218)
HS Graduates	0.370	0.572	0.674	0.724	5.462	2.853	1.397	0.713
	(.016)	(.023)	(.026)	(.028)	(.214)	(.128)	(.126)	(.125)
College Graduates	0.588	0.939	1.129	1.214	9.105	5.198	2.600	0.911
	(0.028)	(.041)	(.047)	(.051)	(.384)	(.230)	(.219)	(.222)
Panel D: Quadratic Model with no J^W Function								
HS Dropouts	0.211	0.383	0.515	0.607	3.830	3.034	2.238	1.442
	(.013)	(.025)	(.036)	(.044)	(.253)	(.221)	(.193)	(.174)
HS Graduates	0.237	0.423	0.558	0.643	4.231	3.214	2.197	1.181
	(.007)	(.014)	(.020)	(.025)	(.140)	(.125)	(.114)	(.109)
College Graduates	0.401	0.710	0.928	1.053	7.105	5.263	3.420	1.578
	(.013)	(.024)	(.035)	(.044)	(.244)	(.220)	(.203)	(.197)

Table 6: Estimated Cumulative and Marginal Returns to Seniority

Group	Cumulative Returns				Marginal Returns (in %)			
	Years of Seniority				Years of Seniority			
	2	5	10	15	2	5	10	15
Panel A: Quartic Model								
HS Dropouts	.133	.297	.507	.677	6.128	4.882	3.669	3.242
	(.010)	(.019)	(.024)	(.029)	(.408)	(.236)	(.249)	(.238)
HS Graduates	.127	.283	.475	.616	5.850	4.590	3.215	2.555
	(.006)	(.011)	(.014)	(.018)	(.227)	(.141)	(.150)	(.150)
College Graduates	.136	.294	.477	.615	6.109	4.513	3.035	2.630
	(.010)	(.019)	(.026)	(.034)	(.412)	(.257)	(.279)	(.272)
Panel B: Quadratic Model								
HS Dropouts	0.151	0.266	0.347	0.392	4.966	4.721	4.314	3.906
	(.015)	(.029)	(.040)	(.050)	(.253)	(.216)	(.169)	(.156)
HS Graduates	0.094	0.228	0.435	0.619	4.612	4.347	3.906	3.464
	(.003)	(.008)	(.013)	(.018)	(.153)	(.132)	(.108)	(.105)
College Graduates	0.097	0.236	0.449	0.637	4.775	4.485	4.001	3.517
	(.006)	(.014)	(.025)	(.034)	(.281)	(.247)	(.209)	(.205)
Panel C: Quartic Model with no J^W Function								
HS Dropouts	0.099	0.204	0.304	0.370	4.296	2.774	1.479	1.302
	(.010)	(.018)	(.021)	(.022)	(.385)	(.195)	(.209)	(.200)
HS Graduates	0.092	0.184	0.256	0.280	3.918	2.303	0.778	0.369
	(.006)	(.010)	(.012)	(.013)	(.214)	(.110)	(.122)	(.126)
College Graduates	0.097	0.177	0.212	0.216	3.838	1.675	0.095	0.308
	(.010)	(.017)	(.020)	(.021)	(.381)	(.179)	(.196)	(.195)
Panel D: Quadratic Model with no J^W Function								
HS Dropouts	0.054	0.131	0.251	0.359	2.653	2.510	2.273	2.035
	(.005)	(.011)	(.018)	(.022)	(.216)	(.177)	(.125)	(.114)
HS Graduates	0.042	0.102	0.192	0.268	2.070	1.915	1.656	1.397
	(.003)	(.006)	(.010)	(.013)	(.121)	(.099)	(.074)	(.076)
College Graduates	0.027	0.067	0.133	0.198	1.342	1.331	1.313	1.294
	(.004)	(.010)	(.016)	(.020)	(.201)	(.162)	(.119)	(.130)

Table 7: Estimates of the Stochastic Elements by Education Groups

	Variable	High School Dropouts				High School Graduates				College Graduates			
		Mean	St.dev.	Range		Mean	St.dev.	Range		Mean	St.dev.	Range	
				Min	Max			Min	Max			Min	Max
	Covariance Matrix of White Noises (element of Σ)												
1.	ρ_{uv}	-0.0012	0.0098	-0.0201	0.0180	-0.0019	0.0072	-0.0145	0.0134	-0.0021	0.0103	-0.0222	0.0178
2.	$\rho_{u\xi}$	-0.0251	0.0069	-0.0390	-0.0124	-0.0323	0.0060	-0.0431	-0.0218	-0.0299	0.0135	-0.0538	-0.0032
3.	$\rho_{v\xi}$	0.0100	0.0071	-0.0040	0.0249	0.0064	0.0049	-0.0042	0.0150	0.0210	0.0094	0.0027	0.0386
4.	σ^2_ξ	0.2954	0.0038	0.2882	0.3029	0.2086	0.0016	0.2054	0.2117	0.2048	0.0023	0.2004	0.2093
	Correlations of Individual Specific Effects (elements of Δ_ρ)												
5.	$\rho_{\alpha_y\alpha_m}$	-0.1835	0.1762	-0.5327	0.1487	-0.5024	0.1033	-0.6203	-0.2394	-0.5521	0.0746	-0.6723	-0.4115
6.	$\rho_{\alpha_y\alpha_w}$	0.3572	0.0368	0.2827	0.4235	0.3574	0.0262	0.3017	0.4070	0.2007	0.0508	0.1002	0.3061
7.	$\rho_{\alpha_m\alpha_w}$	-0.1846	0.3343	-0.8439	0.5321	-0.6980	0.1589	-0.9089	-0.4161	-0.7682	0.0739	-0.8729	-0.6083
	Coefficients on Individual Specific Effects in Initial Condition Equations												
8.	δ_y	1.0223	0.0318	0.9609	1.0858	1.0047	0.0241	0.9574	1.0518	1.0019	0.0334	0.9370	1.0684
9.	δ_m	1.5997	4.2696	-8.4120	12.1475	0.9155	0.2572	0.4026	1.4106	1.2460	0.2809	0.7177	1.8419

Table 8: Comparison of Estimates to Other Studies

Type of Analysis	**OLS**				**Altonji-Williams**				**Topel two-step**			
Group	All	<12	12–15	>16	All	<12	12–15	>16	All	<12	12–15	>16
Coefficients:												
Linear tenure	0.0493	0.0579	0.0438	0.0397	0.0398	0.0455	0.0382	0.0374	0.0673	0.0702	0.0601	0.0547
	(.0022)	(.0057)	(.0027)	(.0049)	(.0045)	(.0061)	(.0033)	(.0057)	(.0031)	(.0061)	(.0034)	(.0057)
Linear experience	0.0542	0.0589	0.0512	0.0590	0.0501	0.0547	0.0492	0.0579	0.0643	0.0685	0.0523	0.0614
	(.0033)	(.0097)	(.0041)	(.0077)	(.0043)	(.0106)	(.0048)	(.0142)	(.0042)	(.0104)	(.0043)	(.0081)
Cumulative return to tenure:												
2 years	0.0856	0.0994	0.0779	0.0683	0.0528	0.0621	0.0486	0.0427	0.0973	0.1136	0.0885	0.0779
	(.0035)	(.0088)	(.0043)	(.0074)	(.0038)	(.0100)	(.0045)	(.0082)	(.0039)	(.0102)	(.0048)	(.0082)
5 years	0.1727	0.1974	0.1629	0.1359	0.0988	0.1117	0.0921	0.0776	0.1843	0.2098	0.1733	0.1453
	(.0059)	(.0150)	(.0073)	(.0120)	(.0062)	(.0169)	(.0082)	(.0136)	(.0068)	(.0172)	(.0085)	(.0139)
10 years	0.2442	0.2729	0.2433	0.1887	0.1195	0.1311	0.1189	0.0915	0.2323	0.2584	0.2304	0.1781
	(.0065)	(.0164)	(.0082)	(.0131)	(.0072)	(.0188)	(.0087)	(.0149)	(.0077)	(.0188)	(.0093)	(.0151)
15 years	0.2700	0.3004	0.2773	0.2075	0.1164	0.1310	0.1202	0.0899	0.2343	0.2599	0.2405	0.1807
	(.0071)	(.0165)	(.0090)	(.0146)	(.0077)	(.0178)	(.0096)	(.0155)	(.0083)	(.0182)	(.0101)	(.0167)
20 years	0.2862	0.3278	0.2896	0.2228	0.1287	0.1460	0.1298	0.0992	0.2469	0.2832	0.2501	0.1926
	(.0080)	(.0168)	(.0104)	(.0176)	(.0084)	(.0191)	(.0110)	(.0198)	(.0089)	(.0190)	(.0124)	(.0198)
Cumulative return to experience:												
5 years	0.2210	0.2319	0.2039	0.2552	0.2278	0.2397	0.2100	0.2651	0.2281	0.2401	0.2110	0.2636
	(.0114)	(.0349)	(.0139)	(.0263)	(.0128)	(.0377)	(.0148)	(.0289)	(.0131)	(.0412)	(.0157)	(.0313)
10 years	0.3611	0.3683	0.3273	0.4418	0.3997	0.4072	0.3624	0.4885	0.3983	0.4052	0.3598	0.4892
	(.0155)	(.0498)	(.0184)	(.0354)	(.0166)	(.0523)	(.0201)	(.0406)	(.0184)	(.0588)	(.0211)	(.0424)
30 years	0.5054	0.5370	0.4513	0.6982	0.5847	0.6286	0.5223	0.7756	0.5239	0.5716	0.5025	0.6995
	(.0139)	(.0467)	(.0167)	(.0333)	(.0150)	(.0512)	(.0180)	(.0356)	(.0153)	(.0538)	(.0188)	(.0392)

Figure 1: High School Graduates—Wage Growth Due to Experience and Seniority

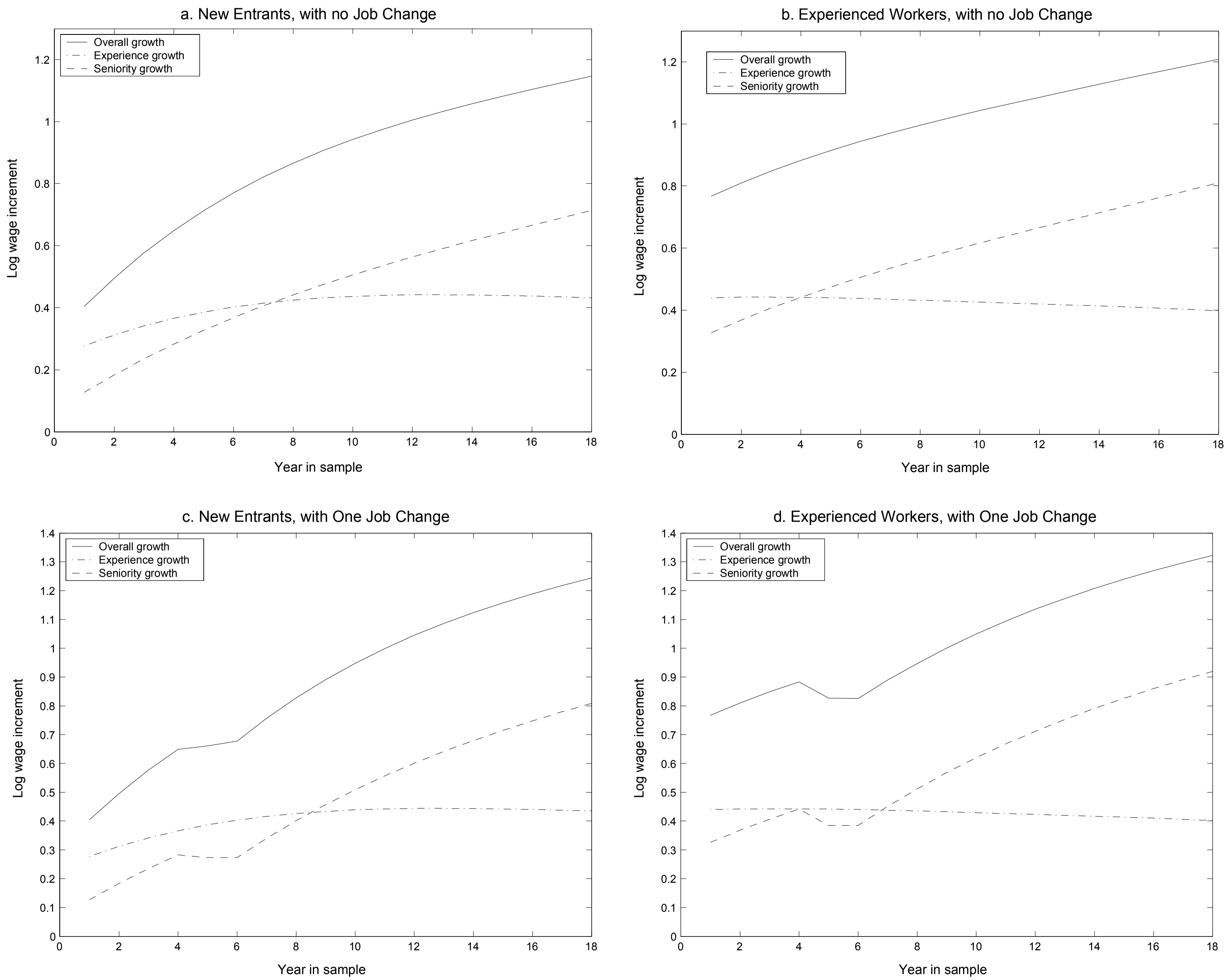

Figure 2: College Graduates—Wage Growth Due to Experience and Seniority