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

Moshe Buchinsky[‡]
UCLA, CREST–INSEE and NBER

Denis Fougère
CNRS, CREST–INSEE, CEPR and IZA

Francis Kramarz
CREST–INSEE, CEPR and IZA

Rusty Tchernis
Georgia State University and NBER

September, 2009

Abstract

In this paper, we expand on the seminal work of Altonji and Shakotko (1987) and Topel (1991) and reinvestigate the returns to seniority in the U.S. We begin with the same wage equation as in previous studies. We extend the model of Hyslop (1999) and explicitly model the participation/employment and interfirm mobility decisions, which, in turn, define the individual's experience and seniority. We introduce into the wage equation a summary of the workers' entire career path. The three-equation system is estimated simultaneously using data from the Panel Study of Income Dynamics (PSID). We find that for each of the three education groups studied the returns to seniority are larger than those previously found in the literature.

[†]We would like to thank Matthew Baird, Claudia Buchinsky, Gary Chamberlain, Andrew Foster, Oded Galor, Hank Farber, Richard Freeman, Jinyong Hahn, Mark Machina, Magali Ponçon-Beffy, Jean-Marc Robin, John Rust, Bernard Salanié, Duncan Thomas and Robert Topel for their comments and discussions. We also thank the editor as well as the three referees whose comments and suggestions considerably improved the paper. The participants of seminars and conferences in Amsterdam, Brown, Crest (Malinvaud seminar), ESEM 2001 in Lausanne, Harvard, Maryland, Michigan, Pennsylvania, Prague, Tel-Aviv, Toulouse, UCLA, UCSD, Wisconsin and Yale, provided very helpful comments. Remaining errors are ours.

[‡]Corresponding author, Department of Economics, University of California, Los Angeles, CA 90095-1477. Email: buchinsky@econ.ucla.edu.

1 Introduction

There is an ongoing debate in the literature about the relative importance of the returns to firm-specific seniority (or tenure) versus general labor market experience. Since the theoretical models proposed in the literature give rise to both alternatives, the resolution of the debate falls to the empirical studies. To date, the empirical literature has struggled to reach a consensus. This paper offers a new perspective to this debate and provides several explanations to conflicting findings in past research. While there has been little disagreement about the importance of the return to experience, there has been an ongoing debate about the contribution of the return to tenure to wage growth among individuals in the U.S. For example, Topel (1991) finds strong evidence for large returns to seniority. In contrast, Altonji and Shakotko (1987) (AS, hereafter) and Altonji and Williams (2005) (AW, hereafter) find little returns to seniority.

In previous literature, much of the focus on the returns to seniority has concentrated on the possible endogeneity of job changes and its effect on the estimated return to tenure. This paper goes one step further and considers the possible endogeneity of labor market experience, and its potential effects on the estimated returns to both tenure and experience. To address this key issue we develop a structural model in which individuals make two key decisions, namely employment (or participation) and interfirm mobility. In turn, these decisions influence the observed outcome of interest, namely wages. Within this model we revisit the issue regarding the magnitude of the returns to seniority in the U.S. and offer new perspectives. Most importantly, the paper concludes that the returns to seniority are higher than those previously reported in the literature, including those reported by Topel (1991).

Our formulation of the wage function, the general Mincer's wage specification, is the same as that taken by virtually all of the literature on this topic. In order to account for the endogenous decisions of participation and mobility, we extend the model of Hyslop (1999) and broaden it to include forward looking dynamic optimization search perspectives. This model gives rise to two *reduced-form* decision equations: (1) a participation/employment;[1] and (2) an interfirm mobility. In this approach, experience and seniority are fully endogenized because they are direct outcomes of the employment and interfirm mobility decisions, respectively. Hence, the

[1] In both the theoretical model and empirical analysis, participation is the same as employment. We, therefore, use these two terms interchangeably throughout the paper.

model accounts for the potential selection biases that stem from these endogenous decisions, and thus allows one to consistently estimate the parameters associated with the wage function, including the returns to tenure and experience.

To account for individuals' unobserved heterogeneity we include person-specific random correlated effects in the estimated equations. Controlling for unobserved heterogeneity in this fashion creates important links between the individual's decisions and the outcome of interest (i.e., wage) that, if ignored, could lead to severe biases in the estimated parameters of interest. Additional links between the equations are present by the fact that the idiosyncratic random components in the equation are also allowed to be contemporaneously correlated.

Our approach offers a unified framework with which one can address and reexamine results that have been previously found in the literature. This is because the statistical assumptions that are adopted in this study—along with the fact that seniority and experience are endogenized—incorporate the most important elements from many previous studies, particularly those studies by Topel (1991) and AW.[2]

We use data from the Panel Study of Income Dynamics (PSID) and estimate the model for three separate education groups: high school drop-outs, high school graduates and college graduates. We adopt a Bayesian approach and employ Markov Chain Monte Carlo (MCMC) methods for estimating the joint posterior distribution for the model's parameters.

The results indicate that, while the estimated returns to experience are somewhat higher than those previously found in the literature, they are of similar magnitude. In contrast, the estimates of the returns to seniority are much higher than those previously obtained, including those obtained by Topel (1991). Consequently, our estimates of total within-job wage growth are significantly higher than Topel's estimates, and those reported by Abraham and Farber (1987) (AF, hereafter). This is true for all three education groups analyzed in this study.

Our study also sheds light on several important factors leading to the apparent differences between our estimates and those obtained in previous studies. First, the results of our study highlight the importance of explicitly modeling the employment and the mobility decisions, which, in turn, define experience and seniority. Second, we establish the need to account for unobserved heterogeneity prevailing in the participation and mobility decisions, as well as in the wage function. Finally, we demonstrate

[2] There are a number of other papers in the literature, including Dustmann and Meghir (2005) and Neal (1995), which analyzed similar questions.

the need to explicitly control for job-specific components in the wage function, through the introduction of a function that serves as a summary statistic for the individual's specific career path. This function captures the overall effect of the worker's specific career path on his/her market wage. We find that the magnitude of the estimated returns changes markedly when we account for this factor, but the qualitative results remain similar. This indicates that the timing of a job change during the course of one's career is important for his/her wage trajectory.

Our findings are different from those previously obtained in the literature. Hence, we provide a number of robustness checks that are designed to ensure that these results do not stem from reasons that are auxiliary to the substantive issues examined here. First, we use two alternative prior distributions for the model's parameters associated with the returns to tenure and experience, centering them around the results obtained by Topel (1991) and AS; the results remain virtually unchanged. Also, our sample extract is somewhat different from that used by Topel (1991) and AS. We provide clear evidence that the substantially distinct results obtained in this study are not driven by the difference in samples. Specifically, when we apply the methods of Topel (1991) and AS, we are able to closely replicate their results using our data extract.

The remainder of the paper is organized as follows. In Section 2, we provide a brief review of the literature and highlight the crucial differences between the various approaches. In Section 3, we present the model, which extends the model from Hyslop (1999), while also incorporating other vital elements from the search literature. In Section 4, we explain the econometric specification and briefly discuss the estimation method. In Section 5 we provide a brief discussion of the data extract used in this study. In Section 6 we present the empirical results and their implications. In Section 7 we present concluding remarks. In a web appendix we provide mathematical proofs, details about the numerical algorithms used and further descriptive analysis of the modelling approach.[3]

2 Literature Review

There has been a long lasting debate in the literature on the magnitude of the returns to seniority in the U.S. The results of Topel (1991), which indicate that there are large returns to seniority, stand in stark contrast to those of AS and AF, which find virtually

[3] The Appendices are available in the Review of Economic Studies website under *Supplementary Material* and in Buchinsky et al. (2005).

no returns to seniority in the U.S. Topel (1991) finds strong evidence that the costs of displacement are highly correlated with prior job tenure, a phenomenon that may stem from the fact that: (a) wages rise with seniority; or (b) tenure merely acts as a proxy for the quality of the job match.[4] In order to examine this phenomenon, Topel uses the following prototype model of wage determination:

$$y_{ijt} = X_{ijt}\beta_1 + T_{ijt}\beta_2 + \epsilon_{ijt}, \tag{1}$$

where y_{ijt} is the log wage of individual i in job j at time t, X_{ijt} denotes experience, and T_{ijt} denotes seniority. The residual term ϵ_{ijt} is decomposed into three components:

$$\epsilon_{ijt} = \phi_{ijt} + \mu_i + v_{ijt}, \tag{2}$$

where ϕ_{ijt} is specific to the individual-job pair, μ_i is a term that reflects the individual's ability and v_{ijt} is an idiosyncratic term representing market-wide random shocks and/or measurement errors. The main problem in estimating the parameters of interest β_1 and β_2 stems from the fact that ϕ_{ijt} is likely to be correlated with experience and/or tenure. In particular, because the match component ϕ_{ijt} is likely to increase with tenure, the estimate of β_2, say $\widehat{\beta}_2$, will be upward biased within a simple ordinary least-squares (OLS) estimation of (1). However, Topel also provides a convincing argument regarding the composition of "movers" and "stayers" in the data set which implied that, if the returns to seniority are positive (i.e., $\beta_2 > 0$), then $\widehat{\beta}_2$ will be downward biased. Under the assumption that experience at the entry level is exogenous and, hence, uncorrelated with the error terms, Topel obtains an unbiased estimate for $\beta_1 + \beta_2$ and an upward biased estimate for β_1 (due to the selection bias induced by not modeling the mobility decisions). Hence, he argues that his estimate of β_2, $\widehat{\beta}_2 = .0545$, provides a lower bound for the returns to seniority.[5] Note that if experience is not exogenous and is positively (negatively) correlated with ϕ_{ijt} because most mobile workers voluntarily (involuntarily) change jobs for better (worse) matches, then the estimate of β_1, say $\widehat{\beta}_1$, will be upward (downward) biased.

In contrast, AS use an instrumental variables approach in which it is assumed

[4] An additional possibility, not explored by Topel, is that workers with long tenures are simply more able.

[5] Topel (1991) also examined two additional sources of potential biases in the estimates of $\beta_1+\beta_2$. Nevertheless, he finds that accounting for these potential biases had a very small effect on the estimate for β_2.

(in Topel's notation) that $\phi_{ijt} = \phi_{ij}$, i.e., the individual job-specific term is time-invariant. Under this assumption, deviation of seniority from its average in a specific job is a valid instrument for seniority. Since this method is a variant of Topel's two-step approach, it is not surprising that AS obtain an estimate for $\beta_1 + \beta_2$ that is similar to that obtained by Topel. Nevertheless, AS's procedure appears to induce an upward bias in the IV estimate for β_1, and hence a downward bias in the estimate for β_2. The problem is potentially magnified by two other factors: (a) measurement error problem in the tenure data used by AS; and (b) differences in the treatment of time trends in the regression.[6]

AF use a somewhat different set of assumptions. In particular, they use completed tenure to proxy for the unobserved dimensions of the individual's, or job's, quality. A problem with their approach is that many of the workers in their data extract have censored spells of employment.[7] While AW specify a model that is closer in spirit to Topel's model, their approach differs in some meaningful way. AW crucially rely on the assumption that the match effect ϕ and time are independent, i.e., $\text{Cov}(t, \phi) = 0$, conditional on experience (or experience and tenure). This assumption is problematic, especially in cases where workers have had more time to find jobs with higher ϕ. Additionally, t may also be correlated with μ in (2) because of changes in the sample composition.

One important conclusion from both Topel (1991) and AW is that individual heterogeneity is an important factor of the wage growth process. It appears that some of the reduction in the upward bias in the estimate for β_1 in Topel (1991) is due to a reduction in the bias that stems from individual heterogeneity.

In a recent paper, Dustmann and Meghir (2005) (DM, hereafter) allow for three different sources of returns due to the accumulation of specific human capital, namely experience, sector-specific seniority and firm-specific seniority.[8] In order to estimate the returns to experience, they use data of displaced workers in their new jobs, assuming that such workers could not predict closure of an establishment more than

[6] Topel uses a specific index for the aggregate changes in real wages by using data from the CPS., while AS used a simple time trend. Consequently, the growth of quality of jobs, due to better matches over time, would cause an additional downward bias in the estimate of β_2.

[7] Another source of difference arises because AF use a quadratic polynomial in experience when estimating the log wage equation, whereas AS and Topel use a quartic specification.

[8] Neal (1995) and Parent (1999, 2000) also focus their investigation on the importance of sector- and firm-specific human capital. We abstract from the sector-specific returns in order to avoid modeling sectorial choices.

a year in advance.[9] Furthermore, under the assumption that displaced workers have preferences for work similar to those that induced their sectorial choices, controlling for the endogeneity of experience also controls for the endogeneity of sector tenure. In a subsequent step DM estimate two reduced-form equations, one for experience and the other for participation. The residuals from these two regressions are used as regressors in the wage regression of the displaced workers. This allows DM to account for possible sample selection biases induced by restricting attention to only the individuals staying with their current employer. Using data from Germany and the U.S., DM find that the returns to tenure for both skilled and unskilled workers are large. The estimated returns to sector-specific tenure are much smaller, but statistically significant.

In his survey paper, Farber (1999) notes the importance of modelling some specific features of the mobility process. First, he shows that in the first few months of a job there is an increase in the probability of job separation, which decreases steadily thereafter. Farber provides strong evidence that contradicts the simple model of pure unobserved heterogeneity, suggesting that one must distinguish heterogeneity from duration dependence. He also finds strong evidence that: (a) firms tend to lay off less senior workers who have lower specific firm-capital; and (b) job losses result in substantial permanent earnings losses.[10]

3 The Model

Our model extends Hyslop's (1999) model by allowing workers to move directly from one job to another, a key feature in the search literature. We show that under some fairly general conditions—depending on the search cost, the mobility cost, and the shape of the utility function—job-to-job transitions may occur. Moreover, this structure leads to a first-order state dependence in the participation and mobility processes. Whether state dependence plays a major role in reality is largely an empirical question that is examined below. The decision to stay non-employed, stay employed in the same firm, move to an alternative job, or to become non-employed, depends on the values of several state-dependent reservation wages. This model gives rise to an econometric model with two selection equations (for the participation and interfirm

[9] For more on the empirical findings in the literature on displaced workers see also Addison and Portugal (1989), Jacobson, LaLonde and Sullivan (1993) and Farber (1999).

[10] See also Gibbons and Katz (1991) for a further discussion of the latter point.

mobility decisions) and a wage equation.

The wage function:

Our theoretical model builds on a specification of the wage function that has been adopted generally in the literature. That is, the observed log wage equation for individual i in job j at time t is:

$$\begin{aligned} w_{ijt} &= w^*_{ijt} \cdot \mathbf{1}(y_{it} = 1), \quad \text{and} \\ w^*_{ijt} &= x'_{wijt}\delta_0 + \varepsilon_{ijt}, \end{aligned} \tag{3}$$

where x_{wijt} is a vector of observed characteristics—including education, labor market experience and seniority (or tenure)—of the individual in his/her current job, $\mathbf{1}(\cdot)$ is the usual indicator function and $y_{it} = 1$ if the ith individual participates in the labor force at time t, and $y_{it} = 0$ otherwise. Consequently, the wage offer, w^*_{ijt}, is observed only if the individual chooses to work.

We decompose the error term ε_{ijt} into three components:

$$\varepsilon_{ijt} = J^W_{ijt} + \alpha_{wi} + \xi_{ijt}, \tag{4}$$

where α_{wi} is a person-specific correlated random effect, analogous to μ_i in (2), and ξ_{ijt} is a contemporaneous idiosyncratic error term. The term J^W_{ijt} is analogous to the term ϕ_{ijt} in (2), only that here it provides a summary statistic for the individual's work history and career. More precisely, J^W_{ijt} captures the timing and magnitude of all discontinuous jumps in the individual's wages that resulted from all job changes until date t.[11]

In principle, this function can be viewed as a full set of variables capturing all observed jumps in the data. However, in the empirical application, this would require estimation of a prohibitively large number of parameters. Thus, we approximate J^W_{ijt} by the following piece-wise linear function of experience and seniority at the time of

[11] In an analysis of matched employer-employee data for France, Abowd, Kramarz and Roux (2006) find that entry wages depend upon seniority in the previous job, as well as the number of previous jobs held by the individual.

a job change:[12]

$$J^W_{ijt} = (\phi^s_0 + \phi^e_0 e_{i0})\, d_{i1} + \sum_{l=1}^{M_{it}} \left[\sum_{k=1}^{4} (\phi_{k0} + \phi^s_k s_{i,t_l-1} + \phi^e_k e_{i,t_l-1})\, d_{kit_l} \right], \tag{5}$$

where d_{1it_l} equals 1 if the lth job of the ith individual lasted less than a year and equals 0 otherwise. Similarly, $d_{2it_l} = 1$ if the lth job of the ith individual lasted between two and five years, and equals 0 otherwise, $d_{3it_l} = 1$ if the lth job lasted between six and ten years, and equals 0 otherwise, $d_{4it_l} = 1$ if the lth job lasted more than ten years and equals 0 otherwise. Finally, M_{it} denotes the number of job changes by the ith individual at time t (not including the individual's first sample year). If an individual changed jobs in the first sample year, then $d_{i1} = 1$; otherwise $d_{i1} = 0$. The quantities e_{it_l} and s_{it_l} denote the individual's experience and seniority in year t_l, respectively, when individual i leaves job l. Note that while the ϕ's are fixed parameters, the size of the jumps (within each of the four brackets of seniority) may differ depending on the level of labor market experience and seniority at the time of a job change.[13]

Note that the J^W_{ijt} function generalizes ϕ_{ijt} in (2), and captures the initial conditions specific to the individual at the start of a new job. In other words, this function provides a measure of the opportunity wage of the worker if he/she were to move to a new job at that point in his/her career. Inclusion of actual (rather than potential) labor market experience as a determinant of the initial earnings at a new job allows one to distinguish between displaced workers, who went through a period of non-employment after displacement, and workers who moved directly from one job to another. Similarly, the inclusion of the seniority level at past jobs allows one to control for the quality of the past job matches. Whether the frequency of changing jobs and the individual's labor market attachment matters is an empirical question that we address below.[14]

[12] In a different context, Light and Ureta (1995) also addressed similar timing issue, but with respect to experience. They discuss the role of experience in the early stages of one's career and highlight the importance of controlling for the exact timing of work experience when estimating wage regressions.

[13] The J^W_{ijt} function contains a total of 13 identifiable parameters corresponding to the four brackets of seniority and the first sample year.

[14] Also, note that J^W_{ijt} is individual-job specific. In general, there are several ways to define a job. Our definition is a particular employment spell in one's career. Hence, it is possible that different individuals will have the same values for J^W_{ijt} even though they may not be employed at the same firm. This definition of a job is consistent with our modelling approach.

The timing of information revelation and decisions:
The model is in discrete time. At the *beginning* of period t, an employed worker receives a wage offer from his/her current firm. This wage offer consists of two components. While one component is certain, the other is random. The first part is $x'_{wijt}\delta_0 + J^W_{ijt} + \alpha_{wi}$. The second part, ξ_{ijt}, is revealed just before the *end* of period t. Its expected value is zero and it can be interpreted as a match-specific productivity shock or as an incentive component of pay with the outcome being determined by the worker's behavior during period t. Hence, at the *end* of period t, after the two components of w_{ijt} have been revealed, the worker decides whether to stay in the firm, move to another firm, or become non-employed.[15] It is clear that at the end of period t the worker can only form an expectation about the exact wage that he/she will be paid in the next period (period $t+1$) either at a potential new job j', or at the current job, j. These expectations, which the worker is fully informed of, are given by $x'_{wij',t+1}\delta_0 + J^W_{ij',t+1} + \alpha_{wi}$ and $x'_{wij,t+1}\delta_0 + J^W_{ijt+1} + \alpha_{wi}$, respectively. Note that while seniority in a new job j' is, by definition, zero, it may still be the case that $J^W_{ij',t+1}$ would be larger than $J^W_{ij,t+1}$. For simplicity we assume that, if one chooses to change jobs an outside wage offer arrives in the next period with probability 1.[16]
The cost structure:
A worker who moves to a new firm at the end of period t has to incur the moving cost of c_M. In reality, some components of these costs are paid at time t, while others are paid in period $t+1$. Typically, costs that are incurred at time $t+1$ are transaction and non-monetary costs associated with reconstructing social capital in a new workplace and family environment. For ease of exposition, we assume that all the costs are incurred at $t+1$. This assumption has no impact on the conclusions derived here. At any point in time a non-participant individual may search for a job at a cost γ_1 per period, assumed to be strictly lower than c_M, paid at the beginning of the next period. While the assumption that $\gamma_1 < c_M$ is not testable, it is reasonable to assume that moving to a new firm entails higher costs than getting a job while unemployed. After all, a working individual has to first quit his/her current job, and then has to

[15] Our model (and the data) does not allow us to distinguish between layoffs and resignations. Thus, a more appropriate reference to use is separation.

[16] One can also account for the costs associated with searching for an outside wage offer, but the main theoretical results remain unchanged. There are a number of equilibrium search models (in continuous time) with costly on-the-job search and outside job offers in the literature, e.g. Bontemps, Robin and Van den Berg (1998) and Postel-Vinay and Robin (2002). Nagypal (2005) shows that such models do not generally match the extent of job-to-job transitions in observed data. Consequently, she develops an alternative theoretical framework that has some similarities with our model.

incur the associated moving costs, which are at least as high as the cost of moving from unemployment to a new job. Indeed, the literature provides strong support for this assumption. For example, Hardman and Ioannides (2004) argue that the high moving costs, both in terms of out-of-pocket costs and the loss of location-specific social and human capital, explain why there are infrequent moves of individuals.

The optimal behavior:

We assume that hours of work are constant across all jobs, and concentrate only on the extensive margin of the participation process, y_{it}, where $y_{it} = 1$ if the ith individual participates in period t, and $y_{it} = 0$ otherwise. Each individual maximizes the discounted present value of the infinite lifetime (intertemporally) separable utility function given by (omitting the i and j subscripts):

$$U_t = \sum_{s=t}^{\infty} \beta^{s-t} E_t \left[u \left(C_s, l_s; x_s \right) \right], \tag{6}$$

$$\text{Subject to: } C_t = z_t + w_t^* y_t - \gamma_1 \left(1 - y_{t-1} \right) - c_M m_{t-1}, \tag{7}$$

where $u(\cdot)$ is the per-period utility from consumption C_t and leisure $l_t = 1 - y_t$, conditional on the vector of (observed and unobserved) individual characteristics, x_t. The term β denotes the discount factor. The notation $E_t(\cdot)$ denotes that the expectation is taken conditional on the information available at time t. The budget constraint assumes neither borrowing nor lending. The price of consumption in each period is normalized to 1. The quantity z_t is non-labor income, w_t^* is the individual's wage offer, $m_t = 1$ if the individual changes jobs at the end of t, and $m_t = 0$ otherwise.

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

$$V_t \left(y_{t-1}, m_{t-1}; x_t \right) = \max_{y_t, m_t} \left\{ u \left(C_t, y_t; x_t \right) + \beta E_t \left[V_{t+1} \left(y_t, m_t; x_{t+1} \right) \right] \right\}. \tag{8}$$

One can derive more detailed expressions for the value functions in each of the three possible states at t (i.e., not participating, participating without moving, participating and moving). It is then possible to derive the reservation wages (i.e., the wage threshold values) that determine the decision rules in each of the three situations.[17] For example, the reservation wages for a stayer in period $t-1$, $w_{11,t}^*$ and $w_{12,t}^*$, denote the wage levels that equate the value function of non-participation with

[17] We provide this in Appendix A.1.

those of participation, with and without interfirm mobility, respectively (at t), that is,

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

The superscript k in $V_t^k\left(y_{t-1},m_{t-1};x_t\right)$ indexes the value functions for non-participation ($k=0$), participation without mobility ($k=1$), and participation with mobility ($k=2$). For a stayer at $t-1$ the wage $w_{13,t}^*$ denotes the reservation wage that equates the value function of participation without interfirm mobility in period t with that of participation with interfirm mobility in period t, namely,[18]

$$V_t^1\left(1,0;x_t\mid w_{13,t}^*\right)=V_t^2\left(1,0;x_t\mid w_{13,t}^*\right).$$

Summary of events and decisions:

Given the existence of such reservation wages (see the proof in Appendix A.1) and their relationships (see Appendix A.2), we provide in Appendix A.3 sufficient conditions under which: (1) interfirm mobility may occur; and (2) the participation and mobility processes exhibit first-order state dependence. These sufficient conditions require that the mobility cost, c_M, should be sufficiently high relative to γ_1 and that the utility function u be weakly concave. Then, the optimal decisions are as follows:

1. A worker who is not employed in period $t-1$ has to pay the search cost, γ_1, at the beginning of period t for receiving a wage offer, w_t, at time t (from the wage offer function defined in (3)–(4)). The individual will accept the wage offer at t if it is greater or equal to a given threshold value $w_{02,t}^*$. Otherwise, the individual will decline the offer and continue to search.

2. An employed worker who changes jobs at the end of period $t-1$ pays the mobility cost, c_M, at the beginning of period t. At the beginning of period t he/she receives a wage offer from her new employer comprising a non-random and a random components. The worker learns about the random component just before the end of period t, i.e. before his/her next mobility/participation decisions. If the wage offer is lower than the threshold value $w_{22,t}^*$, then the worker becomes a non-participant at the end of period t. If the wage offer is higher than a threshold value $w_{23,t}^*$, then the worker accepts this (inside) offer and stays at the firm for at least one more period. If the wage offer falls between $w_{22,t}^*$ and $w_{23,t}^*$ ($w_{22,t}^*<w_{23,t}^*$), then the optimal strategy is to move to another firm at the end of period t.

[18] These reservation wages and all other reservation wages are also explained in greater detail in Appendix A.1.

3. An employed worker who decided to stay in the same firm at the end of period $t-1$ also receives a wage offer from his/her current employer in period t (with the same timing as just above). If the offer is lower than the threshold value $w^*_{12,t}$, then the worker becomes a non-participant at the end of period t. If the wage offer is higher than the reservation wage $w^*_{13,t}$, then the optimal strategy for this worker is to accept this (inside) offer and stay at the firm for at least one more period. If this offer falls between $w^*_{12,t}$ and $w^*_{13,t}$ ($w^*_{12,t} < w^*_{13,t}$), then the optimal strategy is to move to another firm at the end of period t (i.e., the worker becomes a mover in period t).

Given the structure of the model and the timing of the events we obtain the structure for the participation and mobility decisions. We state it formally in the following proposition.

Proposition 1: *Under the sufficient conditions given in Appendix A, the wage function specification in (3), and its error term in (4), the theoretical model implies the following structure for the participation and interfirm mobility (omitting the* ith index):

$$\begin{aligned} y_t &= \mathbf{1}\left[a_0 J^W_{jt} + x'_{yt}\beta_0 + \beta_y\, y_{t-1} + \beta_m\, y_{t-1}\, m_{t-1} + \alpha_y + u_t > 0\right] \quad \text{and} \qquad (9) \\ m_t &= \mathbf{1}\left[a_1 J^W_{jt} + x'_{mt}\lambda_0 + \lambda_m\, y_{t-1}\, m_{t-1} + \alpha_m + v_t > 0\right] \cdot \mathbf{1}(y_{t-1} = 1, y_t = 1)\ (10) \end{aligned}$$

where (i) y_t and m_t are the indicator variables indicating participation and interfirm mobility at date t, respectively; (ii) J^W_{jt} is the summary statistic for the individual's work history at date t specified in (5) (j being the last occupied job); (iii) x_{yt} and x_{mt} are sets of observable covariates; (iv) α_y and α_m are person-specific effects and (v) u_t and v_t are contemporaneous white noises.

Proof: A complete proof is provided in Appendix A.

The presence of the J^W_{jt} in the wage function is crucial for determining interfirm mobility. A large enough positive change in J^W_{jt} would induce a worker to move in order to "boost" his/her career even though the worker needs to pay the mobility cost c_M . However, some workers with a very negative idiosyncratic shock ξ_{ijt} may not be able to cover this mobility cost. In this case, because the search cost γ_1 is lower than c_M, the worker may choose to be non-employed, search for a new job and, hopefully, move to a new firm in the next period. The presence of experience, seniority, and the J^W_{jt} function in the wage function introduces elements of non-stationarity. Clearly, this model generates endogenous mobility and, therefore, stands in stark contrast with models of on-the-job search in which job offers or job destructions take place at some exogenously given Poisson rate.

Equations (9) and (10) are the reduced-form of our structural model and are therefore robust to some changes in the maintained hypotheses of the underlying model, including the exact timing of the events described above. To see this, compare our model with the standard on-the-job search models, such as Burdett (1978) and Postel-Vinay and Robin (2002). In these types of stationary models, that are also often in continuous time, a worker observes the complete wage offer as soon as the offer is posted. Nevertheless, their solutions imply that mobility is decided based on comparing different reservation wages, and thus their reduced-form counterparts are similar to ours. The model introduced here departs from the traditional on-the-job search models in that it incorporates some more realistic features of wage contracts, namely contracts that have both deterministic and random components. It is important to note though that even if ξ_{ijt} was known at the beginning of the period, as in these models, the reduced-form equations would barely change. However, the solution to the optimal decisions would be a lot more complex, involving differences between the random terms ξ_{ijt} and $\xi_{ij't}$, without changing the main results in any meaningful way.

Finally, technical comments are in order. First, as is discussed in detail below we impose some exclusion restrictions for identification, i.e., some of the covariates in x_{yt} and x_{mt} are excluded from x_{wt}, the regressor vector of the wage function in (3). Second, note that the first-order state dependence on lagged participation and lagged mobility are captured by the parameters β_y, β_m in (9) and λ_m in (10). Finally, note that the person-specific effects, (i.e., α_y, α_m from (9) and (10), respectively, and α_w from (13) below), are correlated by construction (see Appendix A for details).

4 Econometric Specification and Estimation

4.1 Econometric Specification

Before proceeding with the econometric specification, it is convenient to rewrite the model as follows:

The *participation (employment)* equation, at any date $t > 1$, is given by[19]

$$\begin{aligned} y_{it} &= \mathbf{1}(y_{it}^* \geq 0), \\ y_{it}^* &= a_0 J_{ijt}^W + x_{yit}' \beta_0 + \beta_y y_{i,t-1} + \beta_m m_{i,t-1} + \alpha_{yi} + u_{it}. \end{aligned} \tag{11}$$

The *interfirm mobility* equation, at any date $t > 1$, is given by[20]

$$\begin{aligned} m_{it} &= \mathbf{1}(m_{it}^* \geq 0) \cdot \mathbf{1}(y_{i,t-1} = 1, y_{it} = 1), \\ m_{it}^* &= a_1 J_{ijt}^W + x_{mit}' \lambda_0 + \lambda_m m_{i,t-1} + \alpha_{mi} + v_{it}. \end{aligned} \tag{12}$$

The *wage* equation, for any date t, is given by:

$$\begin{aligned} & w_{ijt} = w_{ijt}^* \cdot \mathbf{1}(y_{it} = 1), \quad \text{where} \\ & w_{ijt}^* = x_{wit}' \delta_0 + J_{ijt}^W + \alpha_{wi} + \xi_{ijt}. \end{aligned} \tag{13}$$

The terms y_{it}^* and m_{it}^* denote the two latent variables affecting the employment and mobility decisions, respectively. Note that an obvious implication of the definition of a move is that a worker cannot move at date t unless he/she participated at both $t-1$ and t.

Controlling for initial condition is crucial since not all individuals in the sample are at the same stage in their life cycle. Following Heckman (1981), we approximate the initial condition using a probit specification given by

$$\begin{aligned} y_{i1} &= 1(y_{i1}^* \geq 0), \quad \text{where } y_{i1}^* = a x_{yi1} + \delta_y \alpha_{yi} + u_{i1}, \text{ and} \\ m_{i1} &= 1(m_{i1}^* \geq 0) \cdot 1(y_{i1} = 1), \quad \text{where } m_{i1}^* = b x_{mi1} + \delta_m \alpha_{mi} + v_{i1}, \end{aligned} \tag{14}$$

and where α_{yi} and α_{mi} are the individual specific effects in the participation and mobility equations defined in (11) and (12), respectively. That is, δ_y and δ_m are allowed to differ from 1.

We assume that the individual specific effects are stochastically independent of the idiosyncratic shocks, that is $\alpha_i \perp (u_{it}, v_{it}, \xi_{ijt})$. Furthermore, α_i follows the dis-

[19] Note that the labor market experience is simply the sum of the individual sequence of y_{it}. As is common in the literature, we make no distinction in this specification between unemployment and non-participation in the labor force. In addition, as mentioned earlier, we equate participation and being employed.

[20] Mobility takes place at the beginning of period t. Seniority is the sum of the individual sequence of m_{it} with his/her current employer.

tribution given by

$$\alpha_i \sim N(f(x_{i1},...,x_{iT}), \Gamma_i), \quad \text{where} \tag{15}$$
$$\Gamma_i = D_i \Delta_\rho D_i, \quad D_i = \text{diag}(\sigma_{yi}, \sigma_{mi}, \sigma_{wi}), \quad \text{and}$$
$$\{\Delta_\rho\}_{j,l} = \rho_{\alpha_{ji}\alpha_{li}}, \quad \text{for } j,l = y,m,w.$$

where the function $f(x_{i1},...,x_{iT})$ depends, in principle, on all of the exogenous variables in all of the periods. We also allow for σ_{ji} to be heteroskedastic, i.e., to depend on x_{yit}, x_{mit} and x_{wit}, respectively. That is,

$$\sigma_{li}^2 = \exp(h_l(x_{i1},...,x_{iT})), \quad \text{for } l = y,m,w, \tag{16}$$

where each $h_l(\cdot)$ is a real valued function. Since the ultimate goal is to control for the possible existence of heterogeneity in a parsimonious way, we base our estimation only on a constant term and the first three principle components of the sample averages of the regressors' vector, i.e., $\overline{x}_{li} = (\sum_{t=1}^{T} x_{lit})/T$. That is, we approximated $h_l(x_{i1},...,x_{iT})$ by $\widehat{h}_l(x_{ji1},...,x_{jiT}) = pc_i'\widehat{\gamma}_l$, where pc_i is the vector containing the first three principal components. This significantly reduces the computational burden.[21]

Finally, the idiosyncratic error components from (11), (12) and (13), i.e., $\tau_{it} = (u_{it}, v_{it}, \xi_{ijt})'$, are assumed to be contemporaneously correlated white noises, with

$$\tau_{it} \sim N(0, \Sigma), \quad \text{where} \tag{17}$$

$$\Sigma = \begin{pmatrix} 1 & \rho_{uv} & \rho_{u\xi}\sigma_\xi \\ \rho_{uv} & 1 & \rho_{v\xi}\sigma_\xi \\ \rho_{u\xi}\sigma_\xi & \rho_{v\xi}\sigma_\xi & \sigma_\xi^2 \end{pmatrix}, \tag{18}$$

and the variances of u_{it} and v_{it} are normalized to $\sigma_u^2 = \sigma_v^2 = 1$ for identification purposes.

In the analysis reported below, we adopt a Bayesian approach whereby we obtain the conditional posterior distribution of the parameters, conditional on the data, using Markov Chain Monte Carlo (MCMC) methods as explained below.[22]

[21] The first three principle components account for over 98% of the total variance of $\overline{x}_{ji}$, so that there is almost no loss of information by doing so. Also, note that $\widehat{\gamma}_j$, $j = y,m,w$, have no causal interpretation. We use this structure to control merely for possible heteroskedasticity of the covariance matrix.

[22] One can also use an alternative (frequentist) approach such as Simulated Maximum Likelihood (SML) (e.g. Gouriéroux and Monfort (1996), McFadden (1989) and Pakes and Pollard (1989)).

4.2 Computation of The Posterior Distribution

Since it is analytically intractable to compute the exact posterior distribution of the model's parameters, conditional on the observed data, our goal here is to summarize the joint posterior distribution of the parameters of the model using a Markov Chain Monte Carlo (MCMC) algorithm.[23]

Let the prior density of the model's parameters be denoted by $\pi(\theta)$, where θ contains all of the parameters of the model 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 the latent variables, $z^*_{it} = (y^*_{it}, m^*_{it}, w^*_{ijt})$, where y^*_{it}, m^*_{it} and w^*_{ijt} are defined in (11), (12) and (13), respectively. With this addition, it is easier to implement the Gibbs sampler, which iterates through the set of the conditional distributions of z^* (conditional on θ) and θ (conditional on z^*).[24] For the simulation results reported below, we use 50,000 MCMC repetitions after discarding the first 5,000. We use Gelfand and Smith (1990) to assess the performance of the algorithm.

For the prior distributions of the model's parameters we use conjugate priors. In each of the cases, we used proper, but very diffuse priors, to ensure that the posterior distributions are indeed proper distributions. An extensive sensitivity analysis carried out shows that the choice of the particular prior distribution has hardly any effect on the posterior distribution. In particular, we estimate the model centering the key parameters around the results obtained by Topel (1991) and AS. The resulting posterior distributions obtained under both scenarios were virtually identical.

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) of college graduates. The estimation took several months. Nevertheless, the point estimates for all of the model's parameters obtained by the SML were virtually the same as the mean of the joint posterior distribution of the model's parameter, inducing identical estimates for the main parameters of interest, namely the returns to seniority and experience. The MCMC method can be viewed in this sense as a mechanical device for obtaining SML estimates.

[23] Appendix B provides a detailed description of the implementation of the MCMC simulation.

[24] A presentation of the theory and practice of Gibbs sampling and MCMC methods may be found in Robert and Casella (1999) and Chib (2001). In econometrics, recent applications to panel data include Geweke and Keane (2000), Chib and Hamilton (2002) and Fougère and Kamionka (2003).

4.3 Exclusion Restrictions and Identification

It is worthwhile to note that models such as the one specified above depend on the functional form restrictions (e.g. Chamberlain (1984)). Thus, the robustness of the results depends on the particular situation. Hyslop (1999) discusses the potential problems and identifying restrictions one needs to impose in a similar situation. Here, the identification of state dependence depends on having time varying regressors. As Hyslop (1999) correctly notes, any transitory changes in the relevant regressors (i.e., x_{yit} in (11) and x_{mit} in (12)) will have persistent effects on the probabilities through their effect on y_{it} and m_{it}, respectively. If the only source of dynamics in the model stems from the first-order state dependence, then the participation probability, i.e., $\Pr\left(y_{it} = 1|y_{i,t-1}m_{i,t-1}, J^W_{ijt}, x_{yit}, \alpha_{yi}\right)$, does not depend on the lagged x's , and similarly for the mobility probability. However, any form of serial correlation in the error terms will induce dependence of these probabilities on lagged x's. Our specification assumes that the idiosyncratic term, τ_{it}, in (17) is serially uncorrelated,. but some correlation over time does exist because of the presence of the individual specific random component vector α_i given in (15).[25]

As we discuss in the results section below we impose some exclusion restrictions; that is, some of variables that appear in these equations do not appear in the wage function. These exclusion restrictions provide an additional source of identification beyond that implied by the particular functional form imposed. Our framework allows us to directly test these identifying exogeneity restrictions.

Consider the linear regressions $\alpha_{ji} = z_i'\overline{\delta}_j^A + u_{\alpha_j i}$, for $j = y, m, w$, where $z_i' = (z_{i1}', ..., z_{iT}')$, and z_{it} is a vector containing the potentially exogenous variables for individual i at time t. If z_{it} is indeed exogenous with respect to both the participation and mobility, then it should be uncorrelated with the unobserved individuals component, α_i. Hence, the null hypothesis of exogeneity is first specified as

$$H_0^A\text{: } \overline{\delta}_j^A = 0, \quad \text{for } j = y, m, w. \tag{19}$$

To test this hypothesis we compute the mean of the posterior distribution for the individuals' specific parameters, say $\overline{\alpha}_1, ..., \overline{\alpha}_n$, and regress each component of $\overline{\alpha}_i$, (i.e., $\overline{\alpha}_{yi}$, $\overline{\alpha}_{mi}$, and $\overline{\alpha}_{wi}$) on z_i to obtain the estimates for $\overline{\delta}_j^A$, say $\widehat{\overline{\delta}}_j^A$, for $j = y, m, w$.

[25] An additional problem with discrete choice models is that in order to observe changes in the binary outcome variables, $y_{i,t}$ and $m_{i,t}$, from one period to another, there need to be sufficiently large changes in the underlying latent variables, $y_{i,t}^*$ and $m_{i,t}^*$.

We then conduct Wald tests for the joint significance of the coefficients in $\widehat{\overline{\delta}}_j^A$.

In addition, we also consider the regressions $\alpha_{ji} = x'_{-z_i}\overline{\eta} + z'_i\overline{\delta}_j^B + u_{\alpha_j i}$, for $j = y, m, w$,where x_{-z_i} is the vector of all variables in all periods for individual i, excluding the variables in z_i. Again, we regress each of the components of $\overline{\alpha}_i$ on the regressor vectors z_i and x_{-z_i} and test for

$$H_0^B\colon \overline{\delta}_j^B = 0, \quad \text{for } j = y, m, w. \tag{20}$$

5 The Data

Following Topel (1991) and AW we used the Panel Study of Income Dynamics (PSID) survey. 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) the 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 these households every year, whether or not they were living in the same dwelling or with the same people. While there was some attrition, the PSID has had significant success in its re-contact efforts.[26]

The data extract used in this study comes from the 18 waves of the PSID from 1975 to 1992. The sample is restricted to all heads of households interviewed for at least three years during the sample period and who are between the ages of 18 and 65. We use all of the individuals, including self-employed. We carry out some sensitivity analyses and exclude the self-employed from our sample, but the results remain virtually unchanged. Also, since most of the heads of the households in the sample are men, we conduct an additional analysis dropping the women household heads from the estimation. Again, no visible differences have been detected in the results obtained. We do exclude from the analysis all observations from the poverty sub-sample of the PSID.

[26] For a more detailed description of the PSID, see Hill (1992).

There are some difficulties with some of the key variables of our analysis and especially with the tenure variable. 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. There is also a large number of inconsistencies in the data causing tremendous spurious year-to-year variance in reported tenure on a given job. For example, between two years of a single job, tenure sometimes fell (or rose) by more than one year. There are many years with missing tenure followed by years in which a respondent reported more than 20 years of seniority.

Since these errors can basically determine the outcome of the analysis, we reconstruct the tenure and experience variables using the exact procedure suggested by Topel (1991). Specifically, for jobs that begin in the panel, tenure was started at zero and incremented by one for each additional year in which the person worked for the same employer. For jobs that began before the first year that a person is in the sample, a different procedure was followed. In such cases, the starting tenure is inferred from the longest sequence of consistent observations. If no such sequence exists, then we begin from the maximum tenure on the job, provided that the maximum is less than the age of the person minus his/her education minus 6. If this is not the case, then we begin from the second largest value of recorded tenure. Once the starting point has been determined, tenure is incremented by one for each additional year with the same employer. Similar principles have been also employed for determining initial experience. Once the starting point has been computed, experience is incremented by one for each year in which the person works. Using this procedure, we have managed to eliminate all inconsistencies.[27]

Summary statistics of some of the key variables in the extract used here are reported in Table 1. Note that due to the nature of the PSID data collection strategy, the average age does not increase much over time. The average education rises by more than half a year between 1975 and 1992 because the new individuals entering the sample tend to have a higher level of education than those interviewed in the past. Also, experience and seniority increase over the sample from 20.8 and 5.2 years in 1975 to 24.2 and 7.3 years in 1992, respectively.

The experience variable reported in our study seems to be very much in line with

[27]We have also taken a few other additional cautionary measures. For example, we check that: (i) the reported unemployment matched the change in the seniority level; and (ii) no peculiar changes in the reported state of residence and region of residence exist, etc. The programs are available from the corresponding author upon request.

the experience levels reported in other studies. However, there is a substantial difference between the tenure variable used here and that used by Topel and especially AW (see AW, Table 1). The main reason for this discrepancy is the fact that we restrict the sample to individuals for whom we have consistent data for three consecutive years. Individuals with lower seniority are also those who tend to leave the PSID sample more frequently. Nevertheless, we provide strong evidence below that the differences in the results are not due to the use of different data extracts, but rather due to the differences in the modeling and estimation approaches.

The mobility variable indicates that, in each of the sample years, between 6.5% and 14.1% of the individuals change jobs. The mobility of 23.4% in the first year of the sample is largely due to measurement error. This is precisely the reason why it is necessary to control for initial conditions as is explained above.

Consistent with other data sources, the average real wages increased slightly over the sample years. This increase is mainly caused by individuals entering the sample whose wages increased over time in real terms, while those who left the sample had wages that decreased somewhat over the sample years. More importantly, as has been documented in the literature, the wage dispersion has increased across years. Also, consistent with other data sources, the participation rate of the individuals in the sample (almost exclusively men) decreases steadily over the sample years from about 94% in 1975 to 86% in 1992. The PSID over-sampled non-white individuals. However, since the results change very little when we use a representative sample of non-white individuals, we have included all individuals satisfying the conditions described above.

The data reflect the general finding in the literature about the changing structure of the typical American family. The fraction of married individuals declined significantly over the sample period, as did the total number of children in the family. Consequently, the fraction of individuals having young children also declined over the sample years.

Note that there are also substantial changes in the distribution of cohorts during the sample period. While the fraction of individuals in the youngest cohort increased steadily, the fraction of individuals in the older cohorts decreased over the sample period. We also observe some changes in the age and race structure. All these changes are largely due to the strong geographic mobility of the younger workers that the PSID has lost contact with.

6 The Results

Below, we present a set of results covering various aspects of the model. After discussing the specifications we highlight the sources for identification of the model's parameters. After a brief discussion of the results for the participation and mobility we concentrate our discussion on the returns to human capital, namely education, experience, and the returns to seniority, the key element of this study. We then present an in depth discussion of the implications of the results, as well as a comparison to previous literature.

6.1 Specifications and Sources of Identification

The estimation is carried out for three separate education groups. The first group includes all of the individuals with less than 12 years of education, referred to as *high school (HS) dropouts*. The second group consists of those who are high school graduates, but have not completed a four-year college degree. We refer to this group as *high school graduates*. The third group, the *college graduates*, consists of those individuals with at least 16 years of education and who have earned a college degree.

The participation equation includes the following variables: a constant, education, quartic in lagged labor market experience, a set of three regional dummy variables, a dummy variable for residence in an SMSA, family unearned income, dummy variables for being African American and Hispanic, county of residence unemployment rate, a set of variables providing information about the children in the family, a dummy variable for being married, a set of four dummy variables for the cohort of birth and a full set of year indicators. We also include the J^W function, as explained in Section 3. The mobility equation includes the same variables as the participation equation, and, in addition, a quartic polynomial in lagged seniority on the current job and a set of nine industry indicator variables.

The dependent variable in the wage equation is the log of the deflated annual wage. The set of regressors in the wage function is the same as in the mobility equation, with two key differences. First, the quartic polynomials in the wage equation are in current (rather than lagged) experience and seniority. Second, as discussed in Section 4.3, we impose a number of exclusion restrictions on the wage function. The excluded variables are: family unearned income, the number of children in the family, the number of children between one and two, the number of children between three and five, and the marital status dummy variable, all of which appear in both the participation and mobility equations.[28]

It is worthwhile to note that the identification of the model's parameters comes from both the cross-sectional and the time-series dimensions. If there were no individual specific effects the cross-sectional dimension would provide enough variation for us to estimate all the model's parameters, although one might rightly question the quality of the estimate for the return to experience and, especially, the return to tenure. The time-series dimension helps identify these last two key parameters. This comes from the fact that there are many individuals who changed jobs during the sample period, and did so at different points of their life-cycle. The time-series dimension also allows us to control for the individual specific effects. This is why we restrict the data extract to include only individuals who were observed for at least three consecutive years. Since experience and seniority are fully endogenized, we need not impose any further restrictions on the data extract (e.g. restricting attention to only exogenously displaced workers) as is done in Dustmann and Meghir (2005) or Topel (1991).

The main results are provided in Tables 2 through 9. It turns out that the marginal distributions of all the model's parameters, as well as all the marginal and cumulative returns to experience and seniority, are very close to those of normal random variables. Thus, it is sufficient to report the mean and standard deviation of the marginal posterior distribution for each of the parameters. For brevity we present in the tables only the results regarding the key variables in the analysis.

6.2 Participation and Mobility

The results for the participation and mobility decisions are presented in Tables 2 and 3, respectively. As explained above, in both the participation and mobility equations we account for duration dependence and unobserved heterogeneity. As is implied by our model, the participation equation includes lagged participation and lagged mobility, whereas the mobility equation only includes lagged mobility. In principle, one also needs to include the J^W function from (5) in both the employment and mobility equations; however, doing so does not change the results in any meaningful way. First, all of the coefficients associated with the elements of the J^W function are statistically

[28] Hyslop (1999), Mroz (1987), among others, provided very strong evidence that these variables are indeed exogenous, at least with respect to the participation decision.

and economically insignificant. Moreover, the inclusion of the J^W function has no effect on any of the other parameter estimates. Hence, we exclude the J^W function from both the participation and mobility equations.

Participation (Table 2):

The estimates of the parameters from the participation equation are in line with those previously obtained in the literature and consistent with the classic human capital theory. Education is an important factor in the participation decision for the more educated individuals, but seem to play no role for the high school dropouts. Similarly, experience is more important for the more educated individuals as well.

Note that lagged employment has a positive and highly significant effect for all three education groups. Individuals who participate in one period are also likely to participate in the following period, regardless of their education level. The positive and significant coefficient on lagged mobility indicate that worker who changes jobs are more likely to participate in the labor force in the following period, and, as predicted by human capital theory, this effect is stronger for the least educated individuals.

The results for the family variables are, in general, consistent with economic theory, especially regarding the marital status and the children variables. The race effect is particularly striking. African-Americans are less likely to participate, relative to their white counterparts, while for Hispanics this appear to be the case only for college graduates.

Mobility (Table 3):

One important finding is that the more senior workers are a lot less likely to move. This is largely because once a person leaves a job the accumulated returns to seniority (representing firm-specific human capital) are lost . Moreover, a move in one period significantly reduces the probability of a move in the subsequent period, and more so for the more educated individuals. This may simply mean that more educated individuals are more likely to be in jobs that are better matches, or, alternatively it may be that the signalling effect of repeated frequent moves are more severe for the more educated individuals.

Note that experience does not seem to have as strong an effect as that of seniority. Nevertheless, consistent with the general findings in the literature (e.g. Farber (1999)) the results suggest that more experienced workers tend to move less. In turn, the results suggest that one needs to account for the endogeneity of experience, not only seniority, when estimating a wage function. Interestingly, the only family variable that seems to affect mobility is family unearned income. As it turns out, the marginal effect of unearned family income on the probability of a move, when evaluated at the group mean, is roughly the same for all groups. Finally, race does not seem to play a major role in mobility decisions, except maybe for the most highly educated African-American.

6.3 The Wage Function and the Returns to Human Capital

The Returns to Education:

The results for the returns to education are reported in Table 4. The (within-education groups) returns to education are broadly consistent with the human capital theory of Becker (1964). Declining marginal returns to education lead to lower marginal returns to education for the college graduates than for the high school graduates. However, the marginal returns for the high school graduates, 4.8%, are larger than for the high school dropouts, only 2.5%. These results appear to be consistent with, although somewhat lower than, recent findings in the literature (e.g. Card (2001)).[29] Once one controls for the selection effects, due to the participation and mobility decisions, the estimated returns to education are reduced somewhat. This also seems to provide a plausible explanation for the fact that the returns to education are larger for the high school graduates and college graduates than for high school dropouts. We return to this issue below when discussing the estimates of the stochastic terms of the model.

The Returns to Experience:

In Table 5 we provide estimates for the returns to experience, which are based on the parameter estimates reported in Table 4. The results are close, yet generally larger, than those previously reported in the literature. Topel reports an estimate (see Tables 1 and 2 of Topel (1991)) of the cumulative returns to experience at 10 years of experience of .354. The estimates of AS at this level of experience range between .372 and .442, while those of AW range between .310 and .374. Our estimates are .362 for the high school dropouts, .402 for the high school graduates and .661 for the college graduates. Our results indicate that there are substantial differences across the various education groups at all levels of experience. For example, five years after

[29] There is substantial variation in years of education within each group of education as we defined them. This variation allowed us to separately identify the within-group returns to education for each of the three groups.

entering the labor market, the cumulative returns for college graduates are almost twice as large as those for the other two groups.

To examine the sensitivity of the results we consider several alternative model's specifications.[30] First we estimate the model using a quadratic, rather than the quartic, polynomial in experience and mobility. The estimated cumulative returns to experience for this model are substantially lower than those reported in Table 5. For example, the cumulative returns at 10 years of experience are .246, .253 and .446 for the three education groups, respectively.

Next we examine the effect of controlling for past labor market history. To do that we re-estimate the two models discussed above with the J^W function excluded. Omitting the J^W function induces a large increase in the estimated cumulative returns to experience for all education groups. Given that the estimates of the parameters associated with the J^W function (see rows 11–23 of Table 4) are jointly highly significant, highlights the importance of controlling for past job market history. As the results clearly indicate, failing to account for past changes in employment status and job mobility induces a severe upward bias on the estimated returns to experience. Moreover, in our study we control for the possible correlations between the various individual-specific effects in the estimated equation. It turns out that failure to control for these correlations induces downward biased on the estimated returns to experience. After controlling for these two effects, the estimate we obtain are somewhat larger than those previously found by Topel (1991), AS, and AW.

The Returns to Seniority:

The returns to seniority are reported in Table 6 and are based on the parameter estimates reported in Table 4. The estimated returns to seniority, at all levels of seniority and for all education groups, are significantly higher than those previously reported in the literature, including those reported by Topel (1991). In fact, the support of the marginal posterior distributions for the returns to seniority is entirely in the positive segment of the real line for all groups. Also notable is the fact that the returns also increase dramatically with the level of seniority. For example, while at two years of seniority the cumulative returns are around 13% for the three education groups , they rise to about 50% at ten years of seniority. The results for the quadratic model, not reported here, are qualitatively similar to those reported here for the quartic model. One noticeable difference is that the cumulative return to seniority

[30] For brevity, we do not report these results in a form of a table, but they are available from the authors upon request.

for the least educated group flattens at ten years of seniority.

When the J^W function is excluded from the wage equation the resulting estimates of the returns to seniority, for both the quartic and the quadratic models, are significantly lower. Again, this clearly illustrates the need for one to account for past labor market history, as we do via the J^W function. When one does not include the J^W function, the average cumulative returns for the three education groups are, as one might expect, remarkably close to those obtained by Topel.

One key reason for the difference between our results and those of Topel (1991) is that Topel assumes that level of experience at a new job is exogenous. The literature (e.g. Farber (1999)) and the results presented here provide strong evidence that this far from being true. In fact, individuals with low experience tend to change jobs more frequently. Indeed, when we treat experience as exogenous and eliminate the participation equation from the estimation, the resulting returns to experience are significantly reduced and are comparable to those obtained by Topel. The returns to seniority obtained remain still significantly higher than those obtained by Topel. Moreover, when we omit the J^W function from the wage equation, the resulting estimates for the returns to seniority are reduced even further. Nevertheless, they are still larger than those of Topel, largely because he does not account for the possible unobserved heterogeneity in the way we do.

The results for cumulative returns to experience and seniority indicate that the sum of the returns from experience and seniority are quite similar across the various specifications (i.e., quartic versus quadratic, and with and without the J^W function), but larger than previously obtained in the literature. The correlation coefficients between experience and seniority in our data extract are .22, .36, and .46, for the high school dropouts, high school graduates and college graduates, respectively. Thus, it is not surprising to find that larger returns to one component are associated with lower returns to the other. Specifically, our results suggest that accounting for endogeneity of experience is crucial in correctly assessing the effect of tenure, and more so for the more educated individuals.

To contrast our results with those previously found in the literature, we provide in Table 7 a comparison with AW results. In Panel A of the table we present the results of AW (their Table 2). The results in that panel are based on AW's methodology applied to the *Topel replication sample* for the period 1968-1983. In Panel B we present the results for the high school dropouts and the second for college graduates, based on AW's methodology and our model, using *our* sample for the period 1975-1992.

The OLS estimates for the returns to seniority for the two alternative samples are remarkably close, as are the IV estimates for the two samples. The conclusion is thus very clear: the apparent differences in results in our study stem from differences in methodology, not in the data extract. In fact, the most important elements that explain these differences are: (a) endogenizing the participation decision; (b) the explicit control for individual-specific effects in the three equations; and (c) the introduction of the J^W function that captures past employment spells.

6.4 Testing for the Exclusion Restrictions

The results of the testing procedure described in Section 4.3 are provided in Table 8 for two alternatives, namely the quartic and quadratic models (both including the J^W function). The results in the columns that are marked H_0^A and H_0^B refer to the test described in (19) and (20), respectively.

Note that in all cases but one we are unable to reject the null hypothesis that the set of coefficients corresponding to the excluded variables are zero. Even in the single case where we do reject the null hypothesis, the p-value is quite large, .087. These results lend strong support to the validity of the identifying assumptions described in Section 4.1. There are some differences, as one might expect, between the results obtained for the individual specific effect for the wage equation, α_w, and the participation and mobility equations, α_y and α_m. Specifically, it is easier to reject the null hypothesis for α_w than for either α_y or α_m.

6.5 The J^W Function

As explained above, the J^W function is an individual-job specific function that parsimoniously summarizes the changes in one's wage that correspond to a particular career path, providing a measure of the individual's opportunity wage at any point in his/her career. The parameter estimates of the J^W function, provided in rows 11 through 23 of Table 4, clearly indicate the importance of controlling for an individual's career path. While some of the individual parameter estimates are statistically insignificant, the null hypothesis that all coefficients are zeros is rejected at any reasonable significance level for all of the three education groups.

Some important implications emerge from these results. First, the timing of a move in a worker's career matters. While, at the time of a move, a person loses the accumulated returns to seniority, he/she also receives a premium. This premium corresponds to the particular levels of experience and seniority at the previous job, and includes: (a) a fixed increase determine by the length of spell on the previous job (ϕ_{j0}, $j = 1, ..., 4$); (b) a variable change proportional to seniority at the time of the job change (ϕ_j^s, $j = 2, 3, 4$); and (c) variable, *career effect*, change proportional to experience at the time of the job change (ϕ_j^e, $j = 1, ..., 4$).

Consider, for example, a high school dropout with 10 years of seniority and 10 years of experience moving to a new job. The results from Table 6 indicate that he/she would lose .507 in log-wage. However, from Table 4 we see that he/she would gain .215 (line 15) plus $.011 \times 10 = .11$ (line 18) plus $.001 \times 10 = .01$ (line 22), for an overall substantial wage loss of about 17%. This loss captures the so-called *displaced worker effect* analyzed in the literature. The effect of *unemployment spells* during one's career is captured through the (actual) labor market experience.

The pattern for the more highly educated individual is quite different. Similar calculation show that the net effect for a college-educated worker changing jobs after 10 years of seniority and 10 years of experience is a 8% reduction in his/her wage. In contrast, the net gain for a college-educated worker at the beginning of his/her career changing jobs having 2 years of seniority and 2 years of experience is almost 12%.

Note it is possible to estimate the J^W function because we explicitly endogenize both experience and seniority. In turn, the introduction of J^W allows us to control for the indirect effects of experience and seniority that stem from the individual's specific work history. Consistent with the results of Topel and Ward (1992), we find that this non-linear interaction between seniority and experience plays a crucial role in explaining wage growth, particularly at the beginning of one's career, as is illustrated below.

6.6 Estimates of the Stochastic Elements

In Table 9, we provide estimates of the key parameters associated with the stochastic elements of the model, that is, the parameters of Σ in (18) and the correlation parameters of Δ_ρ in (15). Since Σ and Δ_ρ are left largely unrestricted, the results allow us to empirically assess the need to control for the participation and mobility decisions when estimating the wage equation.

Note that correlations between the various person-specific components, i.e., the elements of Δ_ρ, are highly significant and very large in magnitude. This is especially true for the high school graduates and college graduates. Since higher participation rates imply faster accumulation of labor market experience, the estimates of $\rho_{\alpha_y \alpha_w}$ imply that, all other things being equal, high-wage workers (i.e., those with a large α_w) tend also to be more experienced workers. Hence, omission of α_w from the wage equation would induce an upward bias in the estimated returns to experience. In contrast, the positive correlation between α_y and α_m implies that failing to control for the participation decision would induce a downward bias on the estimates of returns to experience. While the net effect of these two conflicting effects is *a priori* not clear, we find that they almost cancel each other, leading to estimated returns to experience which are only somewhat larger than those previously obtained in the literature (e.g. Topel (1991), AS and AW).

The correlations between the participation and mobility components, i.e., $\rho_{\alpha_y \alpha_m}$ and between the mobility and wage components, i.e., $\rho_{\alpha_m \alpha_w}$, are very large and negative for the high school graduates and college graduates. That is, high-mobility workers also tend to be low-wage and low-participation workers. Moreover, since high-wage workers have higher seniority than low-wage workers, omitting α_w from the wage equation would induce an upward bias in the estimated returns to seniority. On the other hand, because $\rho_{\alpha_m \alpha_w} < 0$, failing to control for the mobility decision would induce a downward bias in the returns to seniority. Empirically we find that the net effect—once we control for the mobility and participation decisions, the unobserved heterogeneity, and introduce the J^W function—is huge, leading to estimated returns to seniority which are larger than those estimated by Topel (1991).

Finally, examining the estimated Σ, we see that $\rho_{u\xi}$, the correlation between the idiosyncratic errors in the participation and wage equations, is negative and highly significant for all of the education groups. In contrast, and mostly for the college-educated workers, positive shocks to mobility are associated with positive shocks to wages.

6.7 Implications for Long-run Wage Growth

The complexity of the model and the implied interactions between the various components makes it difficult to evaluate the overall effects of participation and mobility episodes on wage growth. In order to more closely examine this issue, we follow the wage growth for two types of hypothetical workers with two distinct career paths. One group consists of individuals working through the entire sample period in one job, while the other includes those working at one firm for the first four years, then changing to a new job in which they stay for the remainder of the sample period. For both groups we focus our attention on new entrants (with 5 years of experience and 2 years of tenure) and experienced workers (with 15 years of experience and 6 years of tenure.) The results for the high school graduates and college graduates are presented in Figures 1 and 2, respectively.[31] In all of the figures we decompose the wage growth that stems from returns to human capital into the part due to rising experience and the part due to rising seniority. In doing so we also account for all the terms associated with the J^W function, which come into play when an individual changes jobs.

First, consistent with the theory of human capital, wage growth early in one's career is large, and more significant for the more educated individuals. This effect is further reinforced when one takes into account the effect implied by the J^W function. For example, wage growth over the 18-year period for new entrants with a high school degree (Figure 1a) is .74 log points, whereas for new entrants with a college degree it is .94 log points (Figure 2a). This is in addition to the fact that college graduates also start a a higher level. Furthermore, wage growth that stems from increased seniority is larger than wage growth that stems from increased experience. For the college graduates with no job changes (Figure 2a and 2b) the increases due to seniority and experience are .59 and .35 log points, respectively, for the new entrants, and .50 and $-.08$, for the experienced workers. The differences for the high school graduates with no job changes (Figures 1a and 1b) are .59 and .16 log points, respectively, for the new entrants, while for the experienced workers, the respective increases are .48 and $-.04$ log points. The results indicate that the decreases in wages associated with the loss of firm-specific skills, as represented by seniority, are quite substantial for all groups.

These figures illustrate the magnitude of two effects that come into play when a worker changes jobs. First, there are direct losses due to lost returns to accumulated seniority. Second, there are discrete jumps, as implied by the J^W function, in wages at the entry level in a new job. Figures 1 and 2 show that labor market histories have differential effects across the various groups. For example, for high school graduates with little labor market experience, the loss of accumulated seniority is compensated almost entirely by a discrete increase in the base wage (Figure 1c). For new entrants

[31] The results for the high school dropouts are very similar to those of the high school graduates, and are therefore omitted.

with a college degree, the loss of accumulated seniority is more than compensated by the change in the entry wage level (Figure 2c).

In contrast, workers with substantial labor market experience after a relatively long spell within a firm incur substantial wage loss when changing jobs as is observed in Figure 1d. The overall loss is approximately .16 log points for the high school graduate (Figure 1d), while the loss for experienced college graduates a long tenure spell is somewhat smaller (Figure 2d).

7 Summary and Conclusions

Various theories in economics predict that wage compensation should rise with labor market experience and seniority (tenure) in the firm. While the empirical literature finds that experience is indeed important, there is much disagreement about the role of tenure. In this study we adopt recent methodological advancements that allow us to take a new, and more insightful, look at this on going debate. We adopt the same general Mincer's wage specification as in previous literature. However, in contrast to virtually all other studies, we explicitly model the participation (employment) and job mobility decisions, along with the wage equation. That is, we not only take into account the endogeneity of tenure, but we also take into account the possible endogeneity that stems from the (endogenous) employment decisions.

The estimated equations are the reduced-form equations derived from a theoretical model. This model builds on Hyslop (1999) and extends his model to include key features from the job search theory. This approach allows us to re-examine the magnitude of the returns to seniority in the United States, as well as providing explanations for the apparent differences between our findings and those previously found in the literature. Specifically, we control for unobserved heterogeneity by introducing person specific random correlated effects in the three estimated equations. We also control for the individuals' job specific labor market histories. This modeling strategy allows us to trace the implied biases that stem from estimating the wage equation by itself, or by estimating the wage function only controlling for the endogeneity of tenure.

The results unequivocally support Topel's (1991) findings and conjectures. We find that the returns to seniority are large and highly significant. In fact, the support of the posterior distribution of the returns to seniority is entirely in the positive segment of the real line for all three education groups. Various sensitivity analyses demonstrate that the distinct results obtained in our study are not an artifact of the data extract used, but rather stem from our modeling approach. Particularly, our study clearly demonstrates the importance of jointly estimating the wage equation along with the employment and mobility decisions.

We also find that in addition to the direct effects that experience and seniority have on one's wages, there are indirect effects that depend on the individual's specific career path. Because careers differ widely across workers with otherwise similar backgrounds, it seems crucial to take into account this feature when estimating a wage equation. Estimation of the model with and without taking into account the workers' careers paths (summarized by the newly developed non-linear function J^W) provide quantitatively different results, but the qualitative results remain unchanged. That is, the consistently large estimates for the returns to tenure stem largely from the fact that we fully endogenize the employment and mobility decisions, while also accounting for unobserved heterogeneity.

Our estimates of the returns to experience are also somewhat larger than those previously obtained in the literature. However, these findings are not uniform across all education groups, as they are higher for the college graduates than for the other two education groups. Also, the pattern of job-to-job mobility, as summarized by the non-linear function J^W, has differential impact on wages at the entry level in new jobs across the three education groups. In particular, and consistent with the displaced worker literature, while early job changes are most beneficial to college-educated workers, late job changes are more detrimental to workers with lower education.

Overall, wage growth is achieved through a combination of wage increases within the firm and by interfirm mobility. The former is more important for the wage growth of high school dropouts, because of their lower returns to experience. The latter is more important for college graduates, both because the returns to seniority are larger during the first few years in a given job, and because there is no penalty from job-to-job mobility.

This may not be the last word on this important issue, but this study certainly sheds new light on a variety of modeling issues leading to the very different findings obtained in this study. Further work on the issue, which includes cross-country comparisons, are certainly in order and may shed new light on the role of countries' institutions.

8 References

Abowd, J.M., Kramarz, F., and S. Roux (2006): "Wages, Mobility, and Firm Performance: Advantages and Insights from Using Matched Worker-Firm Data," *Economic Journal*, 116, 245-285.

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 (2005): "Do Wages Rise with Job Seniority? A Reassessment," *Industrial and Labor Relations Review*, 58, 370–97.

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

Bontemps, C., Robin, J. M., and G. Van den Berg (1999): "An Empirical Equilibrium Search Model with Continuously Distributed Heterogeneity of Workers' Opportunity Cost of Employment and Firms' Productivities, and Search on the Job," *International Economic Review*, 40, 1039-74.

Buchinsky, M., Fougère, D., Kramarz F. and R. Tchernis (2005): "Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.," IZA Discussion Paper No. 1521.

Burdett, K. (1978): "A Theory of Employee Job Search and Quit Rates," *American Economic Review*, 68, 212-20.

Card, D. (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometrics Problem," *Econometrica*, 69, 1127–60.

Chamberlain, G. (1984): "Panel Data," Chapter 22 in *Handbook of Econometrics*, Volume 2, ed. by Z. Griliches and M.D. Intrilligator. Amsterdam: Elsevier Science Publishers B.V.

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, 3569-49.

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

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

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

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

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

Gelfand, A.E. and A.F.M. Smith (1990): "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, 398-409.

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, 351-80.

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

Hardman, A. and Y.M. Ioannides (2004): "Neighbors' Income Distribution: Economic Segregation and Mixing in US Urban Neighborhoods," *Journal of Housing Economics*, 13, 368–82.

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, 1255-94.

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

Light A. and M. Ureta (1995): "Early-Career Work Experience and Gender Wage Differentials," *Journal of Labor Economics*, 3, 121–54.

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.

Mroz, T.A. (1987): "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica*, 55, 765–99.

Nagypal E. (2005): "On the Extent of Job-to-Job Transitions," mimeo, Department of Economics, Northwestern University.

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, 306-23.

Postel-Vinay, F. and J. M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70, 1295-350.

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

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

Topel, R.H. and M.P. Ward (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107, 439–479.

Table 1: Summary Statistics for the PSID Extract for Selected Years and Variables

Variable	Year		
	1975	1984	1992
Education	11.589	12.155	12.108
	(3.317)	(2.929)	(3.308)
Experience	20.808	21.330	24.169
	(14.364)	(14.148)	(12.877)
Seniority	5.152	5.993	7.285
	(7.686)	(7.142)	(7.563)
Participation	0.942	0.881	0.862
	(0.233)	(0.324)	(0.345)
Mobility	0.236	0.087	0.065
	(0.424)	(0.281)	(0.247)
Log Wage	9.869	9.902	9.924
	(0.789)	(0.877)	(0.842)
African American	0.309	0.320	0.254
	(0.462)	(0.466)	(0.435)
Hispanic	0.036	0.034	0.071
	(0.186)	(0.180)	(0.257)
Family unearned income	0.671	2.042	2.960
	(1.926)	(7.093)	(13.438)
No. of children	1.350	1.042	1.107
	(1.225)	(1.043)	(1.166)
Children 1 to 2	0.216	0.210	0.173
	(0.325)	(0.364)	(0.370)
Children 3 to 5	0.225	0.213	0.196
	(0.333)	(0.368)	(0.386)
Married	0.854	0.789	0.777
	(0.482)	(0.497)	(0.499)
Individual's cohorts by age in 1975:			
Age 15 or less	0	0.256	0.516
	(0)	(0.436)	(0.500)
Age 16 to 25	0.218	0.269	0.186
	(0.413)	(0.443)	(0.390)
Age 26 to 35	0.296	0.187	0.124
	(0.457)	(0.390)	(0.329)
Age 36 to 45	0.173	0.105	0.070
	(0.378)	(0.307)	(0.255)
No. of observations	3,385	4,451	5,397

Table 2: Participation by Education Group

Variable	High School Dropouts Mean	High School Dropouts Std	High School Graduates Mean	High School Graduates Std	College Graduates Mean	College Graduates Std
Constant	-0.507	0.347	0.185	0.296	-1.245	0.650
Education	-0.017	0.015	0.046	0.014	0.130	0.027
Lagged Experience	-0.029	0.043	-0.089	0.027	-0.068	0.061
Lagged Exp.$^2/100$	0.180	0.264	0.794	0.191	0.807	0.413
Lagged Exp.$^3/1,000$	-0.033	0.063	-0.251	0.050	-0.272	0.105
Lagged Exp.$^4/10,000$	-0.000	0.005	0.022	0.004	0.023	0.009
Lagged Participation	1.723	0.061	1.677	0.051	1.890	0.104
Lagged Mobility	0.654	0.132	0.420	0.076	0.372	0.156
Family Variables:						
Family unearned income	0.008	0.006	-0.004	0.003	-0.003	0.002
No. of Children	0.070	0.025	-0.019	0.022	0.193	0.065
Children 1 to 2	-0.054	0.065	-0.119	0.048	-0.158	0.127
Children 3 to 5	-0.122	0.064	-0.008	0.047	-0.062	0.142
Married	0.184	0.064	0.378	0.058	0.173	0.120
Ethnicity:						
African American	-0.464	0.094	-0.434	0.076	-0.224	0.173
Hispanic	-0.026	0.128	-0.051	0.146	-0.738	0.313

Note: Also included in the regression are a full set of year dummy variables, four regional location dummy variables, a dummy variable for living in an SMSA, county of residence unemployment rate, and a set of cohort dummy variables for the cohorts defined in Table 1.

Table 3: Mobility by Education Group

Variable	High School Dropouts Mean	High School Dropouts Std	High School Graduates Mean	High School Graduates Std	College Graduates Mean	College Graduates Std
Constant	-1.201	0.273	-1.293	0.165	-1.248	0.266
Education	-0.000	0.011	0.001	0.008	-0.006	0.010
Lagged Experience	0.002	0.030	-0.012	0.014	-0.014	0.027
Lagged Exp.$^2/100$	-0.186	0.207	-0.106	0.115	-0.232	0.211
Lagged Exp.$^3/1,000$	0.061	0.055	0.048	0.035	0.103	0.064
Lagged Exp.$^4/10,000$	-0.006	0.005	-0.005	0.003	-0.011	0.007
Lagged Seniority	-0.113	0.023	-0.161	0.012	-0.149	0.023
Lagged Sen.$^2/100$	0.729	0.306	1.035	0.155	1.133	0.319
Lagged Sen.$^3/1,000$	-0.215	0.138	-0.291	0.064	-0.393	0.153
Lagged Sen.$^4/10,000$	0.022	0.020	0.029	0.008	0.048	0.023
Lagged Mobility	-0.679	0.069	-0.854	0.037	-1.017	0.062
Family Variables:						
Family unearned income	-0.034	0.009	-0.027	0.003	-0.010	0.002
No. of Children	-0.006	0.015	0.009	0.011	-0.057	0.020
Children 1 to 2	0.063	0.039	0.036	0.021	0.056	0.035
Children 3 to 5	-0.017	0.040	0.004	0.022	-0.017	0.039
Married	-0.006	0.046	-0.091	0.027	-0.065	0.041
Ethnicity:						
African American	-0.027	0.040	-0.005	0.023	0.106	0.046
Hispanic	0.002	0.070	0.018	0.049	-0.019	0.107

Note: Also included in the regression are a full set of year dummy variables, four regional location dummy variables, a dummy variable for living in an SMSA, county of residence unemployment rate, and a set of cohort dummy variables for the cohorts defined in Table 1.

Table 4: Wage Equation by Education Group

Variable	High School Dropouts		High School Graduates		College Graduates	
	Mean	Std	Mean	Std	Mean	Std
Constant	7.662	0.139	7.842	0.082	8.143	0.137
Education	0.025	0.006	0.048	0.004	0.040	0.005
Experience	0.064	0.011	0.076	0.005	0.112	0.008
Exp.$^2/100$	-0.359	0.068	-0.467	0.032	-0.577	0.054
Exp.$^3/1,000$	0.088	0.016	0.123	0.009	0.134	0.015
Exp.$^4/10,000$	-0.008	0.001	-0.012	0.001	-0.013	0.001
Seniority	0.072	0.006	0.069	0.004	0.075	0.006
Sen.$^2/100$	-0.301	0.075	-0.285	0.040	-0.395	0.079
Sen.$^3/1,000$	0.097	0.030	0.077	0.016	0.134	0.034
Sen.$^4/10,000$	-0.011	0.004	-0.007	0.002	-0.016	0.005
Job Switch in 1st Sample Year:						
Job change in 1st year (ϕ_0^s)	-0.006	0.062	0.067	0.027	0.161	0.048
Lagged Experience (ϕ_0^e)	0.013	0.003	0.012	0.002	0.008	0.004
Job switches after:						
Up to 1 year (ϕ_{10})	0.049	0.023	0.109	0.009	0.213	0.017
2 to 5 years (ϕ_{20})	0.137	0.028	0.107	0.013	0.157	0.019
6 to 10 years (ϕ_{30})	0.215	0.076	0.190	0.042	0.327	0.070
Over 10 years (ϕ_{40})	0.057	0.106	0.373	0.054	0.517	0.088
Seniority at Job that Lasted:						
2 to 5 years (ϕ_2^s)	0.026	0.012	0.028	0.005	0.054	0.007
6 to 10 years (ϕ_3^s)	0.011	0.011	0.018	0.006	0.008	0.010
Over 10 years (ϕ_4^s)	0.026	0.005	0.035	0.003	0.002	0.006
Experience at Job that Lasted:						
Up to 1 year (ϕ_1^e)	0.003	0.001	-0.002	0.001	-0.006	0.002
2 to 5 years (ϕ_2^e)	-0.002	0.002	0.0001	0.001	-0.005	0.002
6 to 10 years (ϕ_3^e)	0.001	0.002	-0.003	0.002	-0.001	0.002
Over 10 years (ϕ_4^e)	0.002	0.003	-0.017	0.002	-0.004	0.003

Notes:

1. Also included in the regressions are a full set of year dummy variables, four regional location dummy variables, a dummy variable for living in an SMSA, county of residence unemployment rate, a dummy variable for African American, a dummy variable for Hispanic, and a set of cohort dummy variables for the cohorts defined in Table 1.

2. The ϕ coefficients in parentheses at the bottom part of the table are according to the definiton of the J^W function given in the text.

Table 5: Estimated Cumulative and Marginal Returns to Experience

Group	Cumulative Returns				Marginal Returns (in %)			
	Years of Experience				Years of Experience			
	5	10	15	20	5	10	15	20
HS Dropouts	.241	.362	.410	.420	3.442	1.550	0.482	-0.012
	(.042)	(.062)	(.071)	(.073)	(.597)	(.307)	(.231)	(.232)
HS Graduates	.277	.402	.440	.438	3.776	1.447	0.232	-0.234
	(.017)	(.024)	(.028)	(.031)	(.233)	(.142)	(.135)	(.134)
College Graduates	.430	.661	.762	.786	6.339	3.116	1.117	-0.045
	(.029)	(.044)	(.052)	(.058)	(.421)	(.264)	(.244)	(.244)

Table 6: Estimated Cumulative and Marginal Returns to Seniority

Group	Cumulative Returns				Marginal Returns (in %)			
	Years of Seniority				Years of Seniority			
	2	5	10	15	2	5	10	15
HS Dropouts	.133	.297	.507	.677	6.128	4.882	3.669	3.242
	(.010)	(.019)	(.024)	(.029)	(.408)	(.236)	(.249)	(.238)
HS Graduates	.127	.283	.475	.616	5.850	4.590	3.215	2.555
	(.006)	(.011)	(.014)	(.018)	(.227)	(.141)	(.150)	(.150)
College Graduates	.136	.294	.477	.615	6.109	4.513	3.035	2.630
	(.010)	(.019)	(.026)	(.034)	(.412)	(.257)	(.279)	(.272)

Table 7: Comparison of Alternative Estimates of Cumulative Returns to Seniority

	Panel A. Topel Replication Sample (1968-1983)			
	Years of Tenure			
	2	5	10	15
OLS (AW)	0.092	0.188	0.273	0.315
	(0.013)	(0.022)	(0.025)	(0.026)
IV1 (AW)	0.053	0.097	0.119	0.123
	(0.010)	(0.019)	(0.028)	(0.037)

	Panel B. Our Sample (1975-1992)							
	High School Dropouts				College Graduates			
	Years of Tenure				Years of Tenure			
	2	5	10	15	2	5	10	15
OLS (AW)	0.099	0.197	0.273	0.300	0.068	0.136	0.189	0.208
	(0.009)	(0.015)	(0.016)	(0.017)	(0.007)	(0.012)	(0.013)	(0.015)
IV1 (AW)	0.062	0.112	0.131	0.131	0.043	0.078	0.092	0.090
	(0.010)	(0.017)	(0.019)	(0.018)	(0.008)	(0.014)	(0.015)	(0.016)
Our Model	0.133	0.297	0.507	0.677	0.136	0.294	0.477	0.615
	(0.010)	(0.019)	(0.024)	(0.029)	(0.010)	(0.019)	(0.026)	(0.034)

Note: The first panel estimates are taken from AW, Table 2. The OLS and IV 1 (based on Altonji and Williams methodology) estimates for high school dropouts and college graduates are based on our sample extract. The estimates for our model come from Table 6.

Table 8: Test of Exogeneity for Family Background Variables

Model Specification	Unobserved Heterogeneity Terms α_y H_0^A	α_y H_0^B	α_m H_0^A	α_m H_0^B	α_w H_0^A	α_w H_0^B
Quartic	AC (.1571)	AC (.2516)	AC (.1932)	AC (.2922)	RJ (.0870)	AC (.1812)
Quadratic	AC (.2224)	AC (.3111)	AC (.3210)	AC (.4091)	AC (.2046)	AC (.2891)

Notes:
(i) The test employed is the usual Wald test. AC denotes that the null hypothesis of all coefficients are zero is accepted, while RJ denotes that this null hypothesis is rejected.
(ii) The numbers in parentheses are the corresponding p-values for the tests.
(iii) The specific null hypotheses are defined in (19) and (20).
(iv) The set of variables for which the test is applied are: family unearned income, number of children in the family, number of children age one to two, number of children age three to five, and marital status (see the raw statistics for these variables Table 1).

Table 9: Estimates of the Stochastic Elements by Education Group

Variable	High School Dropouts Mean	High School Dropouts Std	High School Graduates Mean	High School Graduates Std	College Graduates Mean	College Graduates Std
Covariance Matrix of White Noises (Elements of Σ)						
ρ_{uv}	-0.0012	0.0098	-0.0019	0.0072	-0.0021	0.0103
$\rho_{u\xi}$	-0.0251	0.0069	-0.0323	0.0060	-0.0299	0.0135
$\rho_{v\xi}$	0.0100	0.0071	0.0064	0.0049	0.0210	0.0094
σ_ξ^2	0.2954	0.0038	0.2086	0.0016	0.2048	0.0023
Correlations of Individual Specific Effects (Elements of Δ_ρ)						
$\rho_{\alpha_y\alpha_m}$	-0.1835	0.1762	-0.5024	0.1033	-0.5521	0.0746
$\rho_{\alpha_y\alpha_w}$	0.3572	0.0368	0.3574	0.0262	0.2007	0.0508
$\rho_{\alpha_m\alpha_w}$	-0.1846	0.3343	-0.6980	0.1589	-0.7682	0.0739

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

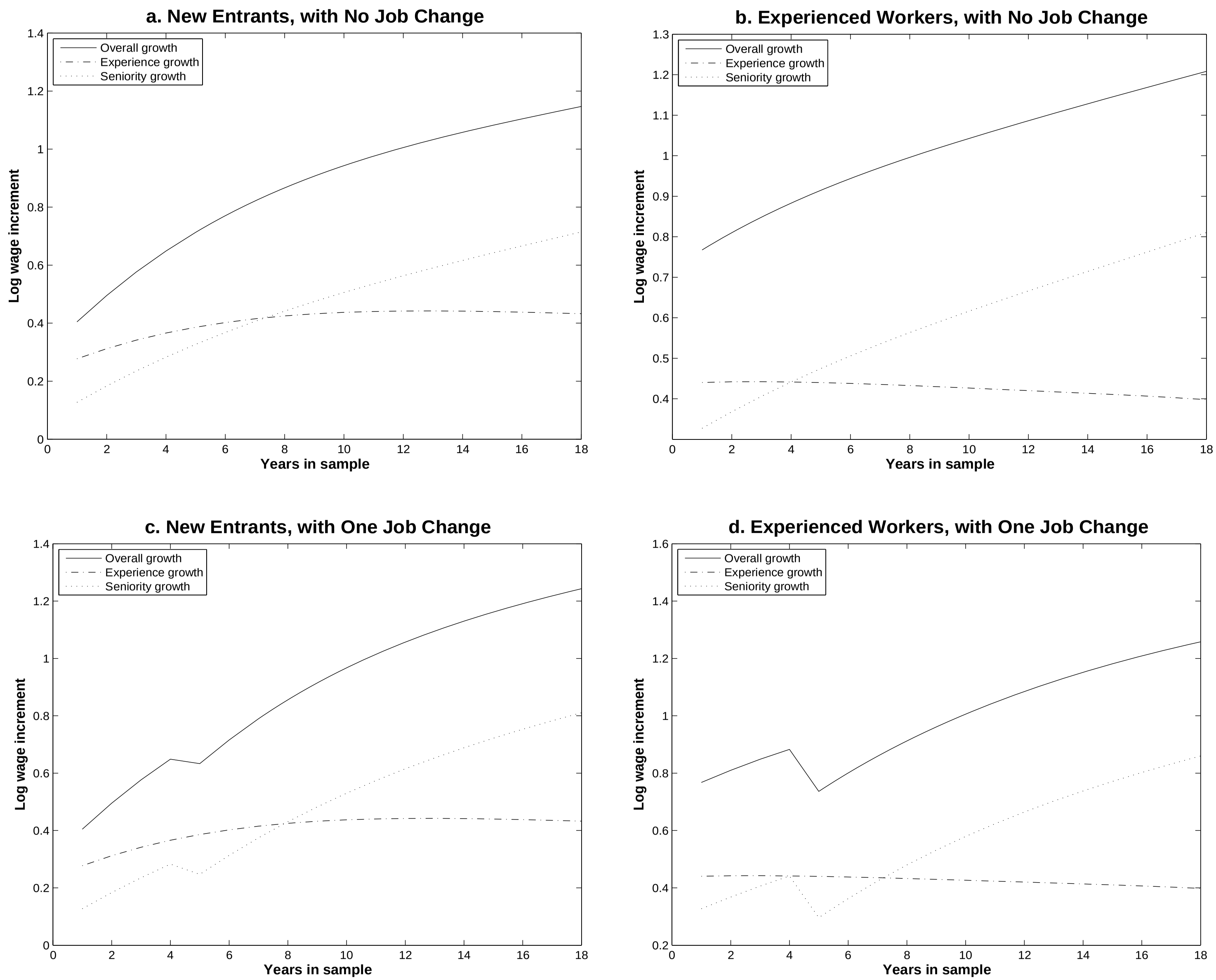

Figure 2: College Graduates–Wage Growth Due to Experience and Seniority