INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2002-29

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

M. BUCHINSKY[1] **D. FOUGERE**[2]
F. KRAMARZ[3] **R. TCHERNIS**[4]

[1] UCLA, CREST-INSEE and NBER.
[2] CNRS, CREST-INSEE, CEPR and IZA.
[3] CREST-INSEE, CEPR and IZA.
[4] Brown University, USA.

Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.[1]

Abstract. This paper presents new estimates of the returns of experience and seniority on individual wages. In contrast to all previous studies, we explicitly model the participation and mobility decisions. Our theoretical model gives rise to a statistical model with 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) to compute the posterior joint distribution of the model's parameters. We find that the effects of seniority and experience differ for all education groups. Our modeling strategy also allows us to examine the individuals' "optimal" mobility patterns for maximizing their wage growth over their lifetime. We find that the optimal job durations differ markedly across education groups.

Résumé. Cet article présente de nouvelles estimations des rendements de l'expérience et de l'ancienneté d'emploi sur les salaires individuels. A la différence de toutes les études antérieures consacrées à ce sujet, nous modélisons ici de manière explicite les décisions individuelles de participation et de mobilité inter-entreprises. Notre modèle théorique produit un modèle statistique à trois équations: (1) une équation de participation, (2) une équation de salaire, et (3) une équation de mobilité inter-entreprises. Dans ce modèle, l'équation de salaire est estimée simultanément avec ces deux dernières équations. Chaque équation inclut un effet aléatoire individuel et un choc idiosyncratique. Les équations de participation et de mobilité incluent par ailleurs les choix passés dans la liste des variables explicatives. Nous utilisons le Panel Study of Income Dynamics (PSID) pour estimer le modèle sur trois strates, correspondant à des niveaux d'éducation différents: (1) les individus qui ont abandonné leurs études avant la fin de leur scolarité secondaire, (2) les diplômés du secondaire, et (3) les diplômés de l'enseignement supérieur (niveau "college" aux Etats-Unis). Nous adoptons une approche bayésienne pour l'estimation du modèle. Pour être plus précis, nous utilisons des méthodes de type MCMC pour faire une inférence sur la distribution jointe a posteriori des paramètres du modèle. Nous trouvons que les effets de l'ancienneté et de l'expérience varient significativement d'un niveau d'éducation à l'autre. Notre stratégie de modélisation nous permet par ailleurs d'examiner les trajectoires de mobilité "optimales", c'est-à-dire celles qui maximisent la progression salariale au cours de la vie active. Nos résultats montrent que les durées optimales d'emploi, ainsi définies, diffèrent fortement selon le niveau d'éducation.

[1]†We thank Andrew Foster, Oded Galor, Tony Lancaster, Magali Ponçon-Beffy, Jean-Marc Robin, and Bernard Salanié for their comments and discussions. Participants of seminars and conferences in Amsterdam, Brown, Crest (Malinvaud seminar), ESEM 2001 in Lausanne, Harvard, Pennsylvania, Tel-Aviv, Toulouse, 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. Only recently, has research 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, because measures of inequality alone are not sufficient to assess changes in the wage determination process. For example, it is vital that we have more information in order for us to be able to understand changes in the respective roles of general- and firm-specific human capital. This is an especially important issue when examining, for example, technological changes induced by computerization and globalization.

While wage inequality increased in the United States during most of the 1980s, in France and some other European countries it was generally stable. Nevertheless, during the same period almost all countries witnessed a sharp decrease in wage mobility.[2] This decline in wage mobility indicates that changes in wage inequality may be worse than previously thought. Furthermore, workers are more likely to be in a worse situation if there is an increase instability of jobs, as has been documented in the United States for males (see, again, Farber (1999), for a discussion of the evidence). Decreased wage mobility and increased job instability makes increasing wage inequality (as in the U.S.) or a high unemployment rate (as in France) less tolerable 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), in that particular firm. 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 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. The basic statistical 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) to compute the posterior joint distribution of the model's parameters.

Our main finding is that returns to seniority are quite high for all education groups. In contrast, the returns to experience appear to be lower than previously thought. While we use a somewhat different sample than the one used by Topel (1991), the results we obtain regarding the return seniority are, qualitatively, similar to Topel's results, while the results for the return to experience differ, most probably because experience is endogenous in our approach. Consequently, our estimate of total within-job growth is lower than Topel's estimate, but closer to other analyses reported in the literature (e.g. Altonji and Shakotko (1987), Abraham and Farber (1987), and Altonji and Williams (1997)). However, our statistical assumptions – by endogenizing both mobility and experience – incorporate elements on

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

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 (2001) and Neal (1995) who analyzed similar questions but took different routes. Specifically, they did not explicitly model the decisions of the individuals that are directly related to the observed endogenous variables. Nevertheless, Dustmann and Meghir (2001) employed data and econometric methods which allowed them to identify the various components of wage growth, namely general, sector-specific, and firm-specific human capital. In contrast to all previous studies, in the current study we explicitly model the participation and mobility decisions. We find that the effects of seniority and experience differ for all education groups. However, in this study, unlike other studies in the literature, our modeling strategy also allows us to examine the individuals' "optimal" mobility patterns for maximizing their wage growth over their lifetime.[3] We find that the optimal job durations differ markedly across education groups.

The remainder of the paper is organized as follows. In Section 2, we outline the model and 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. Section 3 presents the details of our numerical techniques for computing the posterior distribution of the model's parameters. A brief discussion of the data extract used in this study is provided in Section 4. Section 5 presents the empirical results of the estimation procedure. Section 6 follows with a brief conclusion.

2 The Model and the Econometric Specification

2.1 The Model

We consider an extension of the simple dynamic programming model of search behavior proposed by Hyslop (1999). We extend Hyslop's model by allowing a worker to move directly from one job to another.

A worker who is currently employed in a firm receives each period two wage offers without searching: one offer comes from the current employer, while the other comes from another firm.[4] If an interfirm mobility occurs at the end of period $t-1$ the worker incurs the cost c_M, paid at the beginning of period t. While all other assumptions are identical to Hyslop's (1999), some modification of the notation are needed.

A nonparticipant may search for a job at a cost γ_1 per period, paid at the beginning of the next period. We assume that hours of work are constant accross 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 an intertemporally separable utility function, where the current period flow utility is defined over consumption C_t and leisure $l_t = 1 - y_t$. The expected present value of discounted utility over an infinite lifetime is therefore

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

where $u(\cdot,\cdot;\cdot)$ is a single period flow utility, while X_t is a vector of exogenous (observed and unobserved) individual characteristics in period t. The term β is simply the discount factor. 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 nonlabor income, w_t is the individual's wage. The variable m_t is a dummy variable that takes the value 1 if the individual moves between jobs at the end of period t, and takes the value 0 otherwise.

[3] By "optimal path" we mean that it is the path that would have maximized the wage growth, had it been followed.

[4] It is to incorporate a situation in which a participant gets an outside wage offer only if he or she undertakes an on-the-job search, with cost equal to γ_2. However, since the main results would not be affected, we simplify here the presentation.

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 defined as

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

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

$$V_t\left(0,0; X_t\right) = \max_{\{y_t, m_t\}} \left[V_t^0\left(0,0;X_t\right), V_t^1\left(0,0;X_t\right), V_t^2\left(0,0;X_t\right)\right] \tag{4}$$

where

$$\begin{aligned}
V_t^0\left(0,0;X_t\right) &= u\left(z_t - \gamma_1, 0; X_t\right) + \beta E_t V_{t+1}\left(0,0;X_{t+1}\right), \\
V_t^1\left(0,0;X_t\right) &= u\left(z_t + w_t - \gamma_1, 1; X_t\right) + \beta E_t V_{t+1}\left(1,0;X_{t+1}\right), \quad \text{and} \\
V_t^2\left(0,0;X_t\right) &= u\left(z_t + w_t - \gamma_1, 1; X_t\right) + \beta E_t V_{t+1}\left(1,1;X_{t+1}\right).
\end{aligned}$$

The quantity $V_t^0\left(0,0;X_t\right)$ denotes the value of the nonparticipation state in period t, the quantity $V_t^1\left(0,0;X_t\right)$ denotes the value of participating, without moving, in period t, while $V_t^2\left(0,0;X_t\right)$ denotes the value of participating and moving in period t.

For a "stayer", namely a participant who stays in the firm at the end of period $t-1$, the value function at the beginning of period t is

$$V_t\left(1,0;X_t\right) = \max_{y_t, m_t} \left\{V_t^0\left(1,0;X_t\right), V_t^1\left(1,0;X_t\right), V_t^2\left(1,0;X_t\right)\right\}, \tag{5}$$

where

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

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

$$V_t\left(1,1;X_t\right) = \max_{y_t, m_t} \left\{V_t^0\left(1,1;X_t\right), V_t^1\left(1,1;X_t\right), V_t^2\left(1,1;X_t\right)\right\}, \tag{6}$$

where

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

Transitions to a nonparticipation 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 in the nonparticipation state with the value functions in the participation states (without or with interfirm mobility). For a nonparticipant in period $t-1$, these two reservation wages, denoted by $w^*_{01,t}$ and $w^*_{02,t}$ respectively, are defined implicitly by

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

It follows immediately from this system of equations that

$$w^*_{02,t} \lesseqgtr w^*_{01,t} \quad \Leftrightarrow \quad E_t V_{t+1}\left(1,0;X_{t+1}\right) \lesseqgtr E_t V_{t+1}\left(1,1;X_{t+1}\right). \tag{8}$$

Note that if $w^*_{02,t} < w^*_{01,t}$, then the decision rule for a nonparticipant is to accept any wage offer greater than $w^*_{02,t}$, and to eventually move to another firm in the next period. If $w^*_{01,t} < w^*_{02,t}$, then the optimal strategy for a nonparticipant is to accept any wage offer greater than $w^*_{01,t}$, and to stay in the firm at least for one more period.[5]

The corresponding reservation wages for a stayer, denoted $w^*_{11,t}$ and $w^*_{12,t}$, are defined by the following system of equations:

$$V_t^0\left(1,0;X_t\right) = V_t^1\left(1,0;X_t \mid w^*_{11}\left(t\right)\right) = V_t^2\left(1,0;X_t \mid w^*_{12}\left(t\right)\right). \tag{9}$$

For a mover, these reservation wages, denoted $w^*_{21,t}$ and $w^*_{22,t}$, are defined by the following system of equations:

$$V_t^0\left(1,1;X_t\right) = V_t^1\left(1,1;X_t \mid w^*_{21}\left(t\right)\right) = V_t^2\left(1,1;X_t \mid w^*_{22}\left(t\right)\right). \tag{10}$$

A comparison of reservation wage expressions for nonparticipants and stayers, from (7) and (9), implies that

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

A first-order Taylor series expansions of the left and right hand sides of (11) around $z_t + w^*_{01,t}$, $z_t + w^*_{02,t}$ and z_t, respectively, gives

$$\begin{aligned} w^*_{1j,t} &\approx w^*_{0j,t} - \gamma_1\left[1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{0j,t}, 1; X_t\right)}\right] \\ &\approx w^*_{0j,t} - \gamma_{1j}, \qquad \text{for } j = 1, 2, \end{aligned} \tag{12}$$

where $u'(\cdot)$ denotes is the marginal utility of consumption, and

$$\gamma_{1j} = \gamma_1\left[1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{0j,t}, 1; X_t\right)}\right], \qquad \text{for } j = 1, 2.$$

As noticed by Hyslop (1999), if utility 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, namely if

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

then γ_{1j} may be positive or negative. More precisely,

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

If the marginal utility of consumption is lower when working, then γ_{1j} is always negative, because

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

Similarly, a comparison of reservation wages for nonparticipants and movers, from (7) and (10), implies that

$$\begin{aligned} u\left(z_t + w^*_{21,t} - c_M, 1; X_t\right) - u\left(z_t + w^*_{01,t} - \gamma_1, 1; X_t\right) & \\ = u\left(z_t + w^*_{22,t} - c_M, 1; X_t\right) - u\left(z_t + w^*_{02,t} - \gamma_1, 1; X_t\right), & \\ = u\left(z_t - c_M, 0; X_t\right) - u\left(z_t - \gamma_1, 0; X_t\right). & \end{aligned} \tag{13}$$

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

Again, first-order Taylor series expansions of the left and right hand sides of (13) around $z_t + w^*_{0j,t}$, $z_t + w^*_{2j,t}$ $(j = 1, 2)$ and z_t, respectively, give

$$\begin{aligned} u\left(z_t + w^*_{2j,t} - c_M, 1; X_t\right) &\simeq u\left(z_t + w^*_{2j,t}, 1; X_t\right) - c_M\ u'\left(z_t + w^*_{2j,t}, 1; X_t\right), \quad (14) \\ u\left(z_t + w^*_{0j,t} - \gamma_1, 1; X_t\right) &\simeq u\left(z_t + w^*_{0j,t}, 1; X_t\right) - \gamma_1\ u'\left(z_t + w^*_{0j,t}, 1; X_t\right), \\ u\left(z_t - c_M, 0; X_t\right) &\simeq u\left(z_t, 0; X_t\right) - c_M\ u'\left(z_t, 0; X_t\right), \\ u\left(z_t - \gamma_1, 0; X_t\right) &\simeq u\left(z_t, 0; X_t\right) - \gamma_1\ u'\left(z_t, 0; X_t\right), \quad \text{and} \\ u\left(z_t + w^*_{2j,t}, 1; X_t\right) - u\left(z_t + w^*_{0j,t}, 1; X_t\right) &\simeq \left(w^*_{2j,t} - w^*_{0j,t}\right)\ u'\left(z_t + w^*_{0j,t}, 1; X_t\right) \ \text{for } j = 1, 2. \end{aligned}$$

Substitution of the expressions from (14) back into (13) gives

$$w^*_{2j,t} \approx w^*_{0j,t} - \gamma_{1j} + \gamma_{2j} \approx w^*_{1j,t} + \gamma_{2j}, \qquad \text{for } j = 1, 2, \tag{15}$$

where

$$\gamma_{2j} = c_M \left[\frac{u'\left(z_t + w^*_{2j,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{0j,t}, 1; X_t\right)}\right], \qquad \text{for } j = 1, 2,$$

and

$$\gamma_{1j} = \gamma_1 \left[1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{0j,t}, 1; X_t\right)}\right], \qquad \text{for } j = 1, 2.$$

Note that if the marginal utility of consumption is greater when working, namely, if

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

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

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

Suppose now that the mobility cost c_M is strictly greater than the search cost γ_1. It is possible to show then (see the Appendix) that, if $w^*_{02,t} < w^*_{01,t}$, there exist two sets of sufficient conditions under which the participation and mobility equations exhibit first-order state dependence. These sufficient conditions are the following:

Condition 1:

When the marginal utility of consumption is higher when working, we must verify that

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

and that

$$0 < \gamma_{12} < \gamma_{22} < \gamma_{21}.$$

Condition 2:

When the marginal utility of consumption is lower when working, we must verify that

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

and that[6]

$$\gamma_{12} < \gamma_{22} < \gamma_{21} < 0.$$

[6] We prove in the appendix that, in this case, γ_{22} is greater than γ_{12} if and only if

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

This last condition is stronger than $c_M > \gamma_1$, because the term between brackets is strictly greater than 1.

Under either condition 1 or condition 2, a mover (i.e., a worker who move from one firm to another at the end of period $t-1$) becomes nonparticipant at the end of period t if he or she is offered a wage less than $w^*_{22,t}$ where

$$w^*_{22,t} > w^*_{02,t} > w^*_{12,t}.$$

A stayer becomes nonparticipant if he or she is offered a wage less than $w^*_{12,t}$. Thus, the participation decision at period t can be characterized by the equation

$$\begin{aligned} y_t &= \mathbf{1}\left[w_t > w^*_{02,t} - \gamma_{12}\ y_{t-1} + \gamma_{22}\ y_{t-1}\ m_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{02,t} + \gamma_{12}\ y_{t-1} - \gamma_{22}\ y_{t-1}\ m_{t-1} > 0\right] \end{aligned} \tag{16}$$

with $\gamma_{22} > \gamma_{12} > 0$, and where $\mathbf{1}(\cdot)$ denotes the usual indicator function.

A stayer can accept to move to another firm at the end of period t 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 will decide to move again only if the wage offer in t is at least $w^*_{22,t}$. Consequently, the mobility decision at period t can be characterized by the equation

$$\begin{aligned} m_t &= \mathbf{1}\left[w_t > w^*_{12,t} + \gamma_{22}\ y_{t-1}\ m_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{12,t} - \gamma_{22}\ y_{t-1}\ m_{t-1} > 0\right] \end{aligned} \tag{17}$$

with $\gamma_{22} > 0$.

Substitution of the wage function, i.e. the equation for w_t into (16) and (17), gives the equations which are the basis for our econometric specification as detailed below.

2.2 The Econometric Specification

We use a statistical model that is suited to the incorporation of key elements that are important to labor markets and wage setting outcomes. The model consists of three equations. The first equation is a participation equation, reflecting the individual's choice of whether or not to participate in the labor market. The second equation is a mobility equation describing the individual's decision to switch from one firm to another. Finally, a log wage equation specifies individuals' annual earnings function.[7]

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 will become clear from the specifications below.

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

$$\begin{aligned} y_{it} &= \mathbf{1}(y^*_{it} \geq 0), \\ y^*_{it} &= x'_{yit}\beta_0 + \beta_y y_{i,t-1} + \beta_m m_{i,t-1} + \alpha_{yi} + u_{it}, \end{aligned} \tag{18}$$

where y^*_{it} denotes a latent variable that depends on x_{yit}, the observable characteristics for the ith individual at time t. Among other things x_{yit} includes education and actual lagged labor market experience (and its square). This last variable is constructed from the individual sequence of y_{it}. The term α_{yi} is a person specific random effect, while u_{it} a contemporaneous error term. The notation $\mathbf{1}(\cdot)$ is the usual indicator function, that is, y_{it} denotes whether worker i participated at date t. Note also that the equation includes the past realizations of the participation and the mobility processes.[8]

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

$$\begin{aligned} m_{it} &= \mathbf{1}(m^*_{it} \geq 0) \times \mathbf{1}(y_{i,t-1} = 1, y_{it} = 1), \\ m^*_{it} &= x'_{mit}\lambda_0 + \lambda_m m_{i,t-1} + \alpha_{mi} + v_{it}, \end{aligned} \tag{19}$$

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 equation (18))

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

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

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). 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 (18) and (19), is that, by definition, one cannot be mobile at date t unless he/she participates at both dates $t-1$ and t.

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

$$\begin{aligned} w_{it} &= w^*_{it} \times \mathbf{1}(y_{it} = 1), \\ w^*_{it} &= J^W_{it} + x'_{wit}\delta_0 + \alpha_{wi} + \xi_{it}, \end{aligned} \tag{20}$$

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 specific random effect, while ξ_{it} is a contemporaneous error term. Finally, the term J^W_{it} denotes the sum of all wage changes that resulted from the moves that occurred before 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,

$$J^W_{it} = (\phi^s_0 + \phi^e_0 e_{i0})\, d_{i1} + \sum_{l=1}^{M_{it}} \left[\sum_{j=1}^{4} \left(\phi_{j0} + \phi^s_j s_{t_l - 1} + \phi^e_j e_{t_l - 1}\right) d_{jit_l} \right]. \tag{21}$$

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 1 and 5 years, and equals 0 otherwise, $d_{3t_l} = 1$ if the lth job lasted between 5 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. In particular, both the participation and mobility equation 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.

At this point, 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. The only differences come from differences in observables. And, using the conceptual framework of the evaluation literature, there is no way we can compute the expected wage of the treated (say those who indeed move) had they not been treated (i.e. had they stayed in their origin firm). Each worker in our statistical model draws a unique idiosyncratic random term in the wage equation. Indeed, had we two equations as in Robinson (1989) for the union and non-union sector, the construction of a counterfactual wage would have been possible. However, in our approach, the specification of two equations, one for the origin firm, 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. Another solution would be to draw one shock for each firm. But, in the PSID, this is unfeasible since individual firms cannot be identified. Hence, in our results section, we will describe movers in terms of their origin and destination industries, wage changes, J^W_{it} function.

[9] This specification for the term J^W_{it} produces thirteen different regressors in the wage equation (20). These regressors are: a dummy for job change in year 1, experience in year 0, the numbers of switches of jobs that lasted less than 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.

2.3 Stochastic Assumptions

In this subsection, we specify the stochastic structure of the random terms in equations (18)–(20) and provide the distributional assumptions for the random terms.

First, the individual specific effects are stochastically independent of the time-varying shocks, that is

$$(\alpha_{yi}, \alpha_{mi}, \alpha_{wi}) \perp (u_{it}, v_{it}, \xi_{it}).$$

Furthermore, we assume that $(\alpha_{yi}, \alpha_{mi}, \alpha_{wi})$ are correlated individual specific effects, with

$$(\alpha_{yi}, \alpha_{mi}, \alpha_{wi}) \sim N(0, \Omega),$$

where

$$\Omega = \begin{pmatrix} \sigma^2_{\alpha_y} & \rho_{\alpha_y\alpha_m}\sigma_{\alpha_y}\sigma_{\alpha_m} & \rho_{\alpha_y\alpha_w}\sigma_{\alpha_y}\sigma_{\alpha_w} \\ \rho_{\alpha_y\alpha_m}\sigma_{\alpha_y}\sigma_{\alpha_m} & \sigma^2_{\alpha_m} & \rho_{\alpha_m\alpha_w}\sigma_{\alpha_m}\sigma_{\alpha_w} \\ \rho_{\alpha_y\alpha_w}\sigma_{\alpha_y}\sigma_{\alpha_w} & \rho_{\alpha_m\alpha_w}\sigma_{\alpha_m}\sigma_{\alpha_w} & \sigma^2_{\alpha_w} \end{pmatrix}.$$

Here, we allow $\sigma^2_{\alpha_y}$, $\sigma^2_{\alpha_m}$, and $\sigma^2_{\alpha_w}$, and consequently Ω, to be heteroskedastic, i.e., the variances are allowed to depend on x_{yit}, x_{mit}, and x_{wit}, that is,

$$\begin{aligned} \sigma^2_{\alpha_y} &= \exp\left(h_1\left(\gamma_y, x_{yi1}, ..., x_{yiT}\right)\right), \\ \sigma^2_{\alpha_m} &= \exp\left(h_2\left(\gamma_m, x_{mi1}, ...x_{miT}\right)\right), \quad \text{and} \\ \sigma^2_{\alpha_w} &= \exp\left(h_3\left(\gamma_w, x_{wi1}, ..., x_{wiT}\right)\right), \end{aligned} \tag{22}$$

for some real valued functions $h_1(\cdot)$, $h_2(\cdot)$, and $h_3(\cdot)$.

Note that the above specification has direct implications for the correlation between the regressor vectors and the person specific random effects. To see this, consider those employees that participate and have $s_{it} = s$, which is similar to $m_{it-s-1} = 1$, $m_{it-s} = 0$, $m_{it-s+1} = 0, \cdots$, $m_{it-1} = 0$, $m_{it} = 0$, $y_{it-s-1} = 1$, $y_{it-s} = 1$, $y_{it-s+1} = 1, \cdots$, $y_{it-1} = 1$, $y_{it} = 1$, or rewrite as (all this is conditional on $x's$ and $\alpha's$):

$$\begin{aligned} l(m_{it-s-1} &= 1, m_{it-s} = 0, \cdots, m_{it} = 0, y_{it-s-1} = 1, y_{it-s} = 1, \cdots, y_{it} = 1) \\ &= l\left(m_{it} = 0, y_{it} = 1 \mid m_{it-1} = 0, y_{it-1} = 1\right) \times \cdots \\ &\quad \times l\left(m_{it-s+1} = 0, y_{it-s+1} = 1 \mid m_{it-s} = 0, y_{it-s} = 1\right) \\ &\quad \times l\left(m_{it-s} = 0, y_{it-s} = 1 \mid m_{it-s-1} = 1, y_{it-s-1} = 1\right) \\ &\quad \times l\left(m_{it-s-1} = 1, y_{it-s-1} = 1\right) \end{aligned}$$

because of the first-order dependence of the mobility and participation processes. From this, we see that $s_{it} = (s_{it-1} + 1) \times 1(m_{it} = 0, y_{it} = 1)$ and by recursion

$$\begin{aligned} s_{it} &= s_{i0} \times \prod_{k=0}^{t} 1\left(m_{ik} = 0, y_{ik} = 1\right) \\ &\quad + \sum_{k=0}^{t-1} \left[1\left(m_{it} = 0, y_{it} = 1\right) \times \cdots \times 1\left(m_{it-k} = 0, y_{it-k} = 1\right)\right] \end{aligned}$$

Therefore, in the wage equation, for those workers that participate, the seniority component s_{it} is correlated with the person-specific effect of the wage equation α_{wi} through the correlations of α_{mi} and α_{yi}, the person-specific effects of the mobility equation and of the participation equation respectively. Similarly, experience as well as J^W_{it}, which are both present in the wage equation, can be shown to be correlated with the person-specific effect α_{wi} affecting this same wage equation, because they include, albeit in a complex fashion, the person-specific effects from the mobility and participation equation, α_{mi} and α_{yi}, which are in our statistical model correlated to α_{wi}. Hence, our statistical model allows correlated random effects in the wage equation between observables and the unobserved person-specific component.

Finally, the idiosyncratic error components $(u_{it}, v_{it}, \xi_{it})$ are assumed to be contemporaneously correlated white noises. Specifically, we assume that

$$\tau_{it} \equiv (u_{it}, v_{it}, \xi_{it})' \sim N(0, \Sigma), \tag{23}$$

where

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

Note that for identification reasons, we set $\sigma_u^2 = \sigma_v^2 = 1$.

Our analysis departs from the existing literature on the return to seniority in a number of crucial ways.[10] The most important deviation is that we explicitly model the participation and mobility decisions. In order to discuss this relation more precisely, the next paragraphs describe the routes taken by previous research.

2.4 Related Literature

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). Therefore, we first describe Topel's strategy. Then, we briefly discuss Altonji and Shakotko as well as Abraham and Farber (1987). Finally, we spend some time on the more recent article of Altonji and Williams (1997) who revisit very carefully this debate. We also discuss related articles.

Starting from the evidence that the costs of displacement are strongly related to prior job tenure, Topel (1991) singles out two potential explanations: wages rise with seniority; tenure acts as a proxy for the quality of the job (the job was well paid all along). He apparently discards the third hypothesis that workers with long tenures are more able. His base equation is:

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

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 where the residual ϵ_{ijt} can be decomposed as

$$\epsilon_{ijt} = \phi_{ijt} + \mu_i + v_{ijt} \tag{25}$$

with the first component specific to the work pair, the second being ability, the last is for marketwide random shocks and measurement error. In particular, an endogeneity problem arises if the first component ϕ_{ijt} is correlated with experience or tenure. This leads Topel to rewrite the first component as

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

Topel makes the following reasoning. If $\beta_2 > 0$ some workers have rejected offers. Since they stay this reduces the average wage of stayers. And similarly, those who move received large wage offers and this raises the average wage of movers. Hence, this generates a negative b_2. Topel therefore stresses upon the selection effects in who are the movers and who are the stayers. He also notes that mobility costs strengthen the bias (those who move are even more selected among high-wage workers).

The estimation technique that Topel suggests is a two-stage procedure. First, he looks at first difference for the stayers:

$$\Delta y_{ijt} = \beta_1 + \beta_2 + \Delta\epsilon_{ijt}$$

[10] In particular, see Farber (1999) for a comprehensive survey of evidence in the literature.

and he writes the initial wage in the job as

$$y_{ij0} = X_{ij0}\beta_1 + \phi_{ij0} + \mu_i + v_{ij0}$$

So, if one estimates the first difference equation and get $\widehat{\beta_1 + \beta_2} = B$, then $y_{ijt} - T_{ijt}B = X_{ij0}\beta_1 + e$. To have no endogeneity bias, one needs $EX_0\epsilon = 0$ or using the first and the third equations, one obtains:

$$E\widehat{\beta_1} = \beta_1 + b_1 + \gamma_{X_0T} \times (b_1 + b_2)$$

where γ_{ZY} is the coefficient of the regression of Y on Z, and

$$E\widehat{\beta_2} = \beta_2 - b_1 - \gamma_{X_0T} \times (b_1 + b_2)$$

It is crucial to note at this point that, in Topel's approach, experience at entry in a job is exogenous and uncorrelated with the error terms.

Finally, Topel notes that if jobs with large wage growth survive more than jobs with low wage growth, the growth in the stayers population may be larger than in the total, and then β_2 would be overestimated. Also, if those who change jobs often are less productive, then those who stay long have a low initial X_0 and therefore β_1 would be biased downwards, and consequently β_2 would be overestimated.

The data used in Topel (1991) is the PSID (from 1968 to 1983). The wage measure is the log average hourly earnings in the calendar year. Unfortunately, the cross-sections of the PSID are not representative of the population. Furthermore, since a fixed population is followed, the time trend is not exogenous (because the average match quality rises with time in the market). Therefore, Topel chooses to deflate the wage data by an index. As for the econometric methodology, Topel claims that when wage growth follows a random walk $\phi_{ijt} = \phi_{ijt-1} + \eta_{ijt}$, the estimates based on first difference are unbiased, and, indeed, this is what he finds.

The application of his method yields the following results. Topel finds $\beta_2 = .0545$ and hence 10 years of seniority imply a 28% wage increase. He also estimates $b_1 + b_2 = .0020$. According to Topel, the above results constitute a lower bound on the average returns to job seniority. But, there are potential biases in the estimates of $\beta_1 + \beta_2$:

a) the 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 growth component (potentially known to the workers). By looking at jobs that he kwows will survive longer, and comparing the resulting estimates of $\beta_1 + \beta_2$, Topel finds no difference.

b) more able or more productive workers may be less mobile. He therefore examines the maintained hypothesis that μ_i person's ability is unrelated to tenure. In fact, when the best people move less, then the potential bias can be seen from the folllowing equation:

$$E\widehat{\beta_2} = \beta_2 - b_1 - \gamma_{X_0T} \times (b_1 + b_2) - (\text{coefficient of the regression of } \mu \text{ on } X_0)$$

To evaluate this bias, Topel needs an instrument. The variable, uncorrelated with μ_i but correlated with X_0, that he suggests is total experience X. Hence, he assumes that the distribution of ability is unrelated to experience. Under this maintained hypothesis, Topel can use IV to estimate the second stage of his approach:

$$E\widehat{\beta_2}^{IV} = \beta_2 - b_1 - \frac{\gamma_{XT}}{1 - \gamma_{XT}} \times (b_1 + b_2)$$

Since he finds that 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) apply from the beginning an instrumental variables technique assuming that $\phi_{ijt} = \phi_{ij}$, i.e. is time-invariant. Hence, $\sum_t (T_{ijt} - T_{ij})(\phi_{ij} + \mu_i) = 0$ and the deviation of seniority from the average seniority in that job is a valid instrument in the level equation. Topel (1991) shows that this IV technique is a variant of his two-step approach. The IV estimate of $\beta_1 + \beta_2$ is similar (taking within-job differences instead of first differences). The resulting estimate of β_1 becomes

$$E\widehat{\beta_1}^{IV} = E\widehat{\beta}_1^{2S} + \left[\frac{\gamma_{XT}}{1-\gamma_{XT}} - \gamma_{X_0T}\right] \times (b_1 + b_2)$$

Empirically the coefficient of the regression of T on X_0 is equal to $-.25$. Therefore AS' procedure appears to induce a downward bias. Another reason is measurement error for downward biases. Finally, the treatment of time is another difference between these two approaches. Topel uses a specific index for the aggregate changes in real wages based on the CPS. AS use a time trend. If there is growth of quality of the jobs because of better matches across time, then this growth causes an additional downward bias in returns to tenure. Abraham and Farber (1987) have different assumptions and use completed tenure to proxy unobserved dimensions of persons or jobs quality, but many have censored spells. They also use a wage equation that is quadratic in experience and linear in tenure.

Recently Altonji and Williams (1997) have set up their model following Topel's. Their article is an in-depth attempt to outline the consequences of various modelling assumptions on the estimated returns to seniority. First, they write the person specific effect μ_i (see above) as

$$\mu = Xc_1 + Tc_2 + \omega$$

with $cov(\mu, X) = 0$ and

$$T = \mu d_1 + Xd_2 + \nu$$

Since $cov(\mu, X) = 0$, then d_2 is the coefficient of the regression of T on X , denoted γ_{XT}. Then, they show that $c_1 = -\gamma_{XT}c_2 < 0$ and

$$c_2 = d_1 \frac{var(\mu)}{d_1^2 var(\mu) + var(\nu)} > 0$$

Note also that $\gamma_{X_0\mu} = c_2 \dfrac{var(T \mid X)}{var(T)}$. This is in contrast with Topel's assumptions ($EX_0\epsilon = 0$). Our approach is therefore much closer in spirit to AW's since we model experience and allow it to be correlated to wage through its various components, especially the person-specific effect (see the above discussion on correlated effects).

Altonji and Williams (AW, hereafter) replicate Topel's results. They first focus on the specification of the time trends. Topel claims that time trends are problematic because they can be correlated with ϕ when workers have had more time to locate high ϕ jobs. In addition, time may be correlated with μ because of changes in the sample. AW claim that the covariance between t and ϕ is 0 conditional on X, T and X or T and X_0. AW conclude that a way to circumvent the above problems is the use of t in deviation from the person's mean. In addition, AW do not appear to be convinced that the specification of the time trend is a serious problem.

AW test the impact of using the instrumented time trend in place of the Murphy-Welch index based on the CPS that Topel uses. They find no effect on OLS results but find substantial difference in the IV used in AS, i.e. by using the deviation of seniority from mean in the job as an instrument. In addition, they find a small but sizeable upward bias in Topel's estimator.[11]

AW look also at the impact of measurement error in the estimated returns. Remember first that Topel reports that AS's IV estimator goes from .052 to .161 when one moves from AS's measure of tenure to his measure. Note also that AS found that measurement induced changes in the estimated returns without altering the qualitative nature of their conclusions (from .027 to .041).

[11] 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, when there is a job change, all measures of earnings are mixtures of the old and the new job compensation. AW test using year t wages with year t tenure. All available estimates of the returns to seniority decline (say at 10 years, from .223 to .161). In addition, AW examines the impact of the functional form used for tenure in the estimation. AS use a quadratic function plus an indicator for $T > 1$. Topel (1991) uses a quartic (note that, for Topel, given his dating conventions, T is always greater than 1). AW find that the functional form does not matter much.

To conclude, AW summarize the differences in the various estimators that they have examined. First, it appears that all estimators are biased upwards by job match heterogeneity (i.e. correlation between T and ϕ). Accounting for this correlation plus the timing issue, estimates drop from .223 to .126 (at 10 years of seniority). In addition, if one uses the time trend instead of the CPS-based index to deflate and instrument X_0 with X, the effect falls further to .044 (to be compared with .161 when X_0 is considered exogenous). One conclusion is that individual heterogeneity is important (in the growth) and that a large part of the reduction in the (upward bias) in Topel's estimator is due to a reduction in the bias from individual heterogeneity. As a conclusion, AW' best guess for the returns to 10 years of seniority is equal to .11. Interestingly for us, when AW uses more recent data, the returns seem to have increased.

In a recent paper, Dustmann and Meghir (2001) allow for returns to experience, returns to sector-specific seniority, and returns to firm-specific seniority. Those returns may be respectively person-specific, and match specific for the last one. To estimate returns to experience they focus on the new jobs of those displaced. They assume that workers cannot predict closure of an establishment (more than a year in advance), workers and firms have similar information on the quality of the match, for those displaced who start a new job the unobservables that govern preferences for work are the same as those for sector choice. Hence, controlling for the endogeneity of experience also controls for the endogeneity of sector tenure. Finally, conditional on experience, potential experience and education, age does not affect offered wages of young workers.

This implies estimating two reduced-form equations for experience at the beginning of the period and for the participation equation. From these estimates, they get residuals of these equations and include them in the wage equation for the first wage record of the job after displacement. In this equation the residuals from the two previous reduced-form equations are introduced and interacted with potential experience, general experience, sector tenure. Of course this equation has no firm-specific seniority in it (it is equal to zero by construction). A similar procedure is used for estimating returns to seniority. The equivalent problem being selection from staying in the firm. The data are German (IAB) and from the US (NLSY79). Their results show that the returns to tenure for skilled workers are large (2% a year for the first five years), slightly smaller for the unskilled. They also find evidence of heterogeneous returns to tenure. Returns to sector tenure are rather low even though they are significant. Returns to experience are larger than for tenure (4% a year for at least 5 years) and are smaller for the unskilled. Finally, they find that, for both countries, between-job wage growth is larger than within-job wage growth.

The above papers focus on returns to tenure without explicitly specifying the mobility process that generates the observed seniority. However, 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 and then this probability decreases. To understand this process, person heterogeneity and duration dependence must be distinguished. If only pure heterogeneity prevails, the number of jobs previously held is a sufficient statistics for the probability of change. Controlling for experience and the number of prior job changes, the mobility should not be related any more to tenure in this situation. Farber gives evidence that contradict the simple model of pure heterogeneity (see Farber, 1994). Notice also that Farber (1999) 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 literature that are of interest for us. In particular, if specific capital matters, the firm will choose to lay off less senior workers, and Farber found support for this view. Furthermore, job losses result in permanent and substantial earnings loss. In addition, the results in the literature that he surveys show that those with more tenure lose more. Selection of the laid off workers does not seem to explain strong differences (see Gibbons and Katz, 1991). One potential explanation is that some jobs are high-wage jobs. Then average tenure will be higher on those jobs because of the reduced probability of quits. If more stable workers are more productive then long tenure in the past job commands high-reemployment wage. This question of the relation between tenure and the re-employment wage is indeed pervasive in this literature, even though its joint modeling is never explicitly undertaken. Our J^W function allows us to tackle this issue directly. The inclusion in the J_{it}^W function of seniority at the end of the last job is also motivated by the literature on displaced workers. For example, Addison and Portugal (1989) show that wage losses are larger for displaced workers with more tenure (see also Jacobson, LaLonde, and Sullivan (1993) as well as Farber (1999)). The inclusion of job market experience at the previous job as a determinant of the earnings change in the J_{it}^W function 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. Similarly, the inclusion in the J_{it}^W function of the number of past mobilities and the seniority 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 on the related and very important question of sector or firm specific knowledge. Unfortunately, the choice of sector or of firm is left exogenous in our approach. To model the first choice, an additional "change of sector" equation, similar to the mobility equation, should be added to our system of equations; this would make estimation of our model even more difficult. 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 on returns to seniority nor on firm-to-firm mobility, we must mention the recently published article by Geweke and Keane (Geweke and Keane, 2000). These authors include two equations in their system, one for earnings and one for the marital status. Earnings are modelled as an autoregressive process, with a person specific effect and AR(1) process for the idiosyncratic term. The marital status equation has a latent variable that depends on the realization of the status at $t-1$ as well as earnings at $t-1$. From the estimates, they construct a simulated file with the observed characteristics of their sampled individuals. They use their estimates to examine earnings mobility (using quintiles), the present value of lifetime earnings. They show that non-gaussian distributions of shocks do fit the data better than gaussian ones. For instance, the college premium is greater under the non-gaussian model than under the normal one (by one fifth). Their estimation technique shares many features with ours. In addition, the data set that they use, the PSID, is similar to the one we use.

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. In addition, our wage equation specifies quite precisely the re-employment wage through the J^W function, function that depends on past job outcomes.

2.5 The Likelihood Function

In this subsection, we present the likelihood function for our problem. We first specify the likelihood function, conditional on the individual specific effects, and then integrate it with respect to the distribution of the individual specific effects. For convenience of notation, let $\alpha_i^1 = (\alpha_{yi}, \alpha_{mi}, \alpha_{wi})$, and let $x_{it} = (x_{yit}, x_{mit}, x_{wit})$. Conditional on the individual specific effects, the individual's likelihood function is given by

$$\begin{aligned} l\left\{(y_{it}, m_{it}, w_{it})_{t=1,...T} \mid \alpha_i^1, x_{it}\right\} &= l\left\{(y_{iT}, m_{iT}, w_{iT}) \mid \alpha_i^1, x_{it}, \left(y_{iT-1}, m_{iT-1}, J_{iT}^W\right)\right\} \\ &\times l\left\{(y_{iT-1}, m_{iT-1}, w_{iT-1}) \mid \alpha_i^1, x_{it}, \left(y_{iT-2}, m_{iT-2}, J_{iT-1}^W\right)\right\} \\ &\times ... \times l\left\{w_{i1} \mid (y_{i1}, m_{i1}), \alpha_i^1, x_{it}\right\} \times l\left\{y_{i1}, m_{i1}\right\}. \end{aligned}$$

Note that the last term of the right hand side of (26) is the likelihood for the initial state (at time $t=1$) of the ith individual, that is the likelihood of (y_{i1}, m_{i1}). Following Heckman (1981), we approximate this part of the likelihood by a probit specification given by

$$\begin{aligned} y_{i1} &= 1(y_{i1}^* \geq 0), \\ \text{with } y_{i1}^* &= a x_{i1} + \alpha_{yi}^0 + u_{i1}, \end{aligned} \tag{26}$$

and

$$\begin{aligned} m_{i1} &= 1(m_{i1}^* \geq 0) \times 1(y_{i1} = 1), \\ \text{with } m_{i1}^* &= b x_{i1} + \alpha_{mi}^0 + v_{i1}. \end{aligned} \tag{27}$$

The random terms α_{yi}^0 and α_{mi}^0 are assumed to be normally distributed random variables, with mean 0. Furthermore, they are allowed to be correlated with the fixed individual specific components $(\alpha_{yi}, \alpha_{mi}, \alpha_{wi})$. Consequently we assume that

$$\alpha_i \equiv (\alpha^0_{yi}, \alpha^0_{mi}, \alpha_{wi}, \alpha_{yi}, \alpha_{mi}) \sim N(0, \Gamma),$$

where

$$\Gamma = \begin{pmatrix} \sigma^2_{\alpha^0_y} & \rho_{\alpha^0_y \alpha^0_m}\sigma_{\alpha^0_y}\sigma_{\alpha^0_m} & \rho_{\alpha_w \alpha^0_y}\sigma_{\alpha_w}\sigma_{\alpha^0_y} & \rho_{\alpha_y \alpha^0_y}\sigma_{\alpha_y}\sigma_{\alpha^0_y} & \rho_{\alpha_m \alpha^0_y}\sigma_{\alpha_m}\sigma_{\alpha^0_y} \\ \rho_{\alpha^0_y \alpha^0_m}\sigma_{\alpha^0_y}\sigma_{\alpha^0_m} & \sigma^2_{\alpha^0_m} & \rho_{\alpha_w \alpha^0_m}\sigma_{\alpha_w}\sigma_{\alpha^0_m} & \rho_{\alpha_y \alpha^0_m}\sigma_{\alpha_y}\sigma_{\alpha^0_m} & \rho_{\alpha_m \alpha^0_m}\sigma_{\alpha_m}\sigma_{\alpha^0_m} \\ \rho_{\alpha_w \alpha^0_y}\sigma_{\alpha_w}\sigma_{\alpha^0_y} & \rho_{\alpha_w \alpha^0_m}\sigma_{\alpha_w}\sigma_{\alpha^0_m} & \sigma^2_{\alpha_w} & \rho_{\alpha_y \alpha_w}\sigma_{\alpha_y}\sigma_{\alpha_w} & \rho_{\alpha_m \alpha_w}\sigma_{\alpha_m}\sigma_{\alpha_w} \\ \rho_{\alpha_y \alpha^0_y}\sigma_{\alpha_y}\sigma_{\alpha^0_y} & \rho_{\alpha_y \alpha^0_m}\sigma_{\alpha_y}\sigma_{\alpha^0_m} & \rho_{\alpha_y \alpha_w}\sigma_{\alpha_y}\sigma_{\alpha_w} & \sigma^2_{\alpha_y} & \rho_{\alpha_y \alpha_m}\sigma_{\alpha_y}\sigma_{\alpha_m} \\ \rho_{\alpha_m \alpha^0_y}\sigma_{\alpha_m}\sigma_{\alpha^0_y} & \rho_{\alpha_m \alpha^0_m}\sigma_{\alpha_m}\sigma_{\alpha^0_m} & \rho_{\alpha_m \alpha_w}\sigma_{\alpha_m}\sigma_{\alpha_w} & \rho_{\alpha_y \alpha_m}\sigma_{\alpha_y}\sigma_{\alpha_m} & \sigma^2_{\alpha_m} \end{pmatrix}. \tag{28}$$

As for $\sigma^2_{\alpha_y}$, and $\sigma^2_{\alpha_m}$ in (22) we allow $\sigma^2_{\alpha^0_y}$ and $\sigma_{\alpha^0_m}$ to be heteroskedastic, that is

$$\begin{aligned} \sigma^2_{\alpha^0_y} &= \exp\left(h_4(\gamma_{y^0}, x_{yi1}, ..., x_{yiT})\right), \\ \text{and } \sigma^2_{\alpha^0_m} &= \exp\left(h_5\left(\gamma_{m^0}, x_{mi1}, ...x_{miT}\right)\right), \end{aligned} \tag{29}$$

for some real valued functions $h_4(\cdot)$ and $h_5(\cdot)$. Note that each individual in the sample has (potentially) different Γ, say Γ_i, that is

$$\alpha_i \sim N(0, \Gamma_i). \tag{30}$$

For convenience we rewrite Γ as

$$\Gamma = D\Delta_\rho D, \tag{31}$$

where D is a diagonal matrix of the form

$$D = \text{diag}\left(\sigma_{\alpha^0_y}, \sigma_{\alpha^0_m}, \sigma_{\alpha_w}, \sigma_{\alpha_y}, \sigma_{\alpha_m}\right) \tag{32}$$

and

$$\Delta_\rho = \begin{pmatrix} 1 & \rho_{\alpha^0_y \alpha^0_m} & \rho_{\alpha_w \alpha^0_y} & \rho_{\alpha_y \alpha^0_y} & \rho_{\alpha_m \alpha^0_y} \\ \rho_{\alpha^0_y \alpha^0_m} & 1 & \rho_{\alpha_w \alpha^0_m} & \rho_{\alpha_y \alpha^0_m} & \rho_{\alpha_m \alpha^0_m} \\ \rho_{\alpha_w \alpha^0_y} & \rho_{\alpha_w \alpha^0_m} & 1 & \rho_{\alpha_y \alpha_w} & \rho_{\alpha_m \alpha_w} \\ \rho_{\alpha_y \alpha^0_y} & \rho_{\alpha_y \alpha^0_m} & \rho_{\alpha_y \alpha_w} & 1 & \rho_{\alpha_y \alpha_m} \\ \rho_{\alpha_m \alpha^0_y} & \rho_{\alpha_m \alpha^0_m} & \rho_{\alpha_m \alpha_w} & \rho_{\alpha_y \alpha_m} & 1 \end{pmatrix}. \tag{33}$$

Furthermore, we simplify the variances in (22) and (29) to be only a function of the average of the regressors over the sample years.[12] In generic form we have then

$$h_j(\gamma, x_1, ..., x_T) = \bar{x}'\gamma_j, \quad j = 1, ..., 5, \tag{34}$$

where $\bar{x} = (\sum_{i=1}^{T} x_t)/T$. Also we define $\gamma = (\gamma_1', \gamma_2', \gamma_3', \gamma_4', \gamma_5')'$.

Given the above assumptions, the form of the individual's conditional likelihood, given the individual observable characteristics and unobservable individual-specific effects, is given by

$$\begin{aligned} l\left\{(y_{it}, m_{it}, w_{it})\left|\alpha^1_{yi}, y_{i,t-1}, m_{i,t-1}, J^W_{it}\right.\right\} &= \{1 - \Phi(x'_{yit}\beta_0 + \beta_y y_{i,t-1} + \beta_m m_{i,t-1} + \alpha_{yi})\}^{1-y_{it}} \\ &\times \left\{\sigma_\xi^{-1} \times \varphi(\xi_{it} \times \sigma_\xi^{-1})\right\}^{y_{it}} \\ &\times \{\Phi(B_{it}) - \Phi_2\left(A_{it}, B_{it}, R\right)\}^{y_{it} \times (1-m_{it})} \\ &\times \{1 - \Phi\left(A_{it}\right) - \Phi\left(B_{it}\right) + \Phi_2\left(A_{it}, B_{it}, R\right)\}^{y_{it} \times m_{it}} \end{aligned} \tag{35}$$

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

for $t = 2, ...T$, where Φ and φ are the cdf and the density function, respectively, of a standard normal variable,

$$
\begin{aligned}
\xi_{it} &= w_{it} - \{J_{it}^{W} + x'_{wit}\delta_0 + \alpha_{wi}\}, \\
A_{it} &= -\left(x'_{yit}\beta_0 + \beta_y y_{i,t-1} + \beta_m m_{i,t-1} + \alpha_{yi} + \xi_{it}\frac{\rho_{u\xi}}{\sigma_\xi}\right) / \sqrt{1-\rho_{u\xi}^2}, \\
B_{it} &= -\left(x'_{mit}\lambda_0 + \lambda_m m_{i,t-1} + \alpha_{mi} + \xi_{it}\frac{\rho_{v\xi}}{\sigma_\xi}\right) / \sqrt{1-\rho_{v\xi}^2}, \\
\Phi_2(A, B, R) &= \int_{-\infty}^{A}\int_{-\infty}^{B} \frac{1}{2\pi\sqrt{1-R^2}} \times \exp\left(-\frac{x^2 - 2Rxy + y^2}{2(1-R^2)}\right) dxdy, \\
\text{and } R &= \frac{\rho_{uv} - \rho_{u\xi} \times \rho_{v\xi}}{\sqrt{\left(1-\rho_{u\xi}^2\right) \times \left(1-\rho_{v\xi}^2\right)}}.
\end{aligned} \tag{36}
$$

Note that for derivation of the likelihood in (36), we used the fact that

$$
(u_{it}, v_{it}) \mid \xi_{it} \sim N\left[\begin{pmatrix} \xi_{it}\rho_{u\xi}/\sigma_\xi \\ \xi_{it}\rho_{v\xi}/\sigma_\xi \end{pmatrix}, \begin{pmatrix} 1-\rho_{u\xi}^2 & \rho_{uv} - \rho_{u\xi}\rho_{v\xi} \\ \rho_{uv} - \rho_{u\xi}\rho_{v\xi} & 1-\rho_{v\xi}^2 \end{pmatrix}\right].
$$

Similarly, the likelihood function for the initial state is given by

$$
\begin{aligned}
l\{w_{i1} | y_{i1}, m_{i1}, \alpha_i)\} \, l\{y_{i1}, m_{i1}\} &= \{1 - \Phi(x'_{yi1}a + \alpha_{yi}^0)\}^{1-y_{i1}} \left\{\sigma_\xi^{-1} \times \varphi(\xi_{i1} \times \sigma_\xi^{-1})\right\}^{y_{i1}} \\
&\quad \times \{\Phi(B_{i1}) - \Phi_2(A_{i1}, B_{i1}, R)\}^{y_{i1} \times (1-m_{i1})} \\
&\quad \times \{1 - \Phi(A_{i1}) - \Phi(B_{i1}) + \Phi_2(A_{i1}, B_{i1}, R)\}^{y_{i1} \times m_{i1}},
\end{aligned}
$$

where

$$
\begin{aligned}
\xi_{i1} &= w_{i1} - \{J_{i1}^{W} + x'_{wi1}\delta_0 + \alpha_{wi}\}, \\
A_{i1} &= -\left(x'_{yi1}a + \alpha_{yi}^0 + \xi_{i1}\frac{\rho_{u\xi}}{\sigma_\xi}\right) / \sqrt{1-\rho_{u\xi}^2}, \\
\text{and } B_{i1} &= -\left(x'_{mi1}b + \alpha_{mi}^0 + \xi_{i1}\frac{\rho_{v\xi}}{\sigma_\xi}\right) / \sqrt{1-\rho_{v\xi}^2}.
\end{aligned}
$$

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

$$
\begin{aligned}
l\left\{(y_{it}, m_{it}, w_{it})_{t=1,...T}\right\} &= \int \left[\prod_{t=2}^{T} l\left\{(y_{it}, m_{it}, w_{it}) \mid \alpha_i, x_{it}, \left(y_{it-1}, m_{i,t-1}, J_{it-1}^{W}\right)\right\}\right] \\
&\quad \times l\{w_{i1} \mid (y_{i1}, m_{i1}), \alpha_i, x_{it},\} \times l\{y_{i1}, m_{i1}\} \\
&\quad \times (2\pi)^{-5/2} |\Gamma_i|^{-1/2} \exp\left[-0.5 \times (\alpha_i)'\Gamma_i^{-1}(\alpha_i)\right] d\alpha_i.
\end{aligned}
$$

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.[13]

[13] 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).

3 Computation of 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., $\theta = \{\beta, \alpha, \Sigma, \Delta_\rho, \gamma\}$, 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 (18), (19), and (20), 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^*).[14]

Note that in matrix form we can write the model in (18), (19), and (20) as

$$z^*_{it} = \tilde{x}_{it}\beta + L_t\alpha_i + \tau_{it}, \tag{37}$$

for $t = 1, ..., T$, where $\alpha_i \sim N(0, \Gamma_i)$, as is defined in (30), $\tau_{it} \sim N(0, \Sigma)$, as defined in (23),

$$\tilde{x}_{i1} = \begin{pmatrix} x_{yi1} & 0 & 0 & 0 & 0 \\ 0 & x_{mi1} & 0 & 0 & 0 \\ 0 & 0 & x_{wi1} & 0 & 0 \end{pmatrix},$$

$$\tilde{x}_{it} = \begin{pmatrix} 0 & 0 & 0 & x_{yit} & 0 \\ 0 & 0 & 0 & 0 & x_{mit} \\ 0 & 0 & x_{wit} & 0 & 0 \end{pmatrix}, \quad \text{for } t > 1,$$

$$L_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad \text{and}$$

$$L_t = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad \text{for } t > 1.$$

For clarity of presentation we define a few other quantities as follows. The parameter vector β consists of the regression coefficients in (18), (19), and (20), including the parameters from the function J^W_{it} defined in (21), and the parameter vectors from the initial condition equations (26) and (27). The parameter vector γ consists of the coefficients in (34). Note that the covariance matrix for α_i, Γ_i, is constructed from γ and Δ_ρ defined in (31), (32), and (34). 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.[15] Below we explain the sampling of each of the parts in θ (augmented by z^*), conditional on all the other parts and the data.

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 limited sensitivity analysis that we carried out shows that the choice of the particular prior distribution hardly affects the posterior distribution of the parameters. This indicates that the chosen

[14] 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 (2002).

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

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.

Additional evidence that the results are not dominated by the choice of the prior distributions is the fact that the point estimates from the SML procedure were essentially the same as those for the method reported here. Nevertheless, with the SML method one needs to resort to the first-order asymptotic results, which do not provide the exact small sample distribution for the estimated parameters.

3.1 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, $\Pr(w_{it}^*|y_{it}^*, m_{it}^*, \theta)$.

3.2 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

$$\beta \sim N(\beta_0, B_0),$$

then the posterior distribution of β, conditional on all other parameters is

$$\beta|\theta_{-\beta} \sim N(\hat{\beta}, B),$$

where

$$\hat{\beta} = B\left(B_0^{-1}\beta_0 + \sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{x}_{it}'\Sigma^{-1}(z_{it}^* - L_t\alpha_i)\right)$$

and

$$B = \left(B_0^{-1} + \sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{x}_{it}'\Sigma^{-1}\tilde{x}_{it}\right)^{-1}.$$

3.3 Sampling the Individuals' Random Effects α_i

The conditional likelihood of the random effects for individual i is as follows

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

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

$$\alpha_i \sim N(\hat{\alpha}_i, V_{\alpha_i}),$$

where

$$V_{\alpha_i} = \left(\Gamma_i^{-1} + \sum_{t=1}^{T}L_t'\Sigma^{-1}L_t\right)^{-1},$$

and

$$\hat{\alpha}_i = V_{\alpha_i} \sum_{t=1}^{T} L_t' \Sigma^{-1} (z_t^* - \tilde{x}_{it}\beta).$$

3.4 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.

3.5 Sampling the Elements of Γ_i, Δ_ρ and γ

Recall that the covariance matrix Γ_i has the form given by

$$\Gamma_i = \text{diag}(g_{i1}, \ldots, g_{i5}) * \Delta_\rho * \text{diag}(g_{i1}, \ldots, g_{i5})', \tag{38}$$

where

$$g_j = \left(\exp(\bar{x}_{ij}'\gamma_j)\right)^{1/2}$$

and Δ_ρ is the correlation matrix given in (33). 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.

In all the estimations reported below we employed 10,000 repetitions after the initial number of 1,000, which were discarded.

4 The Data

The data for this study comes from the Panel Study of Income Dynamics (PSID). 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 60 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 along the lines 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 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 against 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 the extract used 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 education is very stable, whereas experience and seniority tend to increase. The mobility variable indicates that in each of the sample years approximately 1/10 of the individuals changed jobs. Notice that the mobility is very large in the first year of the sample, certainly because of measurement error; hence, the need for treating initial conditions as separate equations. As a result, the average

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

[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.

seniority is only around 6 years, while the average experience is over 21 years. Consistent with other data sources, the average wage decreases over the sample years, the wage dispersion increases across years, and the participation rate decreases somewhat. Note that a significant fraction of the sample is non-whites. 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 remained quite stable over the sample years. A significant fraction of the sample resides in SMSA, but that fraction tends to decrease.

Looking at the distribution of individuals across the various industries we note that it remains quite stable, even though the fraction of workers employed in manufacturing industries decreases. 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. It is therefore important to control for the cohort composition of the sample in the regression analysis.

5 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 are high school dropouts. The second group consists of those who have are high school graduates, who may have acquired some college education or who 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 dropouts group*, *high school graduates group*, and *college graduates group*, respectively. Below we present the results, for each group separately, from the simultaneous estimation of the three equations, namely participation, mobility and wage equation (together with the initial conditions equations for participation and mobility). For brevity we do not report the estimates for the initial conditions' equations.

The participation equation includes the following right-hand-side variables: a constant, education, lagged labor market experience and its square, a set of three regional dummy variables, a dummy variable for residence in an SMSA, other family income, two dummy variables for being an African American and 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. In addition it also includes: lagged seniority on the current job and its square, and a set of nine industry dummy variables, all of which are listed in Table 1.

The (log) wage equation includes the following right-hand-side variables: a constant, education, experience and its square, seniority on the current job and its square, 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 (21) above, a set of three regional dummy variables, a dummy variable for residence in an SMSA, two dummy variables for being an African American and Hispanic, a set of nine industry dummy variables (the same as in the mobility equation), county of residence unemployment rate, a set of four dummy variables for the cohort of birth (as in the previous two equations), and a full set of year dummy variables. The dependent variable in this equation is the log of deflated annual wage. For individuals that worked less than a full year we annualize their earnings.

Recall that the variance covariance matrix for the individual random effects α_i is given in (38). In order to estimate this matrix for all individuals one needs to obtain estimates for both the elements of Δ_ρ and the coefficient vectors γ_j in $g_j = \left(\exp(\bar{x}_{ij}'\gamma_j)\right)^{1/2}$ $(j = 1, ..., 5)$. As explained above, the numerical

[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.

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 $\bar{x}_{ij}$, we only use the first three principle components of $\bar{x}_{ij}$, as well as a constant term.[19]

High School Dropouts Group:

The results for this group are presented in Tables 2 through 4. Table 2 provides the results for the participation and mobility decisions, while Table 3 contains the results for the wage equation. Table 4 presents the results for the elements of the covariance matrices, namely Σ and Δ_ρ (which is part of Γ_i). For brevity we do not report the estimates for γ_j ($j = 1, ..., 5$). In the discussion below, we focus on the role of education, experience, and seniority in wage determination, when the participation and mobility decisions are endogenously determined. To better evaluate the results, we also provide graphs of the marginal posterior distributions for the variables of interest. In Figure 1 we depict the marginal distribution for the coefficients on education, experience and experience squared in the participation equation. To be able to better compare the results for the three education groups the results for all education group are included. Similarly, in Figure 2 we depict the marginal posterior distribution for the same coefficients and the coefficients on seniority and seniority squared in the mobility equation for all three education groups.

Figures 3 through 5 provide the posterior distributions for the returns to education, experience, and seniority, respectively. While the return to education is simply the coefficient on education, for experience and seniority we need to evaluate the return at some level of experience and seniority, respectively, as indicated in the figures. Finally, in Figure 6 we present the wage paths due to changes in seniority and jobs mobility (the J_{it}^W function). The graphs are depicted for a high school dropout worker with a particular mobility pattern for two levels of experience: (a) a new entrants (Figures 6a); and (b) a mid-career worker (Figure 6b).

It is apparent from Table 2 that the education level is almost irrelevant in this group for either the participation or the mobility decisions. In contrast, all lagged variables are very important predictors of both participation and mobility. Lagged experience has a significant positive effect on participation decision and negative effect on mobility. The same is true for lagged mobility in the mobility equation. That is, high school dropout workers who moved in the recent past tend to stay at their current jobs. Consequently, the average seniority in the sample, over all participating individuals in all years, is about 5.6 years. The probability of a move, when evaluated at the mean level of the regressors, is .078. If seniority increases by 5 years, this probability decreases to .054, i.e., a decline of more than 30%. The probability for those with 15.6 years of seniority (i.e., 10 years above average) is only .044, that is almost half the value at the average seniority. These results are consistent with the results generally obtained in the literature.

The results also indicate that children have almost no effect on either decision.[20] There is also no clear pattern for the effect of place of residence. Being an African American has a significant negative effect on the probability of participation, but there is almost no difference between African Americans and white individuals in terms of the mobility patterns. Also the younger cohorts are more likely to participate in the labor force. Nevertheless, those who do participate have similar mobility patterns to the other cohorts.

Consistent with the general pattern described in Table 1, but somewhat more difficult to interpret, is the general decline in the coefficients on the time dummy variables over time. The decline in these time dummy variables is more pronounced for the participation decision than for the mobility decision. Note also that, as one would expect, mobility differs considerably across the various industries, being higher in industries such as finance and personal services than in industries such as public administration.

Table 3 reports the results for the wage equation, the focus of our investigation. The results clearly indicate that once one controls for jumps in earnings that result from job changes, the effect of seniority is of great importance. This is very much in line with Topel's (1991) results, and in contrast with the results found by Abraham and Farber (1987) and Altonji and Williams (1997). In fact, the point

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

[20]The data includes only heads of households, who are mostly men.

estimates in our study are almost identical to those obtained by Topel (1991).[21] Our results indicate that 10 years of job seniority for a typical high school dropout worker increases his earnings by 59.5% (i.e., $100 \cdot (\exp(.467) - 1)$, where .467 is the implied cumulative return to job tenure). We also note that the range of this parameter in the marginal posterior distribution is rather small, namely .0455, to .0580.

As noted above, the returns to education, experience, and seniority, are also presented in Figures 3 through 5, respectively. Figure 5 clearly indicates that the return to seniority is quite high at all levels of seniority, being approximately 4.6% per year at 5 years of seniority and dropping down to 3.7% at 15 years of seniority. In fact, comparing Figures 4 and 5 indicates that the return to seniority is much larger than that for experience at any comparable level of seniority and experience. This finding is somewhat at odds with the results obtained by Topel (1991): Topel's total within-job wage growth is 0.126, larger than the 0.080 estimate obtained here. However, we model both the participation and mobility, so that Topel's estimates can be viewed as biased estimates due to the endogeneity of experience. The return to education, measured within the generally low-wage group, (see Figure 3) lies exactly between the returns to experience and the returns to seniority.

The estimates for the parameter of the "switching" function J_{it}^{W} are reported in lines 7-19 of Table 3. The estimates indicate that those workers who change jobs frequently, i.e. after less than a year, apparently do so in order to increase their wages. The lump sum gain is about 10%, while there is no loss in wages due to loss of seniority in the previous job. High school dropout workers who move after more than one year, lose approximately 5% for every year of seniority they accumulated on their last job due to the loss on the accumulated returns to seniority. On the other hand these workers gain about 10-24% for each move they had in the past (depending on how long they remained at the previous job), in addition to an increase of 2-3.5% per year of seniority in their last job. The net effect on the individual's annual earnings is negative if the employment spell with the last employer exceeded 5 years. If the worker's experience at the last job exceeded 10 years there is an additional decline 0.9% per year of experience. This feature is specific to high school dropouts and may well relate to the fact that most of their acquired human capital is firm specific.

Figure 6 depicts the wage path of an individual with a particular history of job mobility, that is, it represents the part of the individual's annual earning that resulted from the returns to seniority (i.e., within-job change in wage due to seniority) and the changes in the J_{it}^{W} function due to job changes (i.e., between-job change in wage). Note that each time a worker moves to a new firm, he/she loses the seniority accumulated on the previous job, and gains a certain amount according to his/her specific job history (i.e., the accumulated experience, level of seniority in the job that was left, the number of past moves, etc.) through the J_{it}^{W}. In order to account for the constraints implied by the mobility and wage equations, we depict the wage path for a typical mobility pattern for workers who moved. This mobility path was calculated from the data, for workers with 0-2 years of experience and for workers with 10-12 years of experience at the start of their sample period.[22]

The high school dropout workers move rather frequently, particularly earlier in their career. For both experience groups, short employment spells induce positive between-jobs wage changes, whereas long employment spells induce negative changes. Most of the wage increases are due to within-job rather than between-jobs wage changes. As Figure 6 indicates, a typical path does have periods with wage losses, but the trend over the life cycle is of general increase in real annual earnings.

Note also that inter-industry wage effects (see lines 31-39 in Table 3) are very much in line with what is known in the literature on low-wage employment; manufacturing is a high-wage industry whereas services are low-wage industries for the high school dropouts.

Finally, in Table 4 we report the parameters of the covariance matrices. As the estimates indicate, the correlation between the errors across equations for the individual specific effects are almost all significant, especially for those not in the equations controlling for initial conditions (see lines 12 to 14). For the idiosyncractic parts only some are significant. As expected, participation and wage equations are negatively correlated through the white noises (-.035), but are positively correlated through the

[21] In particular, see Table 3 of Topel (1991). Unlike Topel's case, we find greater effects at higher levels of seniority, but this can be largely explained by the fact that we do not include more than quadratic terms in seniority and experience.

[22] For the 0-2 experience group the mobility sequence used is 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, while for the 10-12 experience group it is 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, where 1 denotes a move and 0 denotes the person stayed in the same job as last year. Each of the sequences is for 18 years, the same length as the sample period.

individual specific effects (.296). Clearly, the correlation between individual specific effects is the more dominant one, and it indicates that individuals who tend to participate more also tend to have higher wages. Interestingly, mobility and wages are negatively correlated both through the idiosyncratic shocks and through the individual specific effects of these two equations. This implies that a large negative unexpected change in wage is likely to induce workers to move to a new job. This is also true for the individual specific effects, that is, high-wage workers are less likely to move than low-wage workers having the same observed characteristics.

Also, for the participation and mobility, the correlation between idiosyncratic parts is positive (.003), but insignificantly different from zero, while the correlation between the individual specific terms is negative (-.21), and statistically significant. That is, the results indicate that the type of individuals who tend to participate less also tend to move more, when they do participate. Overall, the results clearly demonstrate the importance of directly accounting for participation and mobility decision. A failure to do so is likely to lead to substantial bias in the estimated returns.

High School Graduates Group:

The results for the high school graduates group are reported in Tables 5 through 7. Table 5 provides the results for the participation and mobility equations, while Table 6 contains the results for the wage equation. Table 7 presents the estimates for the elements of the covariance matrices. Similar to the previous group, Figures 1 through 5 also provide the marginal posterior distribution for some of the key parameters, as well as for the returns to education, experience, and seniority. In Figure 7 we present wage paths resulting from a typical mobility pattern for two experience groups, similar to the ones presented in Figure 6.

Table 5 indicates that, in sharp contrast to the high school dropout group, the level of education is a very important factor in the participation decision. In fact, the results indicate that workers who have some university education can extract some benefits from their additional investment in human capital relative to those with only high school education. Furthermore, and in sharp contrast to the high school dropouts group, the geographical location variables are, in general, statistically important, and especially residency in an SMSA. Most other variables have similar effects on the participation and mobility decisions as for the high school dropouts group. In particular, there is a cumulative effect of participation, i.e., past participation has positive effect on future participation through lagged experience. On the other hand there is also the opposite effect of mobility, that is, higher seniority and lagged mobility reduce the probability of a job move. The probability of a job switch for workers in this education group is 0.0983, when it is evaluated at the mean level of the regressors. The value of seniority at the mean is approximately 4.7 years. The probability of mobility for a person with 9.7 years of seniority is only 0.0623, and for a person with 14.7 years of seniority, it decreases further to 0.0468. A closer look at the marginal posterior distributions for the coefficients on experience and seniority shows quite a dense distribution around the reported parameter estimates for both the participation and mobility equations.

Next we turn to the results of the wage equation, which are reported in Table 6 and Figures 3 through 5. Note that the effect of seniority is smaller than for the high school dropout group, but the marginal return declines at a slower rate. As a result the mean return at low levels of seniority (say 5 years) is higher for the high school dropouts group (see Figure 5a), but at high levels of seniority (say 15 years) the relationship is reversed (see Figure 5c). In any case, the return to seniority is clearly large and statistically very significant, with a cumulative return that exceeds that for the high school dropouts group. Furthermore, for this group, the return to experience is twice as large as it is for the high school dropouts group, at all levels of experience as is clear from a comparison of the graphs depicted in Figure 4. Hence, the sum of the linear components of the returns to seniority and experience, 0.924, is larger than for high school dropout workers, but somewhat smaller than Topel's (1991) findings for the whole population.

Lastly, we describe the results for the J_{it}^{W} function. First, we observe that the number of job changes always has a strong positive effect on the individual's wage in the new job, except for those jobs that lasted more than 10 years. Furthermore, if seniority at the last job change was between 6 and 10 years, about 35% of the loss is recovered as a lump sum, but not because of the level of accumulated seniority in the last job. In contrast, workers for whom seniority in the last job was either between 2 and 5 years or over 10 years recover roughly 3% for each year of seniority. It therefore appears that, in comparison with a high school dropout, an optimal move should take place before a person becomes too acquainted with the job (i.e., after 2-5 years on the job), or after acquiring a significant amount of experience on the job (i.e., after 10 years with the same employer). Note that, in contrast with high school dropout workers, experience at the last job does not have a negative effect on the wage change, at all levels of experience. This implies that job movements later in one's career seems more beneficial for the high school graduate workers.

As in Figure 6, Figure 7 depicts the wage path of an individual with a particular history of job mobility, that is, it represents the part of the individual's annual earnings that resulted from the returns to seniority and the changes in the J_{it}^{W} function due to job changes. As for the high school dropouts group we consider workers at two experience levels, namely 0-2 and 10-12 years of experience.[23] In contrast with Figure 6, job mobility causes almost no loss in earnings. Furthermore, earlier in one's career, job changes seems to induce larger wage increases than are obtained due to returns to seniority per se. Nevertheless, this effect attenuates through time. For example, for workers with 10-12 years of experience most of the wage increases are due to within-job increases, even though job changes do come with large lump sum increases. Also, the return to education for the high school graduates group is somewhat higher than for the high school dropouts, and, as apparent from comparison of graphs depicted in Figure 3, the posterior distribution is less spread than for the high school graduates group.

Looking at the estimates of the correlations presented in Table 7, we note that most of the estimated correlation coefficients are highly significant. They are generally similar and they all have the same sign as for the high school dropouts group. However, some key correlation coefficients are much larger; see especially the estimates in lines 12 to 14 of Table 7. As for the high school dropouts the individual specific effects from the participation and wage equations are positively correlated (.335) and the coefficient is only slightly larger. The correlation coefficients between the individual specific effects for the mobility and wage, and participation and mobility, are much larger in absolute terms than the corresponding coefficients for the high school dropouts group. For the mobility and wage the correlation coefficient is -.523, while for the participation and mobility equations it is -.430. That is, qualitatively the two groups demonstrate similar characteristics, but the high school graduates who tend to have larger wages tend to move even less than the high school dropouts. Similarly, the results indicate that the type of individuals who tend to participate less also tend to move more, when they do participate, and even more so for the high school graduate than for the high school dropouts.

College Graduate Group:

The results for the college graduate group are provided in Tables 8 through 10. Table 8 presents the results for the participation and mobility equations, while Table 9 contains the results for the wage equation. Table 10 presents estimates of terms of the covariance matrices, similar to those presented in Tables 4 and 7, for the lower education groups. As indicated above, the results are also presented graphically in Figures 1 through 5 and Figure 8. In Figure1 we depict the marginal distribution for the coefficients on education, experience and experience squared in the participation equation, along with the results for the other two educational groups. Similarly, in Figure 2 we depict the marginal posterior distributions for the same coefficients, as well as the coefficients on seniority and seniority squared, in the mobility equation. In Figures 3 through 5 we provide the marginal posterior distributions of the returns to education, experience, and seniority, respectively. Finally, similarly to Figures 6 and 7, Figure 8 presents the wage change due to changes in seniority and the J_{it}^{W} function.

From the three education groups, the college-educated workers are most likely to have general, rather than firm-specific human capital. In addition one would expect within-group heterogeneity to be larger for this group than for the other two groups because of more pronounced differences in career paths, hierarchical positions in the firm, etc. Table 8 indeed confirms that assertion: The effects of the various variables are larger for this group than for the other two education groups. For example, the probability of moving, conditional on a move in the preceding period, is much lower than for the other two groups, indication of a more stable career attachment for the more highly educated workers.

[23]For the 0-2 experience group the mobility sequence used is 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, while for the 10-12 experience group it is 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0. As in Figure 6, 1 denotes a move and 0 denotes the person stayed in the same job as last year. Each of the sequences is for 18 years, the same length as the sample period.

Estimates of the wage equation, presented in Table 9 display similar features. The within-group return to education is comparable to that for the high school graduates group. However, we note that the constant, a measure of between-group wage differentials—attributed to education—is larger (8.3) than those for the other education groups (approximately 7.9). In addition, returns to experience, a measure of the returns to general human capital, is also much larger for the college graduate (5.8%) than for the other two groups (3.7% and 4.0%, for the high school dropouts and high school graduate groups, respectively). The return to seniority is larger than for the high school graduate, but almost the same as for the high school dropouts. Nevertheless, the career incentives, and therefore the observed mobility pattern, are very different across the three education groups, as is apparent from close examination of the various components of the J_{it}^{W} function. First, frequent job-to-job mobility induces large wage increases, but not as large as those observed for longer spells. A job change after one year is associated with a wage increase (for the remainder of the individual's career) of approximately 25%, a much larger increase than for the other two groups. A job move after 2 to 5 years is associated with a smaller increase in wage, i.e., 18%, but the increase is augmented by an additional increase of 5.8% for each year of seniority in the previous job.

The results associated with moves after more than six years are markedly different from those obtained for the less educated individuals (see Tables 3 and 6 in comparison with Table 9). In particular, the wage compensation is not proportional to the wage loss due to loss of seniority. For instance, a person leaving his/her employer after 6 years would lose 30% due to the loss of accumulated seniority on that job, but will gain wage increase of almost 40% (i.e., $100 \cdot (\exp(.3231 + .0111) - 1) = 39.7$) due to that move. After 8 years of seniority, the equivalent numbers would be 40% and, as before 40%, respectively. In contrast, job movements that occur after spells that last more than 10 years entail wage losses.

In Figure 8 we depict, similarly to Figures 6 and 7 for the other two education groups, the results for the wage path for an individual with a particular history of job mobility. As before, the implied changes in the annual earnings are due to the changes in the returns to seniority and the changes in J_{it}^{W} function due to job changes. As for the other groups the wage path is computed for two experience levels, namely 0-2 and 10-12 years of experience.[24] It is apparent that the results are similar to those obtained for the high school dropouts and high school graduates groups. One difference that is worth noting is that the size of the between-jobs jumps are somewhat smaller, whereas the within-job growth is larger than for the other two groups.

In addition to the dynamic considerations discussed above, we also see that some markets offer high wages. For instance, there is a premium to those who live in the Northeast region or in an SMSA. In contrast, college graduate workers employed in the North Central region receive lower wages. Similar structure is also observed across the various industries. For example, the wholesale and retail trades and personal services industries are low-wage sectors for the college-educated, while the manufacturing and finance are high-wage industries.

Table 10 provides the estimates for the various elements of the covariance matrices for the college graduate groups. As is clearly seen, many of the estimated correlations are very similar to those estimated for the high school graduates group. In particular, the correlation coefficients between the individual specific terms are almost the same. Nevertheless, college-educated workers with a higher tendency to participate have even less tendency to move than the high school graduates. This is yet more evidence that career concerns are more important for the college-educated.

6 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 above the appropriateness of the methods used,

[24] For the 0-2 experience group the mobility sequence used is 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, while for the 10-12 experience group it is 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0. As in Figures 6 and 12, 1 denotes a move and 0 denotes the person stayed in the same job as last year. Each of the sequences is for 18 years, the same length as the sample period.

to assess such theories. In a seminal paper, Topel (1991) concludes that there is a significant return to seniority and hence strong support for the theoretical literature on human capital. This finding was in stark contradiction to most previous studies in the literature that concluded that there is no evidence for return 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.

Here we reinvestigate the interrelations between participation, mobility and wages, while examining several questions central to labor economics. Specifically, we model the joint decision of participation and job mobility, while allowing for potential sample selection bias to exist when estimating the equation of interest, namely the wage function. This allows us to address, in a more satisfactory way, a topic which has been in the center of attention over the past fifteen years, namely the return to seniority in the United States. We provide new evidence on the returns to seniority, and experience, as well as some evidence on "optimal" job-to-job mobility patterns. To do so, 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 return to seniority within a job, as well as discrete changes in the starting wage at the beginning of a new job. The results clearly demonstrate the importance of this joint estimation of the wage equation and the participation and mobility decisions. These two decisions have significant effects on observed outcomes, namely the 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. However, the total wage growth is somewhat smaller than implied by Topel's study. Our estimates of the returns to experience are lower than those estimated by Topel, but 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 optimal patterns of job-to-job mobility, patterns that differ markedly across education groups. In particular, we see that job changes are important elements of wage growth for the most educated group. Furthermore, wage losses after a job change is much less likely for college graduate group than for workers with lower education. Hence, mobility through the wage distribution is achieved through a combination of wage increases within the firm and across firms. The former is the more important for wage growth of the high school dropouts, while the latter is more important for the college graduates.

APPENDIX: Proof of Conditions 1 and 2

All along the proof, we assume that the mobility cost is strictly greater than the search cost:

$$c_M > \gamma_1$$

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

$$\frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} < \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} \tag{A1}$$

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

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

Thus

$$\frac{u'\left(z_t + w^*_{22,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} \tag{A2}$$

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 (A1) implies

$$\gamma_{11} < \gamma_{12} < 0$$

Then, equation (12) implies that

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

Moreover, as $w^*_{02,t} < w^*_{22,t}$,

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

Due to equation (A2), this inequality implies that, if $c_M > \gamma_1$,

$$\gamma_{22} < \gamma_{12} < 0$$

Thus equation (13) implies that $w^*_{22,t} < w^*_{02,t}$, which is in contradiction with 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 (12) implies that

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

Two cases must then be distinguished.

- 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 equation (13) implies that $w^*_{22,t} < w^*_{02,t}$, which is in contradiction with the initial assumption.

- 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

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

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, equation (14) implies that $w^*_{22,t} > w^*_{02,t} > w^*_{12,t}$. If $w^*_{21,t} > w^*_{22,t}$,[25] a mover (i.e. a worker who moved to another firm at the end of period $t-1$) becomes nonparticipant at the end of period t if he or she is offered a wage less than $w^*_{22,t}$. A stayer becomes nonparticipant 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

$$\begin{aligned} y_t &= \mathbf{1}\left[w_t > w^*_{02,t} - \gamma_{12}\ y_{t-1} + \gamma_{22}\ y_{t-1}\ m_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{02,t} + \gamma_{12}\ y_{t-1} - \gamma_{22}\ y_{t-1}\ m_{t-1} > 0\right] \end{aligned} \tag{A3}$$

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 accept 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

$$\begin{aligned} m_t &= \mathbf{1}\left[w_t > w^*_{12,t} + \gamma_{22}\ y_{t-1}\ m_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w^*_{12,t} - \gamma_{22}\ y_{t-1}\ m_{t-1} > 0\right] \end{aligned} \tag{A4}$$

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 (A1) implies that

$$\gamma_{11} < 0 < \gamma_{12}$$

Thus, equation (12) implies that

$$w^*_{12,t} < w^*_{02,t} < w^*_{01,t} < w^*_{11,t}$$

- 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 (13) implies that $w^*_{22,t} < w^*_{02,t}$, which is in contradiction with the initial assumption.

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

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

[25] This is satisfied if $\gamma_{21} > \gamma_{22}$.

Thus

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

We find the same condition than previously. Under this equation (14) 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 equations (A3) and (A4) 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

$$\frac{u'\left(z_t+w_{22,t}^{*},1;X_t\right)-u'\left(z_t,0;X_t\right)}{u'\left(z_t+w_{02,t}^{*},1;X_t\right)} > 1-\frac{u'\left(z_t,0;X_t\right)}{u'\left(z_t+w_{02,t}^{*},1;X_t\right)} \tag{A5}$$

- If $u'\left(z_t+w_{22,t}^{*},1;X_t\right) < u'\left(z_t,0;X_t\right)$, then equation (A5) implies that $0>\gamma_{22}>\gamma_{12}$ (if $c_M > \gamma_1$). Then we deduce from equation (13) that $w_{22,t}^{*} > w_{02,t}^{*}$, which is in contradiction with the assumption.
- If $u'\left(z_t,0;X_t\right) < u'\left(z_t+w_{22,t}^{*},1;X_t\right)$, then equation (A5) implies that $\gamma_{22}>\gamma_{12}>0$ (if $c_M>\gamma_1$). Then equation (13) implies that $w_{22,t}^{*} > w_{02,t}^{*}$, which is in contradiction with the assumption.
- Consequently, the assumption $w_{02,t}^{*} > w_{22,t}^{*}$ 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 equation (12) 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

$$\frac{u'\left(z_t+w_{22,t}^{*},1;X_t\right)-u'\left(z_t,0;X_t\right)}{u'\left(z_t+w_{02,t}^{*},1;X_t\right)} < 1-\frac{u'\left(z_t,0;X_t\right)}{u'\left(z_t+w_{02,t}^{*},1;X_t\right)} < 0 \tag{A6}$$

Thus, if $c_M>\gamma_1$, equation (A6) implies that $\gamma_{22}<\gamma_{12}<0$. From equation (14), we can deduce that $w_{22,t}^{*} < w_{02,t}^{*}$, which is in contradiction with the assumption.

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 equation (13) 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, equation (14) 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 equations (A3) and (A4) 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

$$\frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{02,t}, 1; X_t\right)} \tag{A7}$$

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

$$\frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} \tag{A8}$$

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 equation (A7) implies

$$\gamma_{12} < \gamma_{11} < 0$$

Then, equation (12) implies that

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

Moreover, as $w^*_{01,t} < w^*_{21,t}$,

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

Due to equation (A8), this inequality implies that, if $c_M > \gamma_1$,

$$\gamma_{21} < \gamma_{11} < 0$$

Thus equation (13) implies that $w^*_{21,t} < w^*_{01,t}$, which is in contradiction with the initial assumption. Thus $w^*_{21,t}$ 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, equation (12) implies that

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

Two cases must then be distinguished.

- 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 equation (13) implies that $w_{21,t}^* < w_{01,t}^*$, which is in contradiction with the initial assumption.

- 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
$$\begin{aligned} w_{21,t}^* &> w_{01,t}^* \Leftrightarrow \gamma_{21} - \gamma_{11} > 0 \\ &\Leftrightarrow c_M > \gamma_1 \left[\frac{u'\left(z_t + w_{01,t}^*, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w_{21,t}^*, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}\right] \end{aligned}$$
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, equation (14) implies that $w_{21,t}^* > w_{01,t}^* > w_{11,t}^*$. Now let us assume that $\gamma_{22} > \gamma_{21}$, which implies that $w_{22,t}^* > w_{21,t}^*$. In that case, there is no interfirm mobility: a nonparticipant moves to employment (respectively, stays in the nonparticipation 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 nonparticipant (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
$$\begin{aligned} y_t &= \mathbf{1}\left[w_t > w_{01,t}^* - \gamma_{11}\ y_{t-1}\right] \\ &= \mathbf{1}\left[w_t - w_{01,t}^* + \gamma_{11}\ y_{t-1} > 0\right] \end{aligned} \tag{A9}$$
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 equation (A7) implies that
$$\gamma_{12} < 0 < \gamma_{11}$$
and equation (12) implies that
$$w_{11,t}^* < w_{01,t}^* < w_{02,t}^* < w_{12,t}^*$$

- 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 equation (13) implies that $w_{21,t}^* < w_{01,t}^*$, which is in contradiction with the initial assumption.

- 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
$$\begin{aligned} w_{21,t}^* &> w_{01,t}^* \Leftrightarrow \gamma_{21} - \gamma_{11} > 0 \\ &\Leftrightarrow c_M > \gamma_1 \left[\frac{u'\left(z_t + w_{01,t}^*, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w_{21,t}^*, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}\right] \end{aligned}$$
We find the same condition than previously. Under this condition, equation (14) 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 equation (A9) 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

$$\frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} > 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} \tag{A10}$$

- If $u'\left(z_t + w^*_{21,t}, 1; X_t\right) < u'\left(z_t, 0; X_t\right)$, then equation (A10) implies that $0 > \gamma_{21} > \gamma_{11}$ (if $c_M > \gamma_1$). Then we can deduce from equation (13) that $w^*_{21,t} > w^*_{01,t}$, which is in contradiction with the assumption.
- If $u'\left(z_t, 0; X_t\right) < u'\left(z_t + w^*_{21,t}, 1; X_t\right)$, then equation (A10) implies that $\gamma_{21} > \gamma_{11} > 0$ (if $c_M > \gamma_1$). Then equation (13) implies that $w^*_{21,t} > w^*_{01,t}$, which is still in contradiction with the assumption.
- 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

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

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 equation (12) implies that

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

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

$$\frac{u'\left(z_t + w^*_{21,t}, 1; X_t\right) - u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < 1 - \frac{u'\left(z_t, 0; X_t\right)}{u'\left(z_t + w^*_{01,t}, 1; X_t\right)} < 0 \tag{A11}$$

Thus, if $c_M > \gamma_1$, equation (A11) implies that $\gamma_{21} < \gamma_{11} < 0$. From equation (14), we can then deduce that $w^*_{21,t} < w^*_{01,t}$, which is in contradiction with the assumption.

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 equation (13) 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 (14) 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 (A9).

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.

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 (2001): "Wages, Experience and Seniority," The Institute for Fiscal Studies, Working Paper 01/01.

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 (2002): "Bayesian Inference for the Mover-Stayer Model in Continuous Time, with an Application to Labour Market Transition Data," CREST-INSEE Working Paper 2002-23, forthcoming in *Journal of Applied Econometrics.*

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, 653–77

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

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.

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

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	5.4968	5.4141	5.1609	5.5538	5.7433	6.0163	6.2339	6.2294	6.8648
		(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	9.3009	9.1183	9.0478	8.8652	8.7232	8.7794	8.6725	8.7129	8.5501
		(2.4239)	(2.7819)	(2.8645)	(3.1292)	(3.3113)	(3.3359)	(3.4271)	(3.3237)	(3.5153)
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
Geographical Location Characteristics:									
14. Northeast	0.1285	0.0901	0.0935	0.0879	0.0919	0.0875	0.0831	0.0841	0.0715
	(0.6766)	(0.8811)	(0.8434)	(0.8671)	(0.8665)	(0.8884)	(0.9207)	(0.8586)	(0.9130)
15. North Central	0.2106	0.1722	0.1842	0.1707	0.1681	0.1620	0.1510	0.1374	0.1301
	(0.7156)	(0.9148)	(0.8814)	(0.9018)	(0.8985)	(0.9193)	(0.9485)	(0.8824)	(0.9383)
16. South	0.3956	0.3620	0.3740	0.3715	0.3691	0.3677	0.3721	0.3781	0.3634
	(0.7648)	(0.9620)	(0.9278)	(0.9514)	(0.9489)	(0.9705)	(1.0025)	(0.9462)	(0.9993)
17. Living in SMSA	0.6821	0.6850	0.6842	0.6812	0.5704	0.5647	0.5495	0.5981	0.5905
	(0.4657)	(0.4646)	(0.4649)	(0.4661)	(0.4951)	(0.4959)	(0.4976)	(0.4903)	(0.4918)
18. County unemp. rate	8.9048	5.5994	7.0618	9.6538	7.1143	6.6044	5.3981	5.5160	6.9594
	(3.1052)	(2.1196)	(2.5924)	(3.4391)	(3.1559)	(2.6720)	(2.2687)	(2.1410)	(2.2664)
Industry:									
19. Construction	0.0827	0.0931	0.0877	0.0780	0.0820	0.0871	0.0853	0.0899	0.0878
	(0.2755)	(0.2905)	(0.2829)	(0.2682)	(0.2744)	(0.2820)	(0.2793)	(0.2861)	(0.2831)
20. Manufacturing	0.2449	0.2508	0.2483	0.2412	0.2274	0.2162	0.2138	0.2163	0.2010
	(0.4301)	(0.4335)	(0.4321)	(0.4279)	(0.4192)	(0.4117)	(0.4100)	(0.4118)	(0.4008)
21. Trans., Comm., etc.	0.0818	0.0801	0.0858	0.0848	0.0894	0.0814	0.0805	0.0800	0.0793
	(0.2741)	(0.2715)	(0.2802)	(0.2786)	(0.2854)	(0.2735)	(0.2722)	(0.2713)	(0.2702)
22. Wholesale and Retail	0.1108	0.1152	0.1146	0.1180	0.1247	0.1337	0.1395	0.1372	0.1247
	(0.3139)	(0.3193)	(0.3186)	(0.3227)	(0.3304)	(0.3404)	(0.3465)	(0.3441)	(0.3304)
23. Finance	0.0349	0.0323	0.0309	0.0305	0.0333	0.0322	0.0326	0.0344	0.0337
	(0.1835)	(0.1769)	(0.1730)	(0.1720)	(0.1793)	(0.1765)	(0.1775)	(0.1823)	(0.1805)
24. Bus. & Repair Services	0.0381	0.0323	0.0336	0.0378	0.0407	0.0490	0.0480	0.0520	0.0504
	(0.1915)	(0.1769)	(0.1803)	(0.1906)	(0.1975)	(0.2159)	(0.2138)	(0.2221)	(0.2188)
25. Personal Services	0.0304	0.0289	0.0309	0.0188	0.0189	0.0173	0.0225	0.0223	0.0224
	(0.1718)	(0.1674)	(0.1730)	(0.1357)	(0.1361)	(0.1304)	(0.1483)	(0.1478)	(0.1481)
26. Professional	0.0916	0.0918	0.0949	0.0875	0.0869	0.0952	0.0975	0.0883	0.0947
	(0.2885)	(0.2888)	(0.2931)	(0.2826)	(0.2818)	(0.2935)	(0.2966)	(0.2838)	(0.2928)
27. Public Administration	0.0801	0.0759	0.0691	0.0640	0.0694	0.0707	0.0679	0.0693	0.0654
	(0.2714)	(0.2649)	(0.2537)	(0.2448)	(0.2542)	(0.2563)	(0.2516)	(0.2540)	(0.2473)

Table 1: (Continued)

Variable	Year								
	1975	1978	1980	1982	1984	1986	1988	1990	1992
Cohort Effects:									
28. 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)
29. 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)
30. 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)
31. 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 and Mobility Equations for High School Dropouts

	Variable	Participation				Mobility			
		Mean	St. Dev.	Range		Mean	St. Dev.	Range	
				Min	Max			Min	Max
1.	Constant	-0.5880	0.3064	-1.1951	0.0145	-1.0892	0.2139	-1.5020	-0.6618
2.	Education	0.0167	0.0138	-0.0107	0.0437	-0.0109	0.0090	-0.0292	0.0065
3.	Experience at $t-1$	0.0345	0.0101	0.0143	0.0545	-0.0217	0.0059	-0.0336	-0.0102
4.	Experience at $t-1$ squared	-0.0012	0.0002	-0.0016	-0.0008	0.0001	0.0001	-0.0001	0.0004
5.	Seniority at $t-1$	–	–	–	–	-0.0812	0.0115	-0.1007	-0.0605
6.	Seniority at $t-1$ squared	–	–	–	–	0.0018	0.0003	0.0011	0.0024
7.	Participation at $t-1$	1.7349	0.0660	1.5999	1.8530	–	–	–	–
8.	Mobility at $t-1$	0.5295	0.1258	0.3043	0.7836	-0.7190	0.0738	-0.8544	-0.5807
	Family Characteristics:								
9.	Family other income	0.0045	0.0066	-0.0068	0.0181	-0.0350	0.0071	-0.0486	-0.0213
10.	No. of Children	0.0179	0.0256	-0.0319	0.0697	0.0019	0.0146	-0.0268	0.0310
11.	Children 1 to 2	-0.0911	0.0637	-0.2103	0.0361	0.0403	0.0363	-0.0302	0.1106
12.	Children 3 to 5	-0.0858	0.0671	-0.2171	0.0457	0.0176	0.0320	-0.0456	0.0802
13.	Married	0.3186	0.0842	0.1590	0.4834	-0.0692	0.0458	-0.1586	0.0204
	Geographical Location:								
14.	Northeast	-0.1167	0.0695	-0.2459	0.0230	-0.0373	0.0331	-0.1013	0.0285
15.	North Central	-0.0198	0.0738	-0.1600	0.1239	-0.0166	0.0371	-0.0847	0.0551
16.	South	0.1173	0.0808	-0.0314	0.2573	0.0416	0.0426	-0.0334	0.1157
17.	Living in SMSA	-0.0485	0.0646	-0.1766	0.0780	-0.0419	0.0306	-0.1008	0.0182
18.	County unemp. rate	-0.0324	0.0102	-0.0518	-0.0124	0.0074	0.0069	-0.0058	0.0204
	Race:								
19.	Black	-0.3337	0.0827	-0.5005	-0.1780	-0.0300	0.0489	-0.1207	0.0602
20.	Hispanic	-0.0865	0.2805	-0.5438	0.4295	-0.0090	0.0758	-0.1616	0.1380
	Cohort Effects:								
21.	Age 15 or less in 1975	0.9856	0.2591	0.5460	1.5599	-0.0694	0.1171	-0.3072	0.1532
22.	Age 16 to 25 in 1975	0.8421	0.2516	0.4063	1.3342	-0.1047	0.1079	-0.3179	0.1043
23.	Age 26 to 35 in 1975	0.8560	0.1727	0.5269	1.1936	-0.0640	0.0942	-0.2527	0.1171
24.	Age 36 to 45 in 1975	0.7081	0.1742	0.3731	1.0653	0.0007	0.0889	-0.1797	0.1723

Table 2: (Continued)

	Variable	Participation				Mobility			
		Mean	St. Dev.	Range		Mean	St. Dev.	Range	
				Min	Max			Min	Max
	Industry:								
25.	Construction	–	–	–	–	0.7489	0.0686	0.6161	0.8811
26.	Manufacturing	–	–	–	–	0.5901	0.0570	0.4782	0.7047
27.	Trans., Comm., etc.	–	–	–	–	0.6428	0.0694	0.5081	0.7796
28.	Wholesale & Retail Trades	–	–	–	–	0.7124	0.0628	0.5946	0.8376
29.	Finance	–	–	–	–	0.8242	0.1192	0.5898	1.0621
30.	Business & Repair Services	–	–	–	–	0.7008	0.0914	0.5261	0.8783
31.	Personal Services	–	–	–	–	0.7489	0.1090	0.5413	0.9701
32.	Professional	–	–	–	–	0.7290	0.1172	0.5041	0.9441
33.	Public Administration	–	–	–	–	0.5535	0.1086	0.3439	0.7528
	Time Effects:								
34.	Year 1976	1.0338	0.2154	0.6334	1.4204	0.0220	0.1101	-0.1975	0.2311
35.	Year 1977	0.8399	0.1479	0.5498	1.1354	0.4728	0.0879	0.3063	0.6443
36.	Year 1978	0.9913	0.1369	0.7386	1.2741	0.3810	0.0915	0.2014	0.5594
37.	Year 1979	0.8712	0.2301	0.4956	1.3014	0.3748	0.1126	0.1573	0.5892
38.	Year 1980	0.8898	0.1356	0.6417	1.1699	0.3263	0.0991	0.1353	0.5143
39.	Year 1981	0.7150	0.1444	0.4485	1.0144	0.1238	0.1045	-0.0774	0.3281
40.	Year 1982	0.6566	0.1851	0.3308	1.0018	0.0730	0.0919	-0.1071	0.2565
41.	Year 1983	0.4078	0.1472	0.1367	0.7005	0.0360	0.1191	-0.1929	0.2574
42.	Year 1984	0.3594	0.1620	0.0627	0.6661	0.1093	0.0863	-0.0607	0.2759
43.	Year 1985	0.3685	0.1167	0.1442	0.5987	0.0399	0.0974	-0.1552	0.2248
44.	Year 1986	0.3775	0.1663	0.0607	0.6693	0.1243	0.0988	-0.0690	0.3182
45.	Year 1987	0.3129	0.1104	0.1019	0.5372	0.1635	0.0855	-0.0075	0.3280
46.	Year 1988	0.1356	0.1538	-0.1451	0.4216	0.1672	0.1372	-0.0862	0.4099
47.	Year 1989	0.1268	0.1051	-0.0786	0.3347	0.1260	0.0873	-0.0455	0.2952
48.	Year 1990	0.0739	0.1189	-0.1461	0.3143	0.1346	0.1039	-0.0672	0.3311
49.	Year 1991	-0.0111	0.0880	-0.1852	0.1607	0.1206	0.1058	-0.0805	0.3158

Table 3: **Wage Equation for High School Dropouts**

	Variable	Mean	St. Dev.	Range	
				Min	Max
1.	Constant	7.9945	0.1519	7.7081	8.2797
2.	Education	0.0366	0.0068	0.0238	0.0492
3.	Experience	0.0283	0.0027	0.0229	0.0334
4.	Experience squared	-0.0007	0.0000	-0.0007	-0.0006
5.	Seniority	0.0517	0.0034	0.0455	0.0580
6.	Seniority squared	-0.0005	0.0001	-0.0008	-0.0003
	Job switch variables in first sample year:				
7.	Dummy for job change job in 1st year	-0.0336	0.0800	-0.1796	0.1108
8.	Experience at $t-1$ if variable 7= 1	0.0152	0.0036	0.0082	0.0221
	No. of switches of jobs that lasted:				
9.	Up to 1 year	0.0923	0.0144	0.0635	0.1203
10.	2 to 5 years	0.0958	0.0219	0.0526	0.1386
11.	6 to 10 years	0.1229	0.1027	-0.0569	0.3076
12.	Over 10 years	0.2457	0.1078	0.0474	0.4606
	Seniority at last job change that lasted:				
13.	2 to 5 years	0.0293	0.0084	0.0127	0.0456
14.	6 to 10 years	0.0213	0.0109	0.0003	0.0422
15.	Over 10 years	0.0350	0.0053	0.0238	0.0444
	Experience at last job change that lasted:				
16.	Up to1 year	0.0009	0.0012	-0.0015	0.0033
17.	2 to 5 years	-0.0007	0.0016	-0.0038	0.0024
18.	6 to 10 years	0.0007	0.0030	-0.0049	0.0060
19.	Over 10 years	-0.0090	0.0029	-0.0150	-0.0035
	Geographical location:				
20.	Northeast	0.0505	0.0192	0.0131	0.0879
21.	North Central	0.0283	0.0184	-0.0094	0.0624
22.	South	-0.0778	0.0184	-0.1135	-0.0424
23.	Living in SMSA	0.0666	0.0164	0.0349	0.0993
24.	County unemp. rate	-0.0042	0.0023	-0.0087	0.0002
	Race:				
25.	Black	-0.2904	0.0368	-0.3596	-0.2213
26.	Hispanic	-0.0669	0.0458	-0.1532	0.0233
	Cohort effects:				
27.	Age 15 or less in 1975	0.4504	0.0660	0.3224	0.5795
28.	Age 16 to 25 in 1975	0.3308	0.0706	0.1943	0.4659
29.	Age 26 to 35 in 1975	0.2756	0.0602	0.1581	0.3915
30.	Age 36 to 45 in 1975	0.2183	0.0682	0.0883	0.3565

Table 3: (Continued)

	Variable	Mean	St. Dev.	Range	
				Min	Max
	Industry:				
31.	Construction	0.2693	0.0190	0.2324	0.3065
32.	Manufacturing	0.3707	0.0158	0.3398	0.4020
33.	Trans., Comm., etc.	0.3554	0.0268	0.3034	0.4077
34.	Wholesale and Retail Trades	0.2196	0.0185	0.1829	0.2553
35.	Finance	0.1801	0.0407	0.0997	0.2601
36.	Business and Repair Services	0.1532	0.0324	0.0934	0.2162
37.	Personal Services	0.1600	0.0328	0.0966	0.2243
38.	Professional	0.1970	0.0310	0.1371	0.2575
39.	Public Administration	0.2484	0.0988	0.1087	0.3846
	Time Effects:				
40.	Year 1975	0.5969	0.0365	0.5241	0.6679
41.	Year 1976	0.5531	0.0347	0.4837	0.6201
42.	Year 1977	0.5120	0.0417	0.4300	0.5929
43.	Year 1978	0.4821	0.0344	0.4133	0.5506
44.	Year 1979	0.4422	0.0314	0.3800	0.5039
45.	Year 1980	0.3975	0.0292	0.3402	0.4550
46.	Year 1981	0.3417	0.0285	0.2833	0.3965
47.	Year 1982	0.3162	0.0276	0.2611	0.3704
48.	Year 1983	0.2596	0.0321	0.1964	0.3210
49.	Year 1984	0.2198	0.0291	0.1630	0.2754
50.	Year 1985	0.2564	0.0263	0.2039	0.3074
51.	Year 1986	0.2370	0.0238	0.1900	0.2825
52.	Year 1987	0.1853	0.0297	0.1277	0.2409
53.	Year 1988	0.1401	0.0333	0.0766	0.2009
54.	Year 1989	0.1414	0.0231	0.0960	0.1864
55.	Year 1990	0.0808	0.0225	0.0376	0.1260
56.	Year 1991	0.0474	0.0202	0.0079	0.0875

Table 4: Estimates of the Stochastic Elements for Hifh School Dropouts

Variable		Mean	St. Dev.	Range	
				Min	Max
Covariance Matrix of White Noises (element of Σ):					
1.	ρ_{uv}	0.0029	0.0077	-0.0117	0.0160
2.	$\rho_{u\xi}$	-0.0346	0.0072	-0.0497	-0.0183
3.	$\rho_{v\xi}$	-0.0055	0.0074	-0.0185	0.0072
4.	σ_{ξ}^{2}	0.2448	0.0064	0.2331	0.2539
Correlations of Individual Specific Effects (elements of Δ_{ρ}):					
5.	$\rho_{\alpha_y^0 \alpha_m^0}$	-0.1020	0.1146	-0.2589	0.1067
6.	$\rho_{\alpha_w \alpha_y^0}$	0.3447	0.0351	0.2732	0.4142
7.	$\rho_{\alpha_y \alpha_y^0}$	0.7548	0.0566	0.6525	0.8747
8.	$\rho_{\alpha_m \alpha_y^0}$	0.0278	0.2007	-0.2908	0.2281
9.	$\rho_{\alpha_w \alpha_m^0}$	0.0646	0.0505	-0.0061	0.1794
10.	$\rho_{\alpha_y \alpha_m^0}$	0.1972	0.0746	0.0260	0.2971
11.	$\rho_{\alpha_m \alpha_m^0}$	-0.0573	0.1666	-0.2619	0.2194
12.	$\rho_{\alpha_y \alpha_w}$	0.2958	0.0282	0.2292	0.3560
13.	$\rho_{\alpha_m \alpha_w}$	-0.2744	0.0799	-0.4083	-0.1348
14.	$\rho_{\alpha_y \alpha_m}$	-0.2100	0.1053	-0.3832	-0.0429

Table 5: Participation and Mobility Equations for High School Graduates

	Variable	Participation				Mobility			
		Mean	St. Dev.	Range		Mean	St. Dev.	Range	
				Min	Max			Min	Max
1.	Constant	0.3645	0.6728	-1.0401	1.7243	-1.6639	0.2736	-2.2260	-1.1390
2.	Education	0.1068	0.0300	0.0494	0.1676	0.0156	0.0134	-0.0100	0.0430
3.	Experience at $t-1$	0.0518	0.0152	0.0225	0.0812	-0.0314	0.0073	-0.0457	-0.0172
4.	Experience at $t-1$ squared	-0.0024	0.0003	-0.0030	-0.0017	0.0004	0.0002	0.0001	0.0008
5.	Seniority at $t-1$	–	–	–	–	-0.0910	0.0078	-0.1059	-0.0753
6.	Seniority at $t-1$ squared	–	–	–	–	0.0021	0.0003	0.0015	0.0026
7.	Participation at $t-1$	1.5108	0.0876	1.3431	1.6801	–	–	–	–
8.	Mobility at $t-1$	0.4362	0.1317	0.1885	0.6832	-0.7715	0.0533	-0.8772	-0.6686
	Family Characteristics:								
9.	Family other income	-0.0105	0.0049	-0.0199	-0.0007	-0.0229	0.0045	-0.0314	-0.0145
10.	No. of Children	0.0064	0.0427	-0.0827	0.0866	-0.0043	0.0183	-0.0409	0.0309
11.	Children 1 to 2	-0.1098	0.0787	-0.2605	0.0409	0.0427	0.0352	-0.0262	0.1121
12.	Children 3 to 5	0.0299	0.0894	-0.1438	0.2122	-0.0434	0.0365	-0.1148	0.0279
13.	Married	0.3479	0.0975	0.1528	0.5460	-0.0408	0.0434	-0.1271	0.0425
	Geographical Location:								
14.	Northeast	0.2372	0.1115	0.0239	0.4663	-0.0380	0.0323	-0.1015	0.0252
15.	North Central	-0.2235	0.0850	-0.4096	-0.0654	0.0264	0.0281	-0.0294	0.0828
16.	South	0.0410	0.0738	-0.1130	0.1863	0.0190	0.0259	-0.0317	0.0689
17.	Living in SMSA	-0.2096	0.1013	-0.4106	-0.0093	0.0119	0.0348	-0.0556	0.0813
18.	County unemp. rate	-0.0163	0.0137	-0.0433	0.0099	-0.0045	0.0063	-0.0171	0.0078
	Race:								
19.	Black	-0.7999	0.1471	-1.0473	-0.4737	0.0540	0.0385	-0.0202	0.1292
20.	Hispanic	0.2959	0.3183	-0.2943	0.9353	0.0411	0.0733	-0.1050	0.1831
	Cohort Effects:								
21.	Age 15 or less in 1975	-0.0277	0.3661	-0.7951	0.6613	0.0280	0.1553	-0.2806	0.3258
22.	Age 16 to 25 in 1975	-0.3305	0.3636	-1.0753	0.3345	0.0262	0.1507	-0.2762	0.3165
23.	Age 26 to 35 in 1975	-0.0513	0.3293	-0.6915	0.5926	0.0450	0.1398	-0.2317	0.3140
24.	Age 36 to 45 in 1975	0.1024	0.2887	-0.4651	0.6583	0.1373	0.1238	-0.1082	0.3765

Table 5: (Continued)

	Variable	Participation				Mobility			
		Mean	St. Dev.	Range		Mean	St. Dev.	Range	
				Min	Max			Min	Max
	Industry:								
25.	Construction	–	–	–	–	1.0302	0.0905	0.8493	1.2041
26.	Manufacturing	–	–	–	–	0.8824	0.0817	0.7240	1.0454
27.	Trans., Comm., etc.	–	–	–	–	0.8285	0.0891	0.6587	1.0064
28.	Wholesale & Retail Trades	–	–	–	–	1.0256	0.0858	0.8608	1.1951
29.	Finance	–	–	–	–	1.0473	0.1016	0.8452	1.2461
30.	Business & Repair Services	–	–	–	–	0.9934	0.1032	0.7951	1.1990
31.	Personal Services	–	–	–	–	1.0269	0.1145	0.7954	1.2569
32.	Professional	–	–	–	–	1.0271	0.0948	0.8460	1.2243
33.	Public Administration	–	–	–	–	0.7635	0.0889	0.5884	0.9475
	Time Effects:								
34.	Year 1976	-0.0306	0.2117	-0.4403	0.3813	-0.2577	0.1215	-0.4953	-0.0247
35.	Year 1977	0.3586	0.2180	-0.0604	0.7816	0.4856	0.1004	0.2935	0.6818
36.	Year 1978	0.4981	0.2207	0.0713	0.9388	0.4142	0.0963	0.2266	0.6029
37.	Year 1979	0.2731	0.2091	-0.1559	0.6720	0.4896	0.0950	0.3021	0.6723
38.	Year 1980	0.4403	0.1966	0.0483	0.8079	0.3365	0.0933	0.1560	0.5180
39.	Year 1981	0.5411	0.1857	0.1877	0.8843	0.0682	0.0953	-0.1185	0.2537
40.	Year 1982	0.2613	0.1861	-0.0946	0.6417	0.0247	0.0992	-0.1749	0.2223
41.	Year 1983	0.1105	0.1774	-0.2483	0.4542	0.0713	0.0915	-0.1072	0.2484
42.	Year 1984	0.0574	0.1635	-0.2621	0.3818	-0.0233	0.0941	-0.2115	0.1589
43.	Year 1985	0.1224	0.1558	-0.1809	0.4202	-0.0713	0.0936	-0.2556	0.1130
44.	Year 1986	0.2696	0.1644	-0.0498	0.6010	0.0606	0.0877	-0.1098	0.2306
45.	Year 1987	-0.1352	0.1569	-0.4468	0.1615	0.1641	0.0873	-0.0052	0.3343
46.	Year 1988	-0.0784	0.1608	-0.4237	0.2132	0.0503	0.0851	-0.1205	0.2148
47.	Year 1989	0.1526	0.1593	-0.1613	0.4585	0.0322	0.0892	-0.1430	0.2035
48.	Year 1990	-0.0012	0.1462	-0.3028	0.2800	0.1633	0.0857	-0.0083	0.3306
49.	Year 1991	0.0504	0.1400	-0.2229	0.3324	0.0901	0.0837	-0.0701	0.2534

Table 6: **Wage Equation for High School Graduates**

	Variable	Mean	St. Dev.	Range	
				Min	Max
1.	Constant	7.8848	0.1321	7.6266	8.1391
2.	Education	0.0397	0.0056	0.0292	0.0510
3.	Experience	0.0498	0.0030	0.0440	0.0555
4.	Experience squared	-0.0011	0.0001	-0.0013	-0.0010
5.	Seniority	0.0426	0.0029	0.0369	0.0481
6.	Seniority squared	-0.0001	0.0001	-0.0002	0.0001
	Job switch variables in first sample year:				
7.	Dummy for job change job in 1st year	-0.0205	0.0592	-0.1321	0.1021
8.	Experience at $t-1$ if variable 7= 1	0.0140	0.0044	0.0054	0.0225
	No. of switches of jobs that lasted:				
9.	Up to 1 year	0.1234	0.0143	0.0951	0.1519
10.	2 to 5 years	0.1671	0.0210	0.1251	0.2075
11.	6 to 10 years	0.3464	0.0667	0.2144	0.4783
12.	Over 10 years	0.1079	0.0874	-0.0638	0.2826
	Seniority at last job change that lasted:				
13.	2 to 5 years	0.0271	0.0072	0.0131	0.0413
14.	6 to 10 years	-0.0071	0.0090	-0.0247	0.0109
15.	Over 10 years	0.0331	0.0059	0.0215	0.0447
	Experience at last job change that lasted:				
16.	Up to1 year	-0.0014	0.0015	-0.0044	0.0016
17.	2 to 5 years	-0.0005	0.0016	-0.0036	0.0026
18.	6 to 10 years	-0.0023	0.0026	-0.0072	0.0028
19.	Over 10 years	-0.0011	0.0039	-0.0086	0.0066
	Geographical location:				
20.	Northeast	0.0425	0.0208	0.0009	0.0821
21.	North Central	-0.0333	0.0170	-0.0672	0.0009
22.	South	-0.0181	0.0139	-0.0450	0.0089
23.	Living in SMSA	0.0526	0.0144	0.0236	0.0809
24.	County unemp. rate	-0.0032	0.0020	-0.0072	0.0008
	Race:				
25.	Black	-0.2640	0.0261	-0.3141	-0.2104
26.	Hispanic	-0.0078	0.0458	-0.0982	0.0781
	Cohort effects:				
27.	Age 15 or less in 1975	0.4329	0.0855	0.2668	0.5982
28.	Age 16 to 25 in 1975	0.2771	0.0833	0.1081	0.4354
29.	Age 26 to 35 in 1975	0.2114	0.0810	0.0525	0.3677
30.	Age 36 to 45 in 1975	0.2193	0.0809	0.0538	0.3799

Table 6: (Continued)

	Variable	Mean	St. Dev.	Range	
				Min	Max
	Industry:				
31.	Construction	0.2266	0.0231	0.1814	0.2715
32.	Manufacturing	0.3650	0.0188	0.3283	0.4006
33.	Trans., Comm., etc.	0.3995	0.0231	0.3544	0.4440
34.	Wholesale and Retail Trades	0.2377	0.0193	0.1996	0.2764
35.	Finance	0.3095	0.0301	0.2523	0.3697
36.	Business and Repair Services	0.1942	0.0258	0.1455	0.2463
37.	Personal Services	0.2071	0.0313	0.1473	0.2687
38.	Professional	0.2423	0.0248	0.1940	0.2901
39.	Public Administration	0.3522	0.0234	0.3064	0.3985
	Time Effects:				
40.	Year 1975	0.5761	0.0425	0.4930	0.6614
41.	Year 1976	0.5659	0.0399	0.4880	0.6434
42.	Year 1977	0.5258	0.0380	0.4524	0.6019
43.	Year 1978	0.4780	0.0356	0.4088	0.5478
44.	Year 1979	0.4382	0.0341	0.3717	0.5067
45.	Year 1980	0.4212	0.0320	0.3580	0.4836
46.	Year 1981	0.3788	0.0301	0.3190	0.4381
47.	Year 1982	0.3263	0.0297	0.2684	0.3851
48.	Year 1983	0.2777	0.0277	0.2229	0.3320
49.	Year 1984	0.2357	0.0262	0.1853	0.2872
50.	Year 1985	0.2647	0.0254	0.2150	0.3151
51.	Year 1986	0.2199	0.0235	0.1741	0.2659
52.	Year 1987	0.1468	0.0232	0.1026	0.1939
53.	Year 1988	0.1231	0.0219	0.0799	0.1660
54.	Year 1989	0.1056	0.0220	0.0632	0.1491
55.	Year 1990	0.0578	0.0202	0.0185	0.0966
56.	Year 1991	0.0348	0.0201	-0.0045	0.0739

Table 7: Estimates of the Stochastic Elements for High School Graduates

Variable		Mean	St. Dev.	Range	
				Min	Max
Covariance Matrix of White Noises (element of Σ):					
1.	ρ_{uv}	0.0090	0.0099	-0.0143	0.0277
2.	$\rho_{u\xi}$	-0.0472	0.0155	-0.0696	-0.0200
3.	$\rho_{v\xi}$	-0.0282	0.0052	-0.0394	-0.0183
4.	σ_{ξ}^2	0.2024	0.0024	0.1976	0.2073
Correlations of Individual Specific Effects (elements of Δ_ρ):					
5.	$\rho_{\alpha_y^0 \alpha_m^0}$	-0.0121	0.0662	-0.1731	0.0831
6.	$\rho_{\alpha_w \alpha_y^0}$	0.4193	0.0521	0.3123	0.5301
7.	$\rho_{\alpha_y \alpha_y^0}$	0.5843	0.0988	0.3482	0.7148
8.	$\rho_{\alpha_m \alpha_y^0}$	-0.0957	0.0865	-0.2260	0.0605
9.	$\rho_{\alpha_w \alpha_m^0}$	0.0114	0.0468	-0.1082	0.0981
10.	$\rho_{\alpha_y \alpha_m^0}$	-0.2517	0.1578	-0.4480	0.0641
11.	$\rho_{\alpha_m \alpha_m^0}$	-0.0560	0.0384	-0.1312	0.0243
12.	$\rho_{\alpha_y \alpha_w}$	0.3351	0.0315	0.2821	0.4046
13.	$\rho_{\alpha_m \alpha_w}$	-0.5234	0.0860	-0.6179	-0.3495
14.	$\rho_{\alpha_y \alpha_m}$	-0.4304	0.1516	-0.6418	-0.1613

Table 8: Participation and Mobility Equations for College Graduates

	Variable	Participation				Mobility			
		Mean	St. Dev.	Range		Mean	St. Dev.	Range	
				Min	Max			Min	Max
1.	Constant	-1.5795	0.5380	-2.6852	-0.5840	-1.0151	0.2494	-1.5113	-0.5269
2.	Education	0.1146	0.0245	0.0690	0.1645	-0.0129	0.0100	-0.0329	0.0062
3.	Experience at $t-1$	0.0660	0.0152	0.0381	0.0959	-0.0368	0.0066	-0.0492	-0.0240
4.	Experience at $t-1$ squared	-0.0021	0.0003	-0.0027	-0.0015	0.0005	0.0002	0.0002	0.0008
5.	Seniority at $t-1$	–	–	–	–	-0.0878	0.0074	-0.1024	-0.0734
6.	Seniority at $t-1$ squared	–	–	–	–	0.0020	0.0003	0.0015	0.0026
7.	Participation at $t-1$	2.0046	0.0944	1.8178	2.1978	–	–	–	–
8.	Mobility at $t-1$	0.3336	0.1646	0.0111	0.6274	-0.9019	0.0552	-1.0133	-0.7953
	Family Characteristics:								
9.	Family other income	-0.0020	0.0018	-0.0052	0.0018	-0.0110	0.0021	-0.0152	-0.0071
10.	No. of Children	0.1615	0.0605	0.0400	0.2737	-0.0673	0.0188	-0.1041	-0.0308
11.	Children 1 to 2	-0.1562	0.1203	-0.3881	0.0822	0.0650	0.0337	-0.0014	0.1309
12.	Children 3 to 5	-0.0880	0.1212	-0.3291	0.1468	-0.0138	0.0392	-0.0911	0.0631
13.	Married	0.0892	0.1126	-0.1410	0.2959	-0.0771	0.0398	-0.1523	0.0006
	Geographical Location:								
14.	Northeast	0.0152	0.0853	-0.1322	0.2038	-0.0121	0.0242	-0.0600	0.0348
15.	North Central	0.0575	0.0879	-0.1250	0.2133	-0.0244	0.0237	-0.0705	0.0221
16.	South	-0.0301	0.0775	-0.1809	0.1237	0.0155	0.0235	-0.0291	0.0612
17.	Living in SMSA	-0.1068	0.0982	-0.2962	0.0876	0.0537	0.0333	-0.0123	0.1173
18.	County unemp. rate	-0.0034	0.0171	-0.0361	0.0311	-0.0009	0.0063	-0.0135	0.0113
	Race:								
19.	Black	-0.2672	0.1477	-0.5592	0.0328	0.0888	0.0439	0.0043	0.1764
20.	Hispanic	-0.6326	0.2704	-1.1633	-0.0897	-0.0444	0.1053	-0.2550	0.1553
	Cohort Effects:								
21.	Age 15 or less in 1975	0.5290	0.2813	0.0050	1.1356	-0.0386	0.1256	-0.2892	0.2046
22.	Age 16 to 25 in 1975	0.3074	0.3123	-0.2943	0.9407	-0.0346	0.1222	-0.2779	0.2015
23.	Age 26 to 35 in 1975	0.8920	0.3007	0.3360	1.5230	-0.0380	0.1098	-0.2535	0.1751
24.	Age 36 to 45 in 1975	0.5458	0.2543	0.0587	1.0604	-0.0039	0.0966	-0.1951	0.1879

Table 8: (Continued)

	Variable	Participation				Mobility			
		Mean	St. Dev.	Range		Mean	St. Dev.	Range	
				Min	Max			Min	Max
	Industry:								
25.	Construction	–	–	–	–	0.8457	0.1050	0.6367	1.0563
26.	Manufacturing	–	–	–	–	0.9323	0.0920	0.7549	1.1143
27.	Trans., Comm., etc.	–	–	–	–	0.9094	0.1010	0.7063	1.1087
28.	Wholesale & Retail Trades	–	–	–	–	1.0468	0.0918	0.8676	1.2280
29.	Finance	–	–	–	–	1.0297	0.0996	0.8408	1.2290
30.	Business & Repair Services	–	–	–	–	1.0211	0.1009	0.8232	1.2238
31.	Personal Services	–	–	–	–	0.9327	0.1199	0.6973	1.1652
32.	Professional	–	–	–	–	0.9019	0.0896	0.7215	1.0770
33.	Public Administration	–	–	–	–	0.7583	0.0973	0.5668	0.9428
	Time Effects:								
34.	Year 1976	0.5046	0.2361	0.0446	0.9692	-0.1597	0.1104	-0.3817	0.0549
35.	Year 1977	0.6716	0.2324	0.2115	1.1274	0.4063	0.0917	0.2310	0.5893
36.	Year 1978	0.5539	0.2332	0.0952	0.9937	0.3363	0.0894	0.1613	0.5135
37.	Year 1979	0.5600	0.2332	0.0959	1.0053	0.3315	0.0876	0.1618	0.5042
38.	Year 1980	0.5975	0.2279	0.1564	1.0636	0.3545	0.0842	0.1880	0.5217
39.	Year 1981	0.7175	0.2317	0.2705	1.1717	-0.0934	0.0900	-0.2730	0.0827
40.	Year 1982	0.4096	0.1978	0.0132	0.7923	-0.0768	0.0872	-0.2448	0.0953
41.	Year 1983	0.5422	0.2042	0.1365	0.9254	-0.1023	0.0871	-0.2716	0.0679
42.	Year 1984	0.3945	0.1770	0.0440	0.7291	-0.0883	0.0850	-0.2500	0.0793
43.	Year 1985	0.2879	0.1776	-0.0611	0.6262	0.0075	0.0838	-0.1615	0.1693
44.	Year 1986	0.1038	0.1698	-0.2194	0.4463	-0.0259	0.0816	-0.1907	0.1325
45.	Year 1987	-0.0052	0.1708	-0.3391	0.3365	-0.0763	0.0833	-0.2433	0.0868
46.	Year 1988	0.1342	0.1645	-0.1795	0.4634	-0.0444	0.0805	-0.2004	0.1133
47.	Year 1989	0.2361	0.1677	-0.1059	0.5560	-0.0272	0.0806	-0.1909	0.1299
48.	Year 1990	0.0821	0.1613	-0.2671	0.3926	0.0485	0.0800	-0.1102	0.2020
49.	Year 1991	-0.0336	0.1486	-0.3204	0.2702	0.0352	0.0770	-0.1159	0.1865

Table 9: Wage Equation for College Graduates

	Variable	Mean	St. Dev.	Range	
				Min	Max
1.	Constant	8.3258	0.1347	8.0614	8.5874
2.	Education	0.0411	0.0054	0.0302	0.0516
3.	Experience	0.0580	0.0032	0.0518	0.0643
4.	Experience squared	-0.0013	0.0001	-0.0015	-0.0012
5.	Seniority	0.0518	0.0029	0.0460	0.0576
6.	Seniority squared	-0.0005	0.0001	-0.0007	-0.0004
	Job switch variables in first sample year:				
7.	Dummy for job change job in 1st year	0.0769	0.0673	-0.0519	0.2062
8.	Experience at $t-1$ if variable 7=1	0.0108	0.0044	0.0020	0.0192
	No. of switches of jobs that lasted:				
9.	Up to 1 year	0.2240	0.0172	0.1905	0.2572
10.	2 to 5 years	0.1648	0.0189	0.1274	0.2018
11.	6 to 10 years	0.3231	0.0683	0.1861	0.4572
12.	Over 10 years	0.4717	0.0869	0.3031	0.6425
	Seniority at last job change that lasted:				
13.	2 to 5 years	0.0567	0.0070	0.0432	0.0709
14.	6 to 10 years	0.0111	0.0097	-0.0079	0.0303
15.	Over 10 years	0.0062	0.0055	-0.0050	0.0166
	Experience at last job change that lasted:				
16.	Up to 1 year	-0.0071	0.0016	-0.0102	-0.0040
17.	2 to 5 years	-0.0058	0.0016	-0.0090	-0.0027
18.	6 to 10 years	-0.0025	0.0025	-0.0073	0.0024
19.	Over 10 years	-0.0026	0.0033	-0.0090	0.0036
	Geographical location:				
20.	Northeast	0.0497	0.0157	0.0188	0.0802
21.	North Central	-0.0501	0.0141	-0.0784	-0.0232
22.	South	-0.0114	0.0131	-0.0368	0.0140
23.	Living in SMSA	0.0359	0.0128	0.0108	0.0616
24.	County unemp. rate	-0.0042	0.0020	-0.0082	-0.0003
	Race:				
25.	Black	-0.2343	0.0358	-0.3065	-0.1660
26.	Hispanic	0.0079	0.0717	-0.1384	0.1478
	Cohort effects:				
27.	Age 15 or less in 1975	-0.0172	0.0793	-0.1751	0.1323
28.	Age 16 to 25 in 1975	-0.2371	0.0822	-0.3947	-0.0794
29.	Age 26 to 35 in 1975	-0.1633	0.0743	-0.3038	-0.0145
30.	Age 36 to 45 in 1975	0.0261	0.0735	-0.1132	0.1717

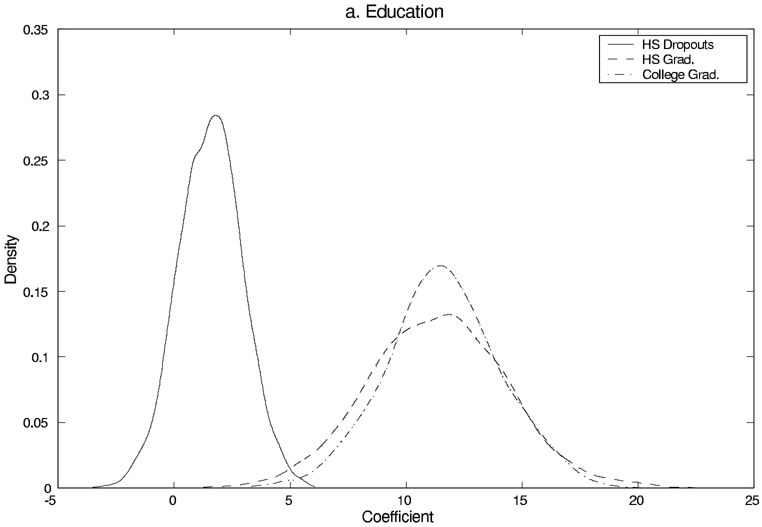
Table 9: (Continued)

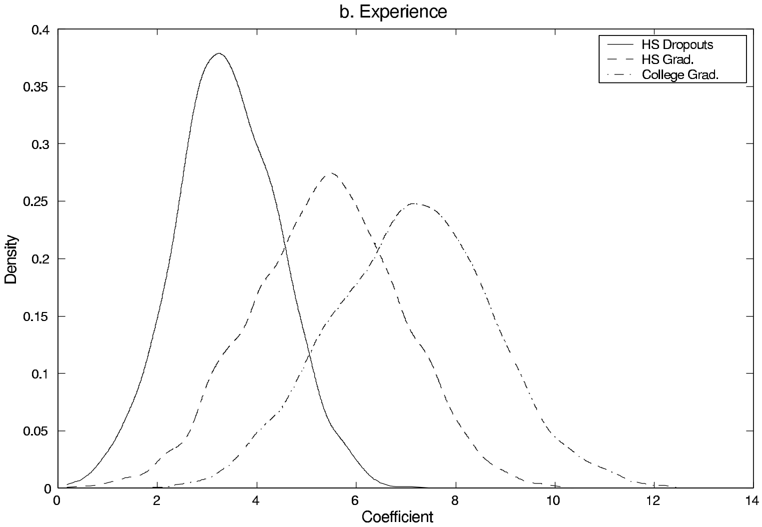
	Variable	Mean	St. Dev.	Range	
				Min	Max
	Industry:				
31.	Construction	0.3495	0.0285	0.2921	0.4050
32.	Manufacturing	0.4559	0.0207	0.4154	0.4963
33.	Trans., Comm., etc.	0.3875	0.0262	0.3360	0.4395
34.	Wholesale and Retail Trades	0.2969	0.0215	0.2546	0.3391
35.	Finance	0.3923	0.0272	0.3394	0.4458
36.	Business and Repair Services	0.3172	0.0254	0.2688	0.3674
37.	Personal Services	0.1864	0.0348	0.1187	0.2520
38.	Professional	0.3774	0.0190	0.3401	0.4143
39.	Public Administration	0.3935	0.0250	0.3438	0.4419
	Time Effects:				
40.	Year 1975	0.3573	0.0430	0.2741	0.4417
41.	Year 1976	0.3543	0.0407	0.2725	0.4332
42.	Year 1977	0.3017	0.0380	0.2255	0.3759
43.	Year 1978	0.2819	0.0365	0.2100	0.3533
44.	Year 1979	0.2510	0.0345	0.1846	0.3200
45.	Year 1980	0.2025	0.0329	0.1394	0.2675
46.	Year 1981	0.1794	0.0304	0.1201	0.2390
47.	Year 1982	0.1789	0.0292	0.1215	0.2362
48.	Year 1983	0.1801	0.0272	0.1263	0.2335
49.	Year 1984	0.1174	0.0253	0.0673	0.1678
50.	Year 1985	0.1388	0.0242	0.0907	0.1865
51.	Year 1986	0.1484	0.0231	0.1037	0.1935
52.	Year 1987	0.1218	0.0219	0.0795	0.1640
53.	Year 1988	0.0939	0.0209	0.0539	0.1348
54.	Year 1989	0.0413	0.0196	0.0030	0.0800
55.	Year 1990	0.0496	0.0191	0.0118	0.0861
56.	Year 1991	0.0065	0.0183	-0.0298	0.0420

Table 10: Estimates of the Stochastic Elements for College Graduates

Variable		Mean	St. Dev.	Range	
				Min	Max
Covariance Matrix of White Noises (element of Σ):					
1.	ρ_{uv}	-0.0005	0.0113	-0.0217	0.0188
2.	$\rho_{u\xi}$	-0.0496	0.0124	-0.0672	-0.0205
3.	$\rho_{v\xi}$	0.0013	0.0075	-0.0111	0.0161
4.	σ_{ξ}^2	0.2062	0.0023	0.2016	0.2104
Correlations of Individual Specific Effects (elements of Δ_ρ):					
5.	$\rho_{\alpha_y^0 \alpha_m^0}$	0.8040	0.0556	0.7024	0.9005
6.	$\rho_{\alpha_w \alpha_y^0}$	0.1335	0.0757	0.0169	0.2714
7.	$\rho_{\alpha_y \alpha_y^0}$	0.5716	0.0286	0.5190	0.6224
8.	$\rho_{\alpha_m \alpha_y^0}$	-0.6044	0.0773	-0.7595	-0.4892
9.	$\rho_{\alpha_w \alpha_m^0}$	-0.1450	0.0884	-0.2586	0.0403
10.	$\rho_{\alpha_y \alpha_m^0}$	0.2896	0.0429	0.2268	0.3845
11.	$\rho_{\alpha_m \alpha_m^0}$	-0.4234	0.0789	-0.5691	-0.2668
12.	$\rho_{\alpha_y \alpha_w}$	0.2174	0.0553	0.1066	0.3017
13.	$\rho_{\alpha_m \alpha_w}$	-0.5352	0.0590	-0.6371	-0.4131
14.	$\rho_{\alpha_y \alpha_m}$	-0.5061	0.0656	-0.6172	-0.3874

Figure 1: Participation Equation





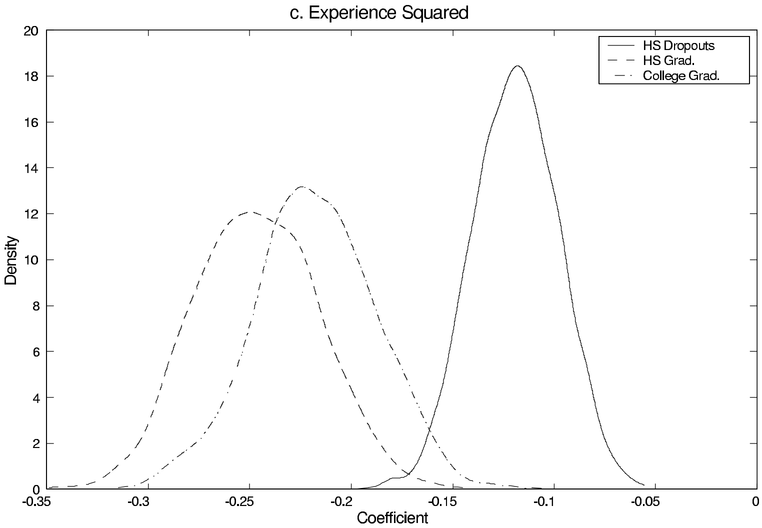


Figure 2: Mobility Equation

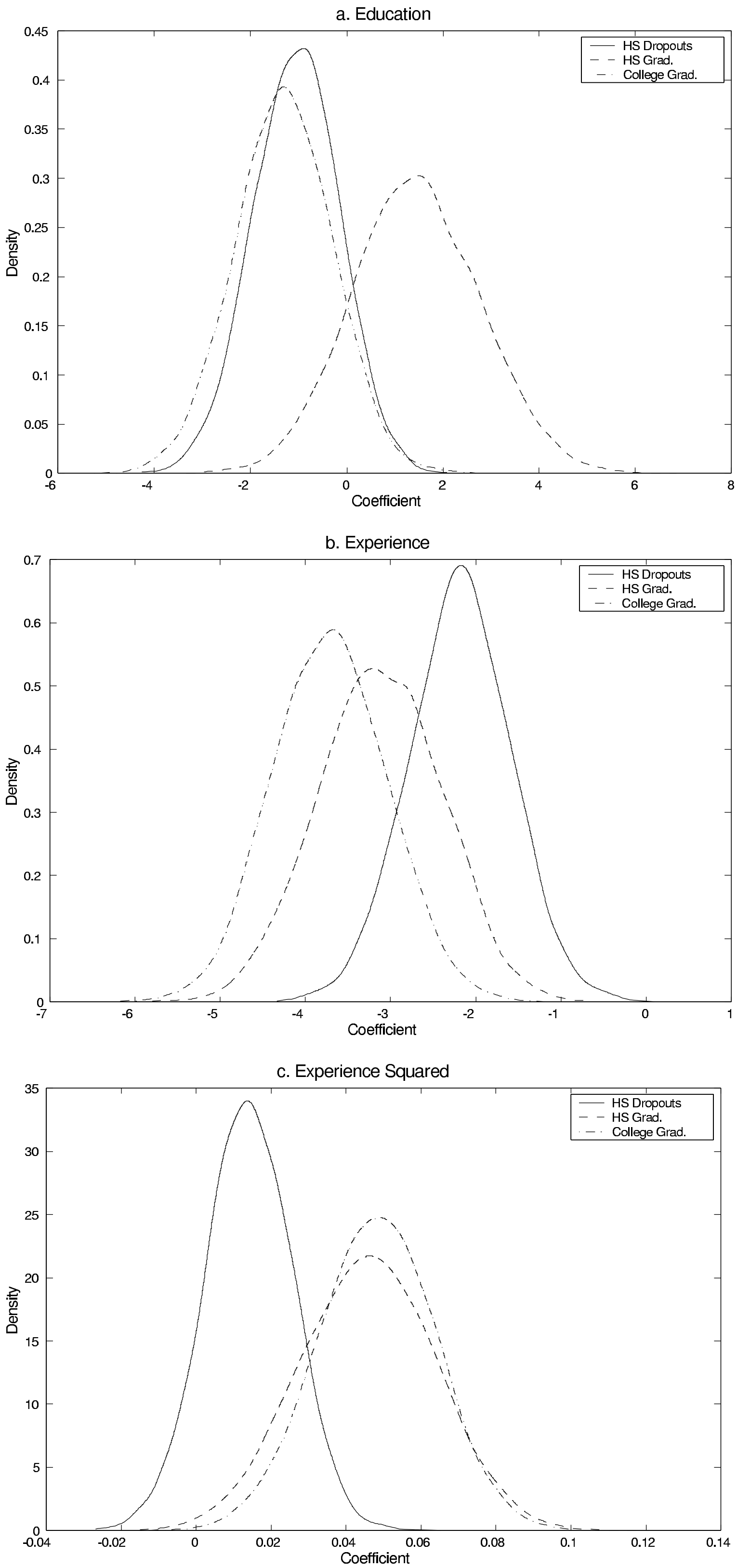
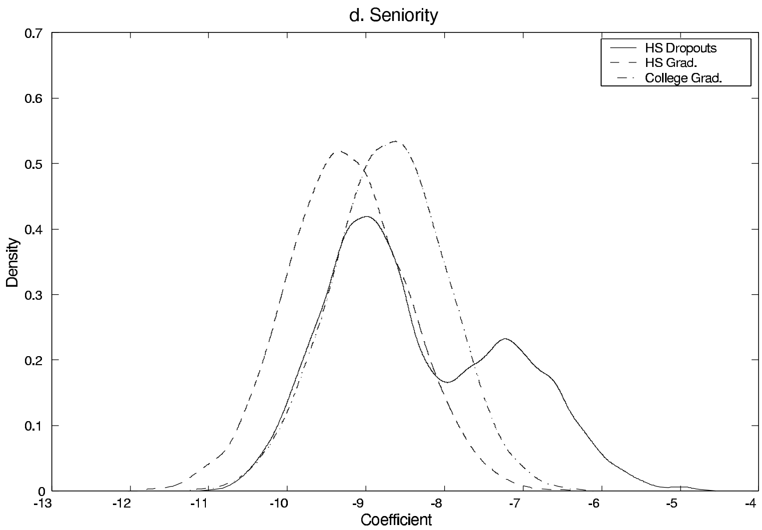


Figure 2: (Continued)



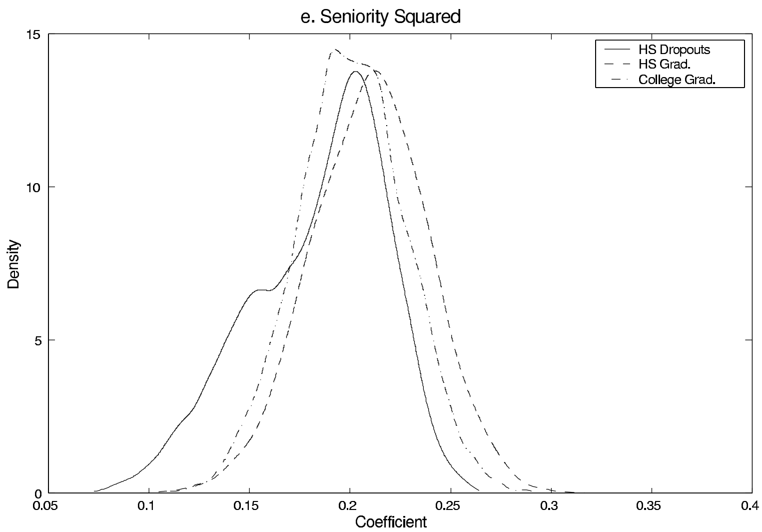


Figure 3: Return of Wage to Education

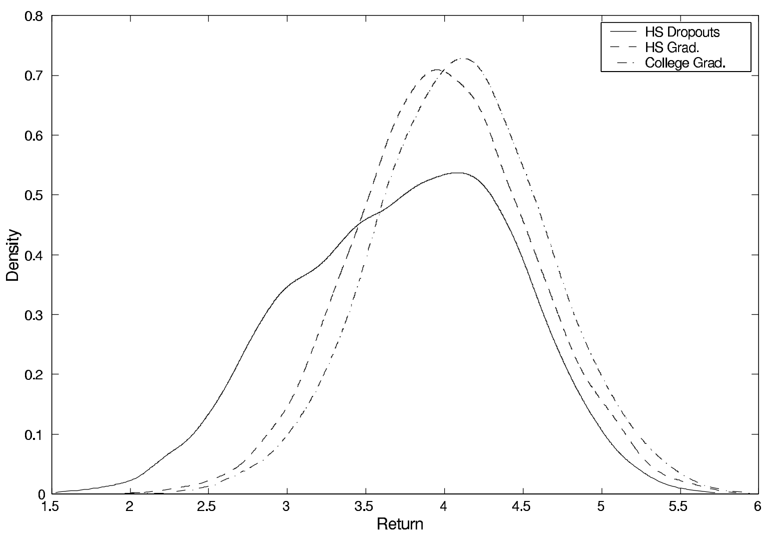
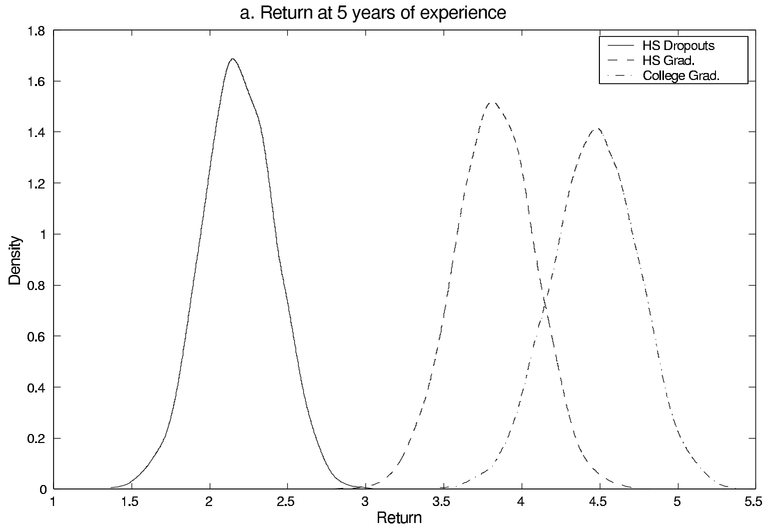
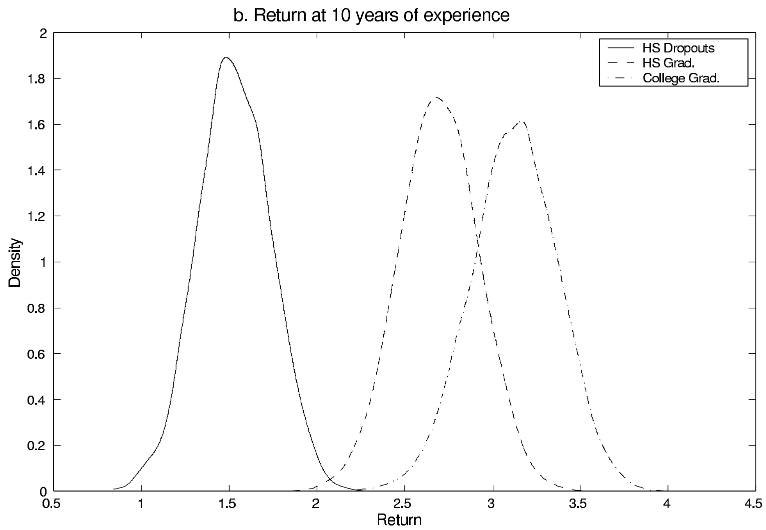


Figure 4: Return of Wage to Experience





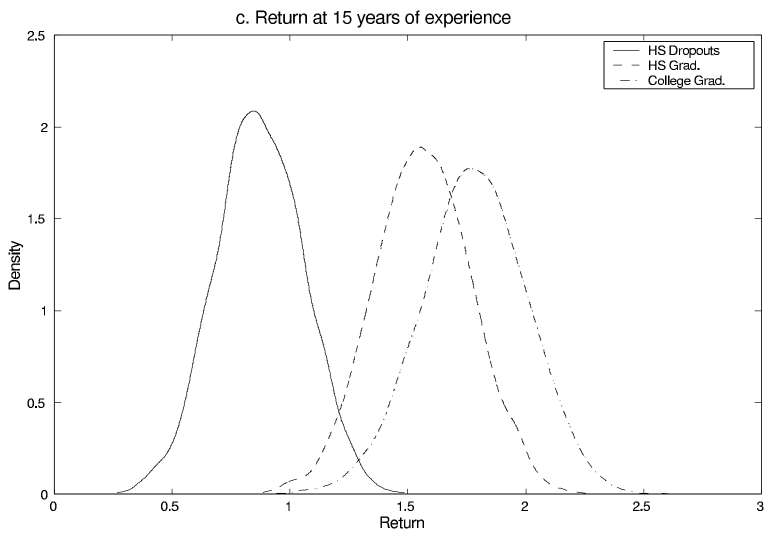


Figure 5: Return of Wage to Seniority

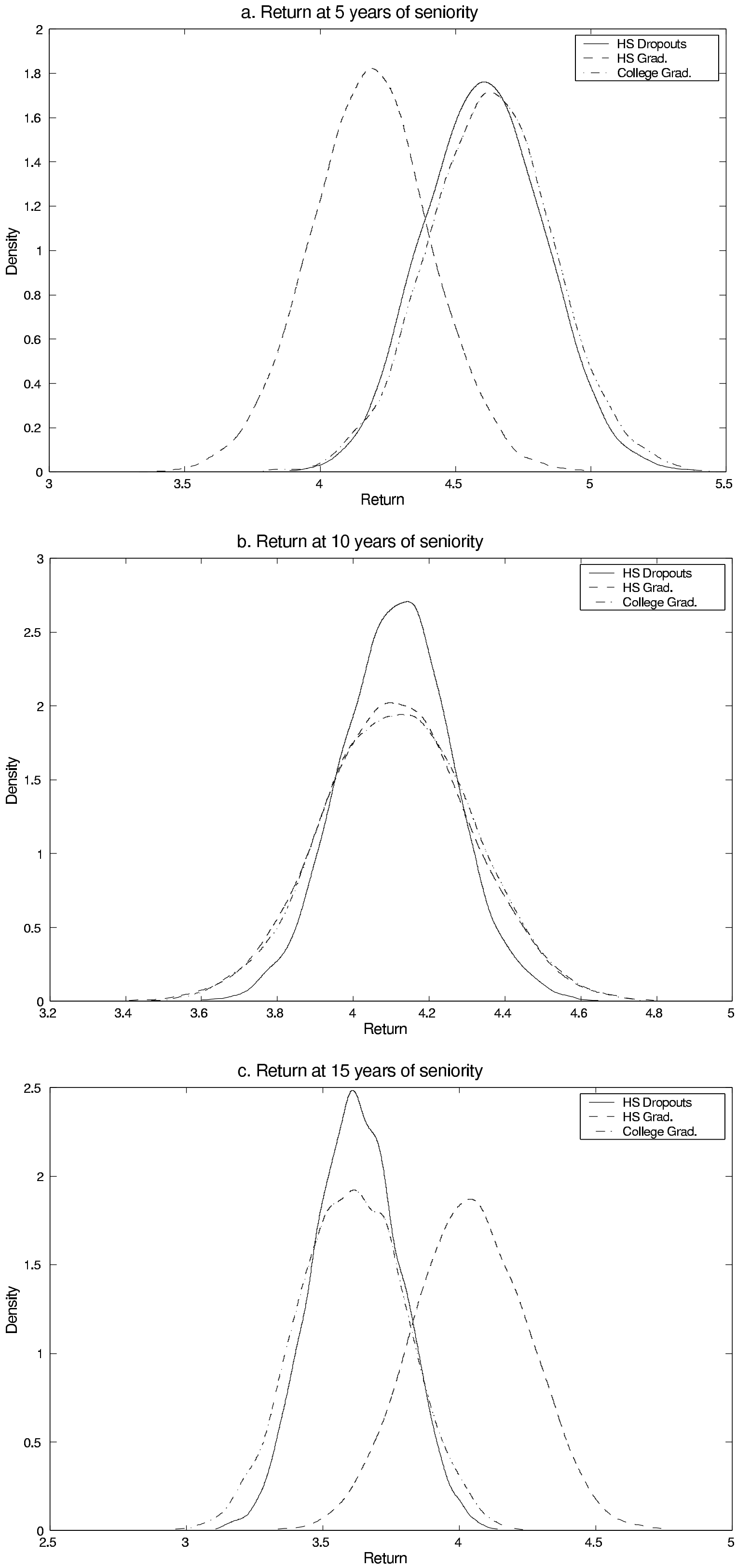


Figure 6: High School Dropouts–Wage Change for Typical Mobility Pattern

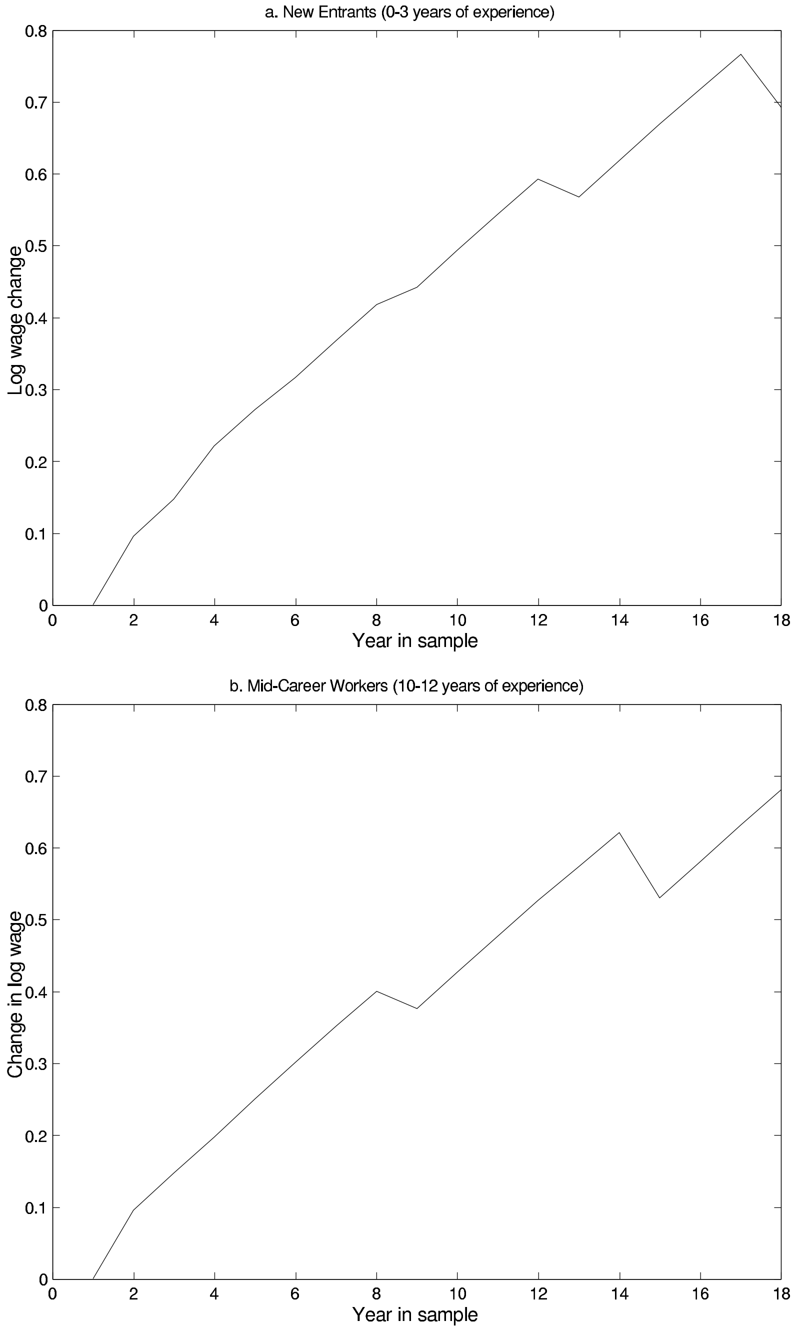


Figure 7: High School Graduates–Wage Change for Typical Mobility Pattern

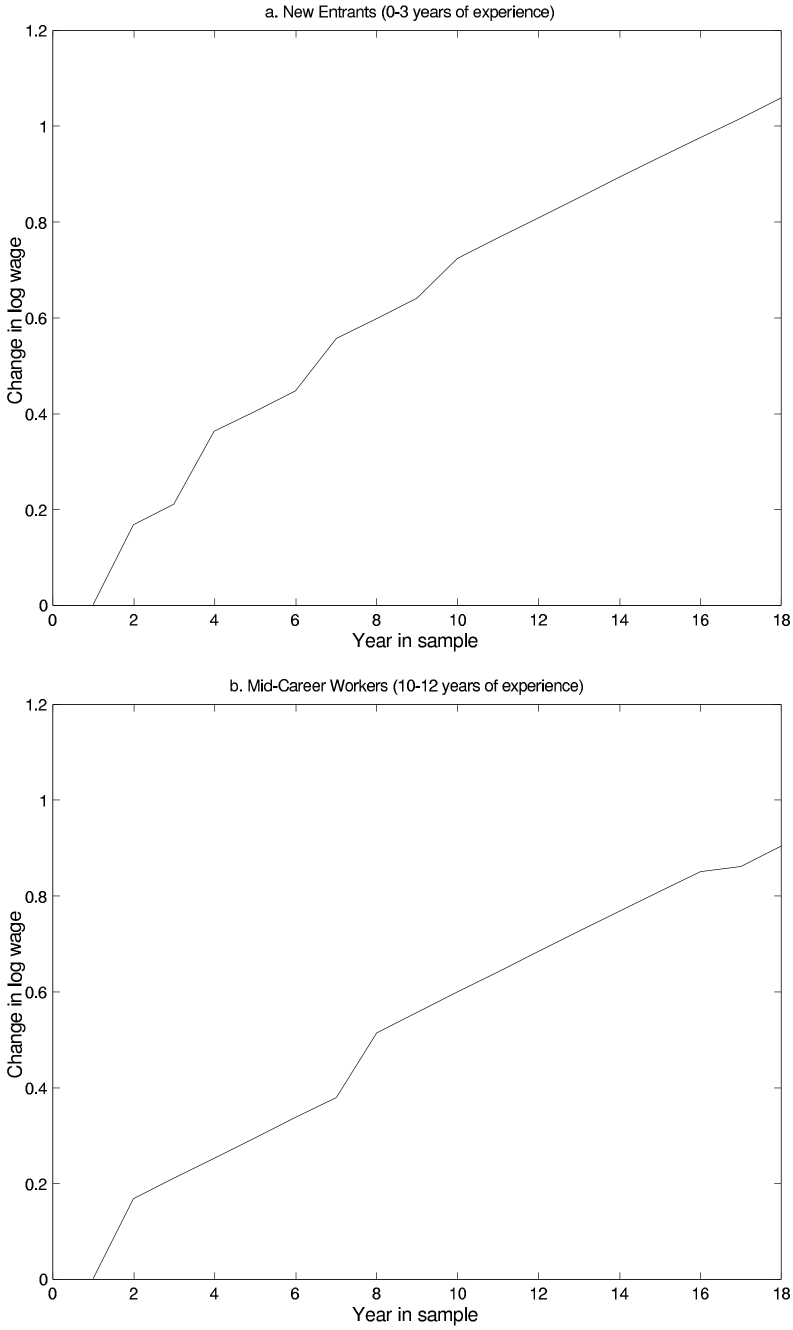


Figure 8: College Graduates–Wage Change for Typical Mobility Pattern