Modeling Earnings Dynamics[1]

Joseph G. Altonji*, Anthony A. Smith**, and Ivan Vidangos***

July 8, 2008

[1]*Yale University and NBER. **Yale University. ***Federal Reserve Board. We are grateful to Richard Blundell, Mary Daly, Costas Meghir, and Luigi Pistaferri for helpful discussions and suggestions. We also thank participants in seminars at UC Berkeley, CEMFI, U. of Chicago, the Federal Reserve Bank of San Francisco, Georgetown U., Pennsylvania State U., Princeton, U. of Rochester, Stanford, Vanderbilt, and Yale, and conference sessions at the Society of Economic Dynamics (June 2005), the World Congress of the Econometric Society (August 2005), the Cowles Foundation Macro/Labor Economics Conference (May 2006), NBER (Nov. 2006) and the Econometric Society Winter Meetings (January 2007) for valuable comments. Our research has been supported by the Cowles Foundation and the Economic Growth Center, Yale University, and by NSF grant SES-0112533 (Altonji). The views expressed in this paper are our own and do not necessarily represent the views of the Federal Reserve Board or of other members of its staff. We are responsible for the remaining shortcomings of the paper.

Abstract

In this paper we use indirect inference to estimate a joint model of earnings, employment, job changes, wage rates, and work hours over a career. Our model incorporates duration dependence in several variables, multiple sources of unobserved heterogeneity, job-specific error components in both wages and hours, and measurement error. We use the model to address a number of important questions in labor economics, including the source of the experience profile of wages, the response of job changes to outside wage offers, and the effects of seniority on job changes. We provide estimates of the dynamic response of wage rates, hours, and earnings to various shocks and measure the relative contributions of the shocks to the variance of earnings in a given year and over a lifetime. We find that human capital accounts for most of the growth of earnings over a career although job seniority and job mobility also play a significant role. Unemployment shocks have a large impact on earnings in the short run as well a substantial long long-term effect that operates through the wage rate. Shocks associated with job changes and unemployment make a large contribution to the variance of career earnings and operate mostly through the job-specific error components in wages and hours.

1 Introduction

In this paper we build and estimate a simultaneous model of earnings. The model consists of equations for transitions into and out of employment, an equation for job to job mobility, a wage equation, an hours equation, and an earnings equation. The model features both observed and unobserved permanent heterogeneity, job specific wage and hours components, a persistent component that affects the wage of workers in all jobs, state dependence in employment and job mobility, tenure and experience effects, and measurement error.

We have three main goals. The first is to advance the literature in labor economics on how employment, hours, and wages are determined over a career. We examine the effects of education, race, experience, job tenure and unobserved heterogeneity, employment shocks, shocks to general skills, and draws of new job opportunities offering different hours and wages. We trace out the response of wages, hours, and earnings to the various shocks and determine the channels through which they operate. Our analysis has implications for a number of longstanding questions in labor economics. For example, we provide estimates of the relative importance of general skill accumulation, job shopping, and job tenure for career wage growth and quantify the specific channels through which an exogenous employment shock affects the path of wage rates, hours, and earnings. We study the effects of shocks on the future variance of earnings changes as well as on the average path.

Our second goal is to provide a comprehensive account of what causes inequality in earnings at a point in time and over the life-time. We measure the contribution of each of the various shocks, permanent unobserved heterogeneity, and education to the variance in earnings, wages, and hours over the course of a career.

Our third goal is to provide a richer model of earnings for use in studies of consumption and saving as well as in dynamic stochastic general-equilibrium models that are a cornerstone of modern macroeconomics and public finance. Such models have been used to study the distribution of wealth, the costs of business cycles, asset pricing, and other important questions.[1] The quantitative implications of the calibrated theoretical models used in these lines of research depend on certain key features of the earnings process, such as the degree

[1]Examples include Huggett (1996), Krusell & Smith (1998), Castañeda, Díaz-Giménez, & Ríos-Rull (2003), Storesletten, Telmer, & Yaron (2004a)) on consumptions and wealth, Imrohoroglu (1989), Krusell & Smith (1999), Storesletten, Telmer, & Yaron (2001a) on the costs of business cycles, and Telmer (1993), Heaton & Lucas (1996), Krusell & Smith (1997), Storesletten, Telmer, & Yaron (2007) on asset pricing.

of earnings uncertainty and the persistence of earnings innovations.[2]

Almost all of the existing structural studies base their modeling and calibration choices for the earnings process on the large empirical literature on univariate statistical models.[3] Much has been learned about the statistical properties of career earnings from this work. However, with only one indicator, univariate models, no matter how richly specified, cannot identify the various sources of earnings fluctuations, their relative importance, their dynamic behavior, or the economics underlying how labor market outcomes are determined. Without such information, it is difficult to think about the potential welfare consequences of specific sources of variation or of policies such as unemployment insurance, employment regulations, wage subsidies, or earned income tax credits that insure against particular types of shocks to income. Furthermore, the innovations in the univariate representation of a multivariate time series process may be aggregates of current and past shocks in the multivariate representation. This will lead to mistakes in characterizing what the surprises to the agent are even under the assumption that the agent's information set is the same as the econometrician's.

Only a few of the studies of earnings dynamics have considered multivariate models. These include Abowd and Card's (1987, 1989) analyses of hours and earnings, and Altonji, Martins, and Siow's (2002) second order vector moving average model of the first difference in family income, earnings, hours, wages, and unemployment. Altonji, Martins and Siow's use of their model to study consumption and labor supply behavior and decompose the variance of innovations in the marginal utility of income into various sources is not entirely successful, but it does illustrate the potential that a multivariate model of the income process provides. The models that we consider, in contrast to those mentioned above, incorporate discrete events such as job changes, employment loss, interactions between job changes and wages, and effects of these discrete events on the variance of wage and hours shocks.[4]

There are two distinct paths that one might take in formulating a multivariate model of earnings. The first approach is the development of a statistical model of the process with little attention to an underlying theory of household decisions and constraints. This

[2]See, for example, Deaton (1991), Aiyagari (1994), Krusell and Smith (1997), Guvenen (2007), and the discussion in Blundell, Pistaferri, and Preston (2008).

[3]Key early contributions include Lillard and Willis (1978), Lillard and Weiss (1979), Hause (1980), MaCurdy (1982). More recent contribution include Baker (1997), Geweke and Keane (2000), Haider (2001), Baker and Solon (2003), Guvenen (2007), and Meghir and Pistaferri (2004). The later paper introduces ARCH shocks.

[4]A number of recent studies provide structural models of wage rates, job mobility, and employment dynamics, including Barlevy (2008), Buchinsky et al (2008), and Bagger et al (2007), who provide references to a few additional studies. Wolpin (1992) is an early effort. We discuss the evidence below.

approach is in the spirit of the literature on univariate earnings processes, but the absence of theory limits what one can learn about how earnings are determined. The second approach is to develop a model that is based on lifetime utility maximization. Grounding the model of the income process in a utility maximization framework provides a foundation for using the results to analyze policies when earnings are partially endogenous. The main disadvantage is the difficulty of specifying and estimating a model that incorporates labor supply choices, job search decisions, hours constraints, voluntary separations, and involuntary job changes. Indeed, we do not know of any papers that have studied work hours and employment using a lifecycle utility maximization model that incorporates job specific hours constraints, let alone job mobility decisions.[5] Estimation of a structural model that is as rich as the one that we work with would require solving an intertemporal model of job search, labor supply (in the presence of hours constraints), and savings as part of the estimation strategy and is probably out of reach at the present time from a computational point of view. Low, Meghir, and Pistaferri (2006) take an important step in this direction by studying earnings risk and social insurance in the context of an intertemporal model of consumption, employment participation, wages, and mobility. However, they work with a simpler model of the earnings process than we do.

Although our model falls short of a fully specified behavior model, the equations can be viewed as approximations to the decision rules relating choices to state variables that would arise in a structural model based on lifetime utility maximization. The parameters of the rules depend on an underlying set of "deep" parameters that characterize consumption preferences, job search technology, etc. The class of models that we consider is rich enough to address a number of core behavior questions in labor economics, but tractable enough (at least the simpler versions) to be used in place of univariate income models that dominate the literature on savings, portfolio choice, etc. Furthermore, it provides a natural path along which to extend the analysis to include other important economic risks that individuals face, including changes in family structure through marriage, divorce, and the death of a spouse.

We estimate the model using data on male household heads from the Panel Study of Income Dynamics. Given the presence of interactions among discrete and continuous variables, unobserved heterogeneity and state dependence in multiple equations, measurement error,

[5]Ham and Reilly (2002) is part of a literature that tests for hours restrictions in an intertemporal labor supply and consumption framework using Euler equations and within period marginal rate of substitution conditions. Blundell and MaCurdy (1999) survey the labor supply literature.

and a highly unbalanced sample, conventional maximum likelihood and method of moments approaches are not feasible. For this reason, we use indirect inference (I-I), which is one of a family of simulation based approaches to estimation that involve comparing the distribution of artificial data generated from the structural model at a given set of parameter values to features of the actual data.[6] A complication arises in our case because our model includes discrete as well as continuous variables. With discrete variables, the simulated values of moments of the artificial data are not continuous in model parameters, which makes gradient based numerical optimization methods problematic. Given our model size, derivative-free methods are also unattractive. Consequently, we use a smoothed version of the procedure suggested by Keane and Smith (2003). Estimation of our model is not straightforward, and a secondary contribution of our research is to explore the feasibility and performance of I-I in large models with a mix of discrete and continuous variables.[7]

There are too many results to concisely summarize, but a few deserve emphasis. First, education, race, and the two forms of unobserved permanent heterogeneity play an important role in employment transitions and job changes. Second, in keeping with a large literature on the labor supply of male household heads, wages have only a small (negative) effect on employment and on annual work hours. Third, even after accounting for unobserved individual heterogeneity and job specific heterogeneity, we find a strong negative tenure effect on job mobility. Fourth, job changes are induced by high outside offers and deterred by the job specific wage component of the current job.

Fifth, unemployment at the survey date is associated with a large decline of .62 log points in annual earnings. About 60 percent of the reduction is due to work hours, which recover almost completely after one year. The other 40 percent is due to a .25 decline in log hourly wage rate. Lost tenure and a drop in the job-specific wage component contribute .064 and .027, respectively, to the reduction. The wage recovers by about .10 after 1 period and more slowly after that.

Sixth, wages do not contain a random walk component but are highly persistent. The persistence is the combined effect of permanent heterogeneity, the job specific wage component, and strong persistence in a stochastic component representing the value of the worker's

[6] The method was introduced, under a different name, in Smith (1990, 1993) and later extended by Gourieroux, Monfort, and Renault (1993) and Gallant and Tauchen (1996). It is closely related to the simulated method of moments.

[7] Other recent papers that apply I-I to panel data include Bagger et al (2007), Nagypal (2007) and Tartari (2006).

general skills.

Seventh, shocks leading to unemployment or to job changes have large effects on the variance as well as the mean of earnings changes. Eighth, job shopping, the accumulation of tenure, and the growth in general skills account for log wage increases of .111, .122, and .580, respectively, over the first thirty years in the labor market.

Finally, job mobility and unemployment play a key role in the variance of career earnings. Job specific hours and wage components, unemployment shocks, and job shocks together account for 36.7%, 48.2%, and 46.8% of the variance in lifetime earnings, wages, and hours, respectively. Job specific hours and wage shocks are about equally important for earnings. Job specific wage shocks dominate for wages, with employment shocks also playing a substantial role. For hours, job specific hours shocks dominate. Education accounts for about 1/3 of the variance in lifetime earnings and wages but makes little difference for hours. In our full sample, unobserved permanent heterogeneity accounts for about 11% of the variance in earnings and about 46% of the variance of hours but matters little for wages, although this breakdown is somewhat sensitive to the model and sample used.

The paper continues in section 2, where we present the earnings model. In section 3 we discuss the data, which is drawn from the Panel Study of Income Dynamics (PSID) and in section 4 we discuss estimation. We present the results in Section 5, beginning with a discussion of the parameter estimates and then turning to an analysis of the fit of the model, impulse response functions to various shocks, and variance decompositions. In Section 6, we briefly discuss results for alternative samples, including whites by education level. In the final section we summarize our main findings and provide a research agenda.

2 Models of Earnings Dynamics

We work with two classes of models, which we call A and B. The main difference between the two is the treatment of job specific wage and hours components. Model A includes an autoregressive wage component as well as a job match component that is fixed within a job and also influences job mobility. The hours equation also includes a job specific component. Model B does not include job specific wage or hours components. However, it allows the dependence of the current wage on the past wage and the variance of wage shocks to depend on whether the individual is continuing an existing job. We first define notation and list the equations of Model A and then discuss the model. We then turn to Model B.

2.1 Model A.

The equations of the model and variable definitions are described below. A word about notation first. We control for economy wide effects using year dummies, but leave them implicit in most of the analysis. The subscript i, which we sometimes suppress, refers to the individual, t_i is potential labor market experience of i for a particular observation. We sometimes refer to it as "time" and usually suppress the i subscript. The subscript $j(t)$ refers to the job that i holds at t. The notation $j(t)$ makes explicit the fact that individuals may change jobs. In particular, $j(t) \neq j(t-1)$ if i experiences a job change without being unemployed at either t or $t-1$ or if i is employed at t but was unemployed at $t-1$. The γ parameters refer to intercepts and to slope coefficients. For each intercept and slope parameter the superscripts identify the dependent variable. The subscripts of slope parameters identify the explanatory variable. We use δ to denote coefficients on the fixed person specific unobserved heterogeneity components μ_i and η_i, the job match heterogeneity wage component $\upsilon_{ij(t)}$, and the job specific hours component $\xi_{ij(t)}$. The superscripts for the δ parameters denote the dependent variable and the subscripts μ and η identify the heterogeneity component. We use ρ with appropriate subscripts to denote autoregression coefficients. The ε_{it}^{k} are iid $N(0, \sigma_k^2)$ random variables where the superscripts k correspond to the dependent variables.

The equations of Model A are as follows.

Employment to Employment Transition (EE)

$$E_{it} = I[\gamma_0^{EE} + \gamma_t^{EE}(t_i - 1) + \gamma_{t^2}^{EE}(t_i - 1)^2 + \gamma_{\hat{w}}^{EE} wage_{it}^{'} + \gamma_{BLACK}^{EE} BLACK_i + \gamma_{EDUC}^{EE} EDUC_i + \gamma_{ED}^{EE} \min(ED_{i,t-1}, 9) + \delta_\mu^{EE}\mu_i + \delta_\eta^{EE}\eta_i + \varepsilon_{it}^{EE} > 0] \text{ given } E_{i,t-1} = 1$$

where E_{it} is an employment dummy, $I(\cdot)$ is an indicator function, $ED_{i,t-1}$ is lagged employment duration and is determined recursively by $ED_{it} = E_{it} \cdot (ED_{i,t-1} + 1)$, and $wage_{it}^{'}$ is what the wage would be in t if the individual were to continue employment in the job held at $t-1$.

Unemployment to Employment Transition (UE):

(1)

$$E_{it} = I[\gamma_0^{UE} + \gamma_t^{UE}(t_i - 1) + \gamma_{t^2}^{UE}(t_i - 1)^2 + \gamma_{BLACK}^{UE} BLACK_i + \gamma_{EDUC}^{UE} EDUC_i + \gamma_{UD}^{UE} UD_{i,t-1} + \delta_\mu^{UE}\mu_i + \delta_\eta^{UE}\eta_i + \varepsilon_{it}^{UE} > 0] \text{ given } E_{i,t-1} = 0,$$

where $UD_{i,t-1}$ is the number of years unemployed at the survey date and $UD_{it} = (1 - E_{it}) \cdot (UD_{i,t-1} + 1)$.

Job Change While Employed (JC):

(2)
$$JC_{it} = I[\gamma_0^{JC} + \gamma_t^{JC}(t_i - 1) + \gamma_{t^2}^{JC}(t_i - 1)^2 + \gamma_{TEN}^{JC} TEN_{i,t-1} + \gamma_{BLACK}^{JC} BLACK + \gamma_{EDUC}^{JC} EDUC$$
$$+\delta_{\upsilon' j'(t)} \upsilon'_{i,j'(t)} + \delta_{\upsilon j(t-1)} \upsilon_{i,j(t-1)} + \delta_\mu^{JC} \mu_i + \delta_\eta^{JC} \eta_i + \varepsilon_{it}^{JC} > 0] \cdot E_{it} \cdot E_{i,t-1}$$

where $\upsilon_{i,j(t-1)}$ is a job specific error component, $\upsilon'_{i,j'(t)}$ is a draw of the job specific component for an alternative job $j'(t)$ in t, and $TEN_{i,t-1}$ is employer tenure at the previous survey date, which evolves according to

$$TEN_{it} = (1 - JC_{it}) \cdot E_{it} \cdot E_{i,t-1} \cdot (TEN_{i,t-1} + 1).$$

Log Wages:

$$wage_{it}^{lat} = \gamma_0^w + \gamma_X^w X_{it} + \gamma_{TEN}^w P(TEN_{it}) + \delta_\mu^w \mu_i + \upsilon_{ij(t)} + \omega_{it} \quad (3)$$
$$\upsilon_{ij(t)} = (1 - S_{it}) \upsilon_{ij(t-1)} + S_{it} \upsilon'_{ij'(t)} \quad (4)$$
$$\upsilon'_{ij'(t)} = \gamma_{JC}^\upsilon JC_{it} + \rho_\upsilon \upsilon_{i,j(t-1)} + \varepsilon_{ij(t)}^\upsilon \quad (5)$$
$$\omega_{it} = \rho_\omega \omega_{i,t-1} + \gamma_{1-E_{it}}^\omega (1 - E_{it}) + \gamma_{1-E_{i,t-1}}^\omega (1 - E_{i,t-1}) + \varepsilon_{it}^\omega \quad (6)$$
$$wage_{it} = E_{it} \cdot wage_{it}^{lat} \quad (7)$$

where $wage_{it}^{lat}$ is the "latent" wage, which we define below, X_{it} is a vector of exogenous variables including t, $BLACK_i$ and $EDUC_i$, $P(TEN_{it})$ is a fourth order polynomial in TEN_{it}, $\upsilon_{ij(t)}$ is the job match specific wage component, ω_{it} is an autoregressive component of the latent wage, $S_{it} = (JC_{it} + E_{it}(1 - E_{i,t-1}))$ is a job separation indicator that equals 1 if JC is 1 or if the individual was unemployed in $t - 1$ and employed in t. The variable $wage_{it}$ is the actual wage rate, which we define as 0 for persons who are unemployed.

Log Annual Work Hours of the Head of Household

$$hours_{it} = \gamma_0^h + \gamma_X^h X_{it} + (\gamma_E^h + \xi_{ij(t)}) E_{it} + \gamma_w^h wage_{it}^{lat} + \delta_\mu^h \mu_i + \delta_\eta^h \eta_i + \varepsilon_{it}^h \quad (8)$$

where $\xi_{ij(t)}$ is a job match specific hours component.

Log earnings

$$\begin{aligned} earn_{it} &= \gamma_0^e + \gamma_X^e X_{it} + \gamma_w^e (wage_{it}^{lat} - \gamma_0^w - \gamma_X^w X_{it}) + \gamma_h^e (hours_{it} - \gamma_0^h - \gamma_X^h X_{it}) + e_{it} \\ e_{it} &= \rho_e e_{i,t-1} + \varepsilon_{it}^e \end{aligned}$$

Error Components and Initial Conditions:

The fixed person specific error components μ_i and η_i are $N(0,1)$, *iid* across i, independent of each other, and independent of all transitory shocks and measurement errors. We parameterize the errors of the various equations so that μ_i may be thought of as the fixed unobserved heterogeneity component of wages. We also allow μ to influence EE, UE, JC, and $hours$. The factor η_i is assumed to have no influence on wages. One may think of it as a factor that is related to labor supply and to job and employment mobility preferences. We impose the sign normalizations $\delta_\mu^w > 0$ and $\delta_\eta^{JC} > 0$.

The job match hours component $\xi_{ij(t)}$ and the innovation $\varepsilon_{it}^{\upsilon}$ in $\upsilon_{ij(t)}$ are $N(0, \sigma_\xi^2)$ and $N(0, \sigma_\upsilon^2)$ respectively. The shocks ε_{it}^{EE}, ε_{it}^{UE}, ε_{it}^{JC}, $\varepsilon_{it}^{\omega}$, ε_{it}^{e} are $N(0, \sigma_k^2)$, where $k = EE, UE, JC, \omega$, and e. They are iid across i and t and independent from one another and all measurement error components defined below.

The initial conditions are

$$\begin{aligned} Employment&: \quad E_{i1} = I[b_{0g} + \delta_\mu^{EE}\mu_i + \delta_\eta^{EE}\eta_i + \varepsilon_{i1}^{EE} > 0] \\ Wages&: \quad wage_{i1}^{lat} = \gamma_0^w + \gamma_X^w X_{i1} + \upsilon_{ij(1)} + \omega_{i1} + \delta_\mu^w \mu_i \\ &\quad\;\; \omega_{i1} \tilde{} N(0, \sigma_{\omega 1,g}^2) \\ Wage\ Job\ Match &: \quad \upsilon_{ij(1)} \sim N(0, \sigma_{\upsilon 1}^2) \\ Earnings\ Error &: \quad e_{i1} \sim N(0, \sigma_e^2) \\ Other\ Initial\ Conditions&: \quad TEN_{i1} = 0, ED_{i1} = E_{i1}, UD_{i1} = 1 - E_{i1}, JC_{i1} = 0. \end{aligned}$$

The intercept b_{0g} of the initial employment condition and the variance of initial wages $\sigma_{\omega 1,g}^2$ depend on the race-education group g, where the groups are ($BLACK$ & $EDUC \leq$

12), ($BLACK$ & $EDUC > 12$), (*not* $BLACK$ & $EDUC \leq 12$), and (*not* $BLACK$ & $EDUC > 12$).

Measurement Error and Observed Wages, Hours, and Earnings:

The observed (measured) variables are:

$$wage^*_{it} = E_{it} \cdot (wage^{lat}_{it} + m^w_{it}) \tag{9}$$

$$hours^*_{it} = hours_{it} + m^h_{it} \tag{10}$$

$$earn^*_{it} = earn_{it} + m^e_{it} \tag{11}$$

The measurement errors m^w_{it}, m^h_{it}, m^e_{it} are $N(0, \sigma^2_{m\tau})$, $\tau = w, h, e$, iid across i and t, mutually independent, and independent from all other errors components in the model.

2.2 Discussion of Model A

The EE equation states that the latent variable that determines E_{it} for previously employed workers depends on a quadratic in t_i, a linear function of $ED_{i,t-1}$ with a ceiling at 9 years, $BLACK_i$, $EDUC_i$ and the error $\delta^{EE}_{\mu}\mu_i + \delta^{EE}_{\eta}\eta_i + \varepsilon^{EE}_{it}$. Early on we experimented with including $TEN_{i,t-1}$ as well as $ED_{i,t-1}$ but in simulation experiments found that we had trouble distinguishing the effects of the two. Standard labor supply models imply that employment at t should depend on the current wage opportunity, which we proxy with $wage'_{it}$. It also depends on the permanent wage heterogeneity component μ_i as well as the hours preference and mobility component η_i.

The UE transition probability has the same form as EE, with unemployment duration $UD_{i,t-1}$ replacing $ED_{i,t-1}$. Because there are relatively few multi-year unemployment spells, we exclude $UD_{i,t-1,}$ restricting γ^{UE}_{UD} to 0 in most of the analysis. We experimented with specifications containing the lagged latent wage rate or the expected value of the period t wage but had difficulty identifying the effects of these variables, perhaps because we observe relatively few unemployment spells. We do include the wage heterogeneity component μ_i as well as η_i.

The JC equation refers to job to job changes for workers who are employed in both t and $t-1$. In our specification of the link between mobility and wages, the main distinction we draw is between job changes from employment and job changes that involve unemployment. We believe that this is the most important distinction both for the determination of wages and annual work hours, although it would be interesting in future work to distinguish between quits and layoffs on the basis of self reports.

Standard job search and job matching models predict a negative coefficient on $\upsilon_{ij(t-1)}$, since higher values of the job match component of the current job should reduce search activity and raise the reservation wage. In the model each worker is assigned a potential draw of $\upsilon_{ij'(t)}$ based on (5), which we discuss momentarily. Search models predict a positive coefficient on $\upsilon_{ij'(t)}$, but the magnitude should depend on the probability that the worker actually receives the offer. That is, the relative magnitudes of the two coefficients should depend on offer arrival rates and need not be equal. We include $TEN_{i,t-1}$ as well as $(t-1)$ because models of firm financed or jointly financed specific capital investment suggest that it will play a role, and the decline in separation rates with $TEN_{i,t-1}$ in cross section data is very strong. Little is known about how much of the association between $TEN_{i,t-1}$ and JC_{it} is causal because of the difficulty of distinguishing state dependence from the individual heterogeneity (μ and η) and job match heterogeneity (υ) in dynamic discrete choice models, particularly when data are missing on early employment histories for most sample members. Indeed, to the best of our knowledge, our study is one of only two to account for both individual and job specific heterogeneity and deal with initial conditions problems when estimating the effects of TEN and t on job changes.[8]

When interpreting results for EE and JC, one must keep in mind that our employment indicator refers to the survey date. We undoubtedly miss short spells of unemployment that fall between surveys. Due to data limitations, we cannot tell whether a person has changed jobs between surveys only once or multiple times. Furthermore, if a person is employed at $t-1$, unemployed for part of the year, and employed in a new job at t, we would count this as a job to job change even if, for example, the job change is due to a layoff into unemployment. A relatively simple alternative would be to make use of information on the number of weeks that the individual was unemployed during the year. However, one would want to distinguish between short spells of unemployment that are associated with temporary layoffs with the strong expectation of recall and unemployment spells due to a permanent layoff, which cannot be done except at the survey date. Fortunately, earnings depend on employment through annual work hours and the transitory error component in

[8] Buchinsky et al (2008) also find negative effects in a simultaneous model of wages, employment, and job changes. Farber (1999) discusses models of the effect of tenure on mobility and surveys the empirical evidence. He presents evidence showing a negative effect of tenure when one uses prior mobility as a control for individual heterogeneity.

the hours equation should capture the effect on hours of unemployment spells of varying duration.[9]

The wage model (3) is unusual in our use of the concept of a latent wage. For employed individuals $wage_{it}^{lat}$ and the actual wage $wage_{it}$ are the same. For an unemployed individual $wage_{it}^{lat}$ captures the process for wage offers that exceed $i's$ reservation wage. At a given point in time the individual might not have such an offer. Our formulation allows us to capture in a parsimonious way the idea that worker skills and worker specific demand factors evolve during an unemployment spell. From a practical point of view, the formulation also allows us to deal with the fact that wages are only observed for jobs that are held at the survey date.

The variable $wage_{it}^{lat}$ depends on five components. The first is the regression index $\gamma_X^w X_{it}$, which captures the effects of potential experience t_i,, education, race, and economy wide variation (through year dummies). The second is tenure. The third is the heterogeneity component μ_i. The fourth is a stochastic component ω_{it}, which depends on $\omega_{i,t-1}$, unemployment, and the error component $\varepsilon_{it}^{\omega}$. The dependence of ω_{it} on the past may reflect persistence in the market value of the general skills of i and/or the fact that employers base wage offers on past wages. We will have more to say about the second mechanism when we turn to model B. The fifth is the job match specific term $\upsilon_{ij(t)}$. When persons leave unemployment or move from job to job without unemployment, they draw a new value of $\upsilon_{ij(t)}$. The new value depends on $\upsilon_{ij(t-1)}$, a mean shift term γ_{JC}^{υ} in the case of a job change without unemployment, and the shock $\varepsilon_{ij(t)}^{\upsilon}$. We set $\gamma_{JC}^{\upsilon} = 0$ when $\upsilon_{ij(t-1)}$ and $\upsilon_{ij(t)}$ are included in the JC_t model (models A.2 and A.3 below), because in that case any shift in the mean of $\upsilon_{ij(t)}$ is accounted for endogenously by the effects of $\upsilon_{ij(t-1)}$ and $\upsilon_{ij(t)}$ on mobility. In standard search models with exogenous offer arrivals, the job specific component of the offer, $\upsilon'_{ij'(t)}$ does not depend on $\upsilon_{ij(t-1)}$ although accepted offers $\upsilon_{ij(t)}$ will. In such models the correlation between accepted offers $\upsilon_{ij(t)}$ and $\upsilon_{ij(t-1)}$ arises because the reservation wage

[9] It is conceptually straightforward to specify the model on a quarterly or monthly basis. Simulated data that matches the periodicity, level time of aggregation, and dating within the calendar year of the various PSID variables could be constructed from the higher frequency data from the model. One could use both measures of weeks of unemployment over the previous calendar year and unemployment at the survey date. One would have to think carefully about the specification of shocks—few employers reset wages on a monthly or quarterly basis. One might also wish to incorporate distributed lags, along the lines of Altonji, Martins and Siow's (2002) quarterly model. A quarterly reformulation would substantially increase computation time by a factor of 4 and would require some substantial programming changes, but is probably feasible with our current computer resources. On the other hand, we believe that there is merit in starting with the simpler specification that we employ.

is a positive function of $\upsilon_{ij(t-1)}$. Nevertheless, we allow offers $\upsilon'_{ij'(t)}$ to depend on $\upsilon_{ij(t-1)}$ through the parameter ρ_v for three main reasons. The first is that employers may base offers to prospective new hires in part on wages in the prior firm, including the firm specific component. Bagger et al (2007), building on Postel-Vinay and Robin (2002) and Postel-Vinay and Turon (2005), is one of a few recent papers in which outside firms tailor offers to surplus in the current job. This surplus will be related to $\upsilon_{ij(t-1)}$ to the extent that $\upsilon_{ij(t-1)}$ is the worker's portion of a job specific productivity component. However, in contrast to those papers, we do not allow the current employers to change $\upsilon_{ij(t-1)}$ in response to outside offers. (Wages do change through ω_{it}.). The second reason $\upsilon'_{ij'(t)}$ will depend on $\upsilon_{ij(t-1)}$ is that $\upsilon_{ij(t-1)}$ is not likely to be entirely job specific in the presence of demand shocks affecting jobs in a narrowly defined industry, occupation, and region. The third is that the network available to an individual may be related to the quality of the job that he is in. As it turns out, our estimates of ρ_v are large—about .60.[10] We were not successful in limited experimentation with estimating models in which the link between $\upsilon_{ij(t)}$ and $\upsilon_{ij(t-1)}$ when $JC = 1$ differs from the link following unemployment, although standard job search models with exogenous layoffs imply that it should.

The equation for $hours_{it}$ includes X_{it}. It also includes η_i, μ_i, and the product of the job specific hours component $\xi^h_{ij(t)}$ and E_{it}. We include $\xi^h_{ij(t)}$because there is strong evidence that work hours are heavily influenced by a job specific component. This componsent presumably reflects work schedules imposed by employers.[11] A new value of $\xi^h_{ij(t)}$is drawn when individuals change jobs. The iid error component ε^h_{it} picks up temporal variation in overtime, multiple job holding, and unemployment conditional on employment status at the survey.

Hours also depend on $wage^{lat}_{it}$ and E_{it}. For most observations, $wage^{lat}_{it}$ is the actual wage. However, many individuals are unemployed at the survey date but work part of the year. We use $wage^{lat}_{it}$ as the measure of the wage the individual would typically receive.

[10] Industry specific and/or occupation specific human capital are not accounted for in the model and are likely to influence estimates of ρ_v more than ρ_ω given that industry and occupation changes tend to occur across employers. They would also affect the estimates of the return to seniority that we import from Altonji and Williams (2005). See Neal (1995), Parent (2000), and Kambourov and Manovskii (2007) for evidence on the importance of occupation-specific, industry specific capital and firm specific human capital. Extending the model to distinguish occupation and/or industry is conceptually straightforward but would require models of occupation and industry transitions and attention to measurement error. We leave this to future work.

[11] See Altonji and Paxson (1986) and Senesky (2005), who show that hours changes are much larger across jobs than within jobs for both quits and layoffs, and that one cannot account for this as a labor supply response to differences in wages, nonpecuniary job characteristics, or changes in labor supply preferences.

Because wage shocks turn out to be highly persistent and because we strongly question the standard labor supply assumption that individuals are free to adjust hours on their main job in response to short term variation in wage rates, we think of the coefficient on the latent wage as a response to a relatively permanent wage change rather than a Frish elasticity. We stick with this interpretation even though we control for μ_i in both the wage and hours equations.

Log earnings $earn_{it}$ depends on (residual) $wage_{it}^{lat}$ and $hours_{it}$. The coefficients γ_w^e and γ_h^e might differ from 1 for a number of reasons, including overtime, multiple job holding, bonuses and commissions, job mobility, and the fact that for some salaried workers the wage reflects a set work schedule but annual hours worked may vary. We also include a first order autoregressive error component e_{it} to capture some of these factors. In previous drafts of the paper we freely estimated γ_w^e and γ_h^e and obtained values close to 1 for most specifications. However, for Model A.2 and A.3 it is helpful to restrict the coefficients to be 1, which we do below.

We have not considered models with an ARCH error structure. However, the model implies that the variance of wage, hours, and earnings changes are state dependent and also depend on t. This is ecause the odds of a job change and an unemployment spell depend on TEN, ED, potential experience, and $\upsilon_{ij(t)}$ and because job changes and unemployment spells are associated with innovations in $\upsilon_{ij(t)}$ and $\xi_{ij(t)}$. The variances also depend on the permanent components of X_{it} (education and race) and on the unobserved heterogeneity components μ_i and η_i.

Many studies of the income process simply ignore the presence of measurement error even though surveys by Bound et al.(2001) and others indicate that it is substantial. Altonji et al.(2002) and some other studies have attempted to directly estimate the variances of measurement error in wages, hours, and earnings under a classical measurement error assumption. Here, we draw loosely upon studies of measurement error in the PSID and other panel data sets to come up with a range of estimates of the measurement error parameters. For most of our models our choices imply that m_{it}^w accounts for 35% of $var(\Delta wage_{it}^*)$, 25% of $var(\Delta hours_{it}^*)$, and 25% of $var(\Delta earn_{it}^*)$. We abstract from measurement error in employment, which we believe is relatively unimportant, as well as in the job change indicator, which we suspect is more serious. (See Altonji and Williams (1998)). Our reported standard errors do not account for uncertainty about the measurement error parameters.[12]

2.3 Model B

The main differences between Model A and Model B are in the wage and JC equations. The wage equation for Model B is

$$
\begin{aligned}
(12)\quad wage_{it}^{lat} &= \gamma_0^w + \gamma_X^w X_{it} + \delta_\mu^\omega \mu_i + \omega_{it} \\
\omega_{it} &= \rho_\omega [1 + \phi_1 S_{it}] \omega_{i,t-1} + \gamma_{JC}^\omega JC_{it} + \gamma_{1-E_t}^\omega (1 - E_{it}) + \gamma_{1-E_{t-1}}^\omega (1 - E_{i,t-1}) \\
&\quad + [1 + \phi_2 S_{it}] \varepsilon_{it}^\omega \\
&\quad \omega_{i1} \tilde{} N(0, \sigma_{\omega_{1g}}^2). \quad \text{(initial condition)}
\end{aligned}
$$

The above wage model does not include the job specific wage component υ or the job specific hours component ξ but introduces the coefficients ϕ_1 and ϕ_2.[13] These allow the degree of persistence and the variance in the wage innovation to shift with a job change or end of an unemployment spell. As noted above, our specification of state dependence in wages captures the fact that many employers use past wage rates, along with other information, in determining wage offers for new hires, as well as the fact that previous wage rates are a reference point for incumbent workers when evaluating an offer. It may also reflect dependence between the productivity of a worker today and productivity last year. One might expect the degree of dependence to be weaker across jobs than within jobs ($\phi_1 < 0$).

The JC equation is the same as (2) with the υ terms excluded. We have estimated versions of Model B with and without $wage_{i,t-1}$ in the EE and JC equations but to save space report the version without the $wage_{i,t-1}$ terms.

Model B has proven to be easier to estimate than Model A. It is also more tractable than Model A for use in dynamic programming models of consumer behavior.[14]

[12] The assumption of normally distributed, classical measurement error runs counter to evidence that actual reports are a mixture of correct responses and responses with error. Furthermore, Bound et al (2001) summarize evidence that measurement error is mean reverting to some extent, with individuals smoothing shocks when they report on economic variables. In principle, our methods can accommodate almost any measurement error assumption. We stick with the simpler formulation for lack of hard quantitative evidence on richer measurement error specifications that we can import into our model.

[13] In early work we experimented with allowing the effect of μ_i to grow linearly with experience, but did not obtain sensible results.

[14] Vidangos (2007) has simulated a model of optimal lifetime consumption using a family income process that embeds a version of model B as the wage process for the household head.

3 Data

We use the 1975-1997 waves of the PSID to assemble data that refers to the calendar years 1975-1996. Because some observations are lost due to the use of lags, the current values of the variables in our model range from 1978 to 1996. We include members of both the SRC stratified random sample and SEO sample. The latter consists primarily of households that were low income in 1968 and substantially overrepresents blacks. We also present results for the SRC sample only. We also include nonsample members who married PSID sample members. The sample is restricted to male household heads. We include both single and married individuals.

Observations for a given person-year are used if the person is between age 18 and 62, was working, temporarily laid off or unemployed in a given year, was not self employed, had valid data on education ($EDUC$) and had no more than 40 years of potential experience. We treat persons on temporary layoff at the survey date as employed. We eliminate a small number of observations in which the individual reports being retired, disabled, a housewife, a student, other, or "don't know".. (See Appendix tables A1 and A2).[15]

Potential experience t_i is $age_{it} - max(EDUC_i, 10) - 5$. $BLACK\ _i$ is one if the individual is black and 0 otherwise. ED_{it} is the number of years in a row that a person is employed at the survey date. In 1975 and for persons who join the sample after 1975, we set ED_{it} to tenure with the current employer.[16]

The variable $UD_{i,t-1}$ is the number of consecutive years up to $t-1$ that the individual has not been employed at the survey date. We set $UD_{i,t-1}$ to 0 if the first time we observe i is in year t. Few unemployment spells exceed 1 year, so the error is probably small. The wage measure $wage^*_{it}$ is the reported hourly wage rate at the time of the survey. It is only available for persons who are employed or on temporary layoff.[17]

[15] We allow persons to come out of retirement and include future observations following a retirement spell if the individual is working, temporarily laid off or unemployed. As reported in Appendix Table A1, 1.85% of the PSID sample reports an employment status as disabled in a given year.

[16] An alternative would be to apply exactly the same censoring that occurs in the PSID in the simulated data. In the simulated data ED_{it} would be set to tenure when t in the simulated case is equal to t for the first value we see in the corresponding PSID case.

[17] This measure is the log of the reported hourly wage at the survey date for persons paid by the hour and is based on the salary per week, per month, or per year reported by salary workers. It is unavailable prior to 1970 and is limited to hourly workers prior to 1976. We account for the fact that it is capped at $9.98 per hour prior to 1978 by replacing capped values for the years 1975-1977 with predicted values constructed by Altonji and Williams (2005). They are based on a regression of the log of the reported wage on a constant and the log of annual earnings divided by annual hours using the sample of individuals in 1978 for whom the reported wage exceeds $9.98.

Finally, we censor reported hours at 4000, add 200 to reported hours before taking logs to reduce the impact of very low values of hours on the variation in the logarithm, and censor observed earnings and observed wage rates (in levels, not logs) to increase by no more than 500% and decrease to no less than 20% of their lagged values. We also censor wages to be no less than $3.50 in year 2000 dollars.

After observations are lost due to construction of lagged values, or missing data, we use information on 4,632 individuals. Each individual contributes between 1 and 19 observations. The 5th, 25th, median, 75th, and 95th percentile values of the number of observations a given individual contributes are 1, 3, 6, 11, and 18 respectively (see Appendix Table A3). Of course, persons who are present for many years contributed disproportionately to the total of 33,933 person-year observations. The number of observations per year varies from 1,200 in 1979 to 2,007 in 1991.

The sample is highly unbalanced. As we have already noted, an advantage of simulation based estimators such as I-I is that by incorporating the sample selection process into the simulation, one can handle unbalanced data. We assume that observations are missing at random, although there is reason to believe that the heterogeneity components and shocks influence attrition from the sample. In principle, one could augment the model with an attrition equation. Alternatively, it would be straightforward to simply use sample weights to reweight the PSID when evaluating the likelihood function of the auxiliary model if suitable weights were available. However, PSID sample weights are designed to keep the data representative of successive cross sections of the US population that originate in the families present in the base year.[18] They do not adjust for factors that alter the US population, such as differences in birth rates by race or education. Furthermore, there are no sample weights for persons who move into PSID households through marriage. Consequently, we do not use weights. In essence, we are assuming both that observations are missing at random and that the model parameters do not vary across demographic groups or over time. The results are fairly robust to restricting the analysis to the SRC sample, as we discuss in Section 6. We also report separate results for SRC whites by education level.

In Table 1a we present the mean, standard deviation, minimum and maximum of the variables used in our structural model. The mean of E_{it} is .97, so we observe relatively few

[18] See Fitzgerald, Gottschalk and Moffit (1998) for an analysis of attrition in the PSID. They conclude that at least through 1989 the PSID is fairly representative of the US population once internal sample weights are used.

unemployment spells. Note also that the mean of EE_{it} is .98. Given these magnitudes, even relatively large movements in the latent variable index determining EE_{it} have only a small effect on whether EE_{it} is 1 or 0. In Table 1b we provide additional information about our sample, including the mean and standard deviation for education, race, potential experience, and the calendar year.

4 Estimation Methodology

We first provide a brief overview of our estimation procedure. We then define the auxiliary model used in the estimation procedure as well as additional moment conditions that we use. Finally we discuss the values of the measurement error parameters that we impose, which are based upon other studies. Note first that to reduce computational complexity, we estimate the coefficients on X in equations with continuous dependent variables by first regressing $hours^*_{it}$, $wage^*_{it}$, and $earn^*_{it}$ on the vector X_{it}. X_{it} consists of a constant, years of education, $BLACK_i$, t_i, t_i^2, t_i^3, and a set of year dummies. We then work with the residuals of these variables when estimating the remaining parameters by I-I.[19]

4.1 Indirect inference

For clarity, we will refer to Model A (or B) above as the "structural" model even though the models do not express the parameters of the decision rules for EE, UE, JC, etc., in terms of preference parameters and parameters governing job search, mobility, and exogenous layoffs. We estimate most of the parameters of the structural model by indirect inference (I-I). The basic idea of I-I is to view both the observed data and the simulated data (generated from the structural model given a set of k structural parameters β) through the "lens" of a descriptive statistical (or auxiliary) model characterized by a set of p auxiliary parameters θ. The $k \leq p$ structural parameters β are then chosen so as to make the observed data and the simulated data look similar when viewed through this lens.

[19]Note that we include the constants γ_0^w, γ_0^h, and γ_0^e in the wage, hours and earnings models even though we also include a constant when construct hours, wage, and earnings residuals. Our reported standard errors do not account for first stage estimation of the γ_X parameters. We doubt, however, that adjustment would make much difference because the estimated standard errors for the elements of γ_X are small. Also, note that the coefficients on the experience profile of wages capture not only the effects of general human capital and age but also average growth in $\upsilon_{j(t)}$ and tenure with t. When we estimate Model A.3 by indirect inference, we include a quadratic in experience in the equation for the wage to account for this. We did not include these terms in A.1, A.2 and B.1, which exclude tenure from the wage equation.

To illustrate the ideas underlying I-I, let the observed data consist of a set of observations on N individuals in each of T time periods: $\{y_{it}\}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$. In addition, there is a corresponding set of exogenous variables $\{x_{it}\}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$. For the moment, suppose that there is no missing data: each individual's history is complete. Finally, assume that the observed data is generated by a structural model characterized by a set of structural parameters β_0.

The auxiliary model can be estimated using the observed data to obtain parameter estimates $\hat{\theta}$. Formally, $\hat{\theta}$ solves:

$$\hat{\theta} = \arg\max_{\theta} \mathcal{L}(y; x, \theta),$$

where $\mathcal{L}(y; x, \theta)$ is the likelihood function associated with the auxiliary model, $y \equiv \{y_{it}\}$ and $x \equiv \{x_{it}\}$.

Given x and structural parameters β, the structural model can be used to generate M statistically independent simulated data sets $\{\tilde{y}_{it}^m(\beta)\}$, $m = 1, \ldots, M$. Each of the M simulated data sets is constructed using the same set of observed exogenous variables x. The auxiliary model can then be estimated using each of the simulated data sets to obtain M estimated parameter vectors $\tilde{\theta}_m(\beta)$. Formally, $\tilde{\theta}_m(\beta)$ solves

$$\tilde{\theta}_m(\beta) = \arg\max_{\theta} \mathcal{L}(\tilde{y}_m(\beta); x, \theta),$$

where the likelihood function associated with the auxiliary model is, in this case, evaluated using the mth simulated data set $\tilde{y}_m(\beta) \equiv \{\tilde{y}_{it}^m(\beta)\}$. Denote the average of the estimated parameter vectors by $\tilde{\theta}(\beta) \equiv M^{-1} \sum_{m=1}^{M} \tilde{\theta}_m(\beta)$. As the observed sample size N grows large (holding M and T fixed), $\tilde{\theta}(\beta)$ converges to a nonstochastic function $h(\beta)$.

Loosely speaking, I-I generates an estimate $\hat{\beta}$ of the structural parameters by choosing β so as to make $\hat{\theta}$ and $\tilde{\theta}(\beta)$ as close as possible.[20] As the observed sample size N grows large, $\hat{\theta}$ and $\tilde{\theta}(\beta_0)$ both converge to the same "pseudo" true value $\theta_0 = h(\beta_0)$.

Implementing I-I requires the choice of a formal metric for measuring the "distance" between $\hat{\theta}$ and $\tilde{\theta}(\beta)$. As described in Keane and Smith (2003) and elsewhere, there are (at least) three possible ways to specify such a metric. Here we choose the structural parameters so as to minimize the difference between the constrained and unconstrained values of this likelihood, evaluated using the observed data. In particular, we calculate:

$$\hat{\beta} = \arg\min_{\beta} [\mathcal{L}(y; x, \hat{\theta}) - \mathcal{L}(y; x, \tilde{\theta}(\beta))].$$

[20] When generating simulated data sets, the seed in the pseudorandom number generator is fixed so that the draws of $\{\tilde{\eta}_{it}\}$ are the same for different values of β.

Gourieroux, Monfort, and Renault (1993) show that $\hat{\beta}$ is a consistent and asymptotically normal estimate of the true parameter vector β_0.

Accommodating missing data in I-I is straightforward: after generating a complete set of simulated data, one simply omits observations in the same way in which they are omitted in the observed data. As we have already discussed, we assume that the pattern of missing data is exogenous. In the simulated data, we simply omit observations according to the same pattern. In some cases, it is convenient to estimate auxiliary models in which missing observations are replaced with some arbitrary value (such as 0). In such circumstances, the same principle applies: use the same arbitrary values in both the simulated and observed data sets.

In our structural model, the observed data y consists of both continuous and discrete random variables. Discrete random variables complicate the calculation of $\hat{\beta}$ because the objective surface (i.e., the difference between the constrained and unconstrained values of the likelihood) is discontinuous in the structural parameters β. Discontinuities arise when applying I-I to discrete choice models because any simulated choice $\tilde{y}_{it}^m(\beta)$ is discontinuous in β (holding fixed the set of random draws used to generate simulated data from the structural model). Consequently, the estimated set of auxiliary parameters $\tilde{\theta}(\beta)$ is discontinuous in β. The non-differentiability of the objective function in the presence of discrete variables prevents the use of gradient-based numerical optimization algorithms to maximize the objective function and requires instead the use of much slower algorithms such as simulated annealing or the simplex method.

To circumvent these difficulties, we use Keane and Smith's (2003) modification to I-I, which they call generalized indirect inference. Suppose that the simulated value of $\tilde{y}_{it}^m$ equals 1 if a simulated latent utility $\tilde{u}_{it}^m(\beta)$ is positive and equals 0 otherwise. Rather than use $\tilde{y}_{it}^m(\beta)$ when computing $\tilde{\theta}(\beta)$, Keane and Smith propose instead to use a continuous function $g(\tilde{u}_{it}^m(\beta);\lambda)$ of the latent utility. The function g is chosen so that as the smoothing parameter λ goes to 0, $g(\tilde{u}_{it}^m(\beta);\lambda)$ converges to $\tilde{y}_{it}^m(\beta)$. Letting λ go to 0 at the same time that the observed sample size goes to infinity ensures that $\tilde{\theta}(\beta_0)$ converges to θ_0, thereby delivering consistency of the I-I estimator of β_0.

Although many functional forms could be chosen for g, here we define g as

$$g(\tilde{u}_{it}^m(\beta);\lambda) = \frac{\exp(\tilde{u}_{it}^m(\beta)/\lambda)}{1+\exp(\tilde{u}_{it}^m(\beta)/\lambda)}.$$

Because the latent utility is a continuous and smooth function of the structural parameters

β, g is a smooth function of β. Moreover, as λ goes to 0, g goes to 1 if the latent utility is positive and to 0 otherwise.

When the structural model contains additional variables that depend on current and lagged values of indicator variables $\tilde{y}_{it}^m$, these additional variables will also be discontinuous in β. In our structural model, for instance, variables such as employment duration and job tenure depend on the history of indicator variables such as employment status and job changes. Since employment duration and tenure are discontinuous in β, they also contribute to creating a discontinuous objective function in the estimation process. Our smoothing strategy, however, ensures that all these variables will also be continuous in β, provided that they depend continuously on $\tilde{y}_{it}^m$. In other words, replacing the indicator functions by their continuous approximations $g(\tilde{u}_{it}^m(\beta); \lambda)$ ensures that all other variables that depend on β through $g(\tilde{u}_{it}^m(\beta); \lambda)$ be continuous. Care must be taken in choosing λ, because approximation error in indicator functions for a particular year accumulate in the approximate functions for employment duration and tenure.

We searched for a combination of the smoothing parameter λ and the number of simulations M that generates sufficient smoothness in the objective function, while keeping bias small and computation time manageable. The larger these parameters are, the smoother the objective function will be, but large values of λ introduce bias and large values of M increase computation time. Based upon simulation experiments, we chose a small value of λ, .05, which is large enough to smooth the objective surface sufficiently given our choice of 20 for M. Our simulation experiments as well as the parametric bootstrap results reported below indicate that the associated bias in the estimates is small for almost all of our parameters.

We use a parametric bootstrapping procedure to conduct inference. Given consistent estimates of the structural parameters, we repeatedly generate "artificial" observed data sets from the structural model, estimate the parameters of the structural model for each such data set, and then calculate the standard deviations of the parameter estimates across the data sets. These standard deviations serve as our estimates of the standard errors of the structural parameter estimates associated with the actual observed data.[21] Standard errors of functions of model parameters, such as the impulse response functions and variance decompositions are constructed as the standard deviation across parametric bootstrap replications.

[21] As a check, we also computed standard errors using a nonparametric bootstrap procedure based on resampling from the PSID for some specifications. We used 100 replications and obtained similar results.

4.2 The Auxiliary Model

Our auxiliary model consists of a system of seemingly unrelated regressions (SUR) with 7 equations and 25 covariates that are common to all 7 equations. We implement the model under the assumption that the errors follow a multivariate normal distribution with an unrestricted covariance matrix. One may write the system as

$$Y_{it} = Z_{it}\Pi + u_{it}; \quad u_{it} \sim N(0,\Sigma); \;\; u_{it} \text{ iid over } i \text{ and } t, \tag{13}$$

where

$$Y_{it} = [E_{it} \cdot E_{i,t-1}, E_{it} \cdot (1 - E_{i,t-1}), JC_{it} \cdot E_{it} \cdot E_{i,t-1}, wage^*_{it}, hours^*_{it}, earn^*_{it}, \ln(1 + wage^{*\,2}_{it})]';$$

and

$$\begin{aligned} Z_{it} = [&Const, (t_i - 1), (t_i - 1)^2, BLACK_i, EDUC_i, ED_{i,t-1}, UD_{i,t-1}, TEN_{i,t-1}, \\ &E_{i,t-1} \cdot E_{i,t-2}, E_{i,t-2} \cdot E_{i,t-3}, E_{i,t-1} \cdot (1 - E_{i,t-2}), E_{i,t-2} \cdot (1 - E_{i,t-3}), \\ &JC_{i,t-1} \cdot E_{i,t-1} \cdot E_{i,t-2}, JC_{i,t-2} \cdot E_{i,t-2} \cdot E_{i,t-3}, \\ &wage^*_{i,t-1}, wage^*_{i,t-2}, hours^*_{i,t-1}, hours^*_{i,t-2}, earn^*_{i,t-1}, earn^*_{i,t-2}, \\ &wage^*_{i,t-1} \cdot (t_i - 1), wage^*_{i,t-1} \cdot (t_i - 1)^2, wage^*_{i,t-1} \cdot JC_{it}, \; wage^*_{i,t-2} \cdot JC_{i,t-1}, wage^*_{i,t-2} \cdot E_{i,t-1}]' \end{aligned} \tag{14}$$

Since Π has 25×7 elements and Σ is a 7×7 covariance matrix with 28 unique elements, the auxiliary model has 203 parameters. In contrast, Model A.3 has only 46 parameters that we estimate by I-I (not counting the measurement error parameters, tenure coefficients, and ρ_ω). As we discuss momentarily, a few of the 46 parameters are estimated all or in part using additional moment conditions rather than exclusively by I-I. Consequently, the number of features of the data used to fit the structural model greatly exceeds the number of parameters.

In estimating the model we use the likelihood function that corresponds to (13). Note that the assumption $u_{it} \sim N(0, \Sigma)$ with u_{it} iid over i and t is false for several reasons, including the fact that Y contains binary variables and that both $wage^*_{it}$ and $\ln(1 + wage^{*\,2}_{it})$ appear.

However, the fact that we use a misspecified likelihood affects the efficiency of our procedure rather than consistency.

Our choice of what to include in the auxilliary model is as much a matter of art as science, but is motivated by the following principles. First, we use a common set of right-hand-side variables in the seven equations of the auxiliary model to avoid having to iterate between Π and Σ to maximize the likelihood function. The disadvantage, however, is that we do not tailor the right-hand-side variables to the particular dependent variable. As a result, the auxiliary model probably contains more parameters than are needed to describe the data. Furthermore, we are restricted in our ability to add additional right-hand-side variables to particular equations, such as additional interactions between $(t_i - 1)$ and other lagged variables, because the total number of variables would get out of hand. Although it would be useful to explore differentiating the equations of the auxiliary model in future work, our simulations indicate that most of our parameters are quite well determined by the auxiliary model that we have chosen.

The second principle is to include variables that appear as explanatory variables in the structural model. This accounts for the presence of $(t_i - 1)$, $(t_i - 1)^2$, $EDUC_i, BLACK_i$, $ED_{i,t-1}$,and $TEN_{i,t-1}$. We also include $UD_{i,t-1}$ even though we constrain γ_{it}^{UE} to equal 0. Since the model is dynamic and includes state dependence terms in most equations, we include two lags of each dependent variable except $\ln(1 + wage_{it}^{*\,2})$. The lags help distinguish between state dependence and heterogeneity. Finally, we include the interaction terms $wage^*_{i,t-1} \cdot (t_i - 1)$, $wage^*_{i,t-1} \cdot (t_i - 1)^2$, $wage^*_{i,t-1} \cdot JC_{it}$, $wage^*_{i,t-2} \cdot JC_{i,t-1}$, and $wage^*_{i,t-2} \cdot E_{i,t-1}$ to help capture any nonstationarity in wages.

One disadvantage of our choice for (14) is that the first three observations for each individual are lost due to lags. This makes it difficult to identify parameters of the models for the initial wage and employment status. It also makes it difficult to identify changes with experience in the variance of shocks at the beginning of a career. In principle, one could add additional equations with 0, 1, or 2 lags to the auxiliary model to accommodate observations with missing data. The cost would be a more complex auxiliary model. (Alternatively, one can set values of missing lags to 0 in both the simulated and actual data.) In simulation experiments we did not find that adding such equations helped a great deal with identification of model parameters. In the next section, we discuss use of the variance of wages when $t \leq 5$ for each race/education group to help identify parameters of the initial condition for wages. We also discuss the initial condition for employment.

4.3 Use of Additional Moments and Other Information Sources to Identify Parameters

In the equation for initial employment we estimate the intercepts b_{0g} as $\hat{b}_{0g} = \hat{b}^*_{0g} \cdot \hat{\sigma}_{E1}$ where $\hat{\sigma}_{E1} = \sqrt{(\hat{\delta}^{EE}_{\mu})^2 + (\hat{\delta}^{EE}_{\eta})^2 + 1}$ and $\hat{b}^*_{0g}$ is the coefficient estimate from a Probit regression of E_{it} on a constant estimated on PSID data for $t \leq 5$ for each of 4 groups g defined by race and whether the person has more than a high school education. We use the first five years rather than simply the first because we have relatively few observations for each group when $t = 1$.[22]

To identify $\sigma_{\omega 1}$, we use the fact that model A implies that the variance of the observed (residual) wage, $wres^*_{i1}$, of an employed individual from race/education group g is

$$Var(wres^*_{i1}; g) \equiv Var(wage^*_{i1} - \gamma^w_0 - \gamma^w_X X_{i1}) = \sigma^2_{\upsilon 1} + (\delta^w_{\mu})^2 + \sigma^2_{\omega 1g} + \sigma^2_{mw}.$$

Because of sample size considerations we estimate $Var(wres^*_{i1}; g)$ as the variance of (residual) wage observations in the PSID corresponding to $t \leq 5$. $\widehat{Var}(wres^*_{i1})$ equals .109 for blacks with a high school degree or less, .124 for blacks with more than a high school degree, .109 for whites with high school or less, and .141 for whites with more than high school. We then obtain $\sigma_{\omega 1g}$ by incorporating the moment condition $\sigma^2_{\omega 1g} = \widehat{Var}(wres^*_{i1}; g) - (\delta^w_{\mu})^2 - \sigma^2_{\upsilon 1} - \hat{\sigma}^2_{mw}$ into the indirect inference procedure, where $\hat{\sigma}^2_{mw}$ is preset to $.0795^2$ on the basis of measurement error studies for the PSID. In Model B $\sigma^2_{\upsilon 1}$ does not appear. Estimates of other parameters are not very sensitive to constraining $\sigma^2_{\omega 1g}$ to be the same for all groups and estimating $\sigma^2_{\omega 1g}$ using $t \leq 3$ rather than 5.

In the case of model A, we also use a large number of moment conditions spanning a much longer time span than the 3 lags in our auxiliary model to identify ρ_{ω} and help distinguish persistence due to ω_{it} from persistence due to μ_i and $\upsilon_{ij(t)}$. For workers who are continuously employed between $t - g$ and $t + j$ and who do not change jobs between t and $t + j$,

$$cov(wres^*_{i,t+j} - wres^*_{it}, wres^*_{i,t-g}) = (\rho^{j+g}_{\omega} - \rho^g_{\omega}) var(\omega_{t-g} | t - g). \quad (15)$$

[22] As it turns out, $\hat{b}^*_{0g}$ is 0.918 for blacks with a high school degree or less, 1.427 for blacks with more than high school, 1.391 for whites with high school or less, and 1.671 for whites with more than high school. We obtained similar results for other model parameters when we constrain $\hat{b}^*_{0g}$ to be the same for all groups and use $t \leq 3$ to estimate it.

We approximate $var(\omega_{t-g}|t-g)$,with a constant plus a third order polynomial in $t-g$. We compute $co\hat{v}(wres^*_{i,t+j}-wres^*_{it}, wres^*_{i,t-g})$ for each $t+j, g$ combination satisfying $1 \leq j \leq j_{\max}$ and $1 < g \leq g_{\max}$. We estimate ρ_ω and the parameters of the polynomial approximation by weighted minimum distance using the number of sample observations used to estimate $co\hat{v}(wres^*_{it+j} - wres^*_{it}, wres^*_{i,t-g})$ as the weights, eliminating moments estimated using fewer than 5 observations. The point estimates and approximate standard errors for $j_{\max}, g_{\max} = 5$, $j_{\max}, g_{\max} = 6$, $j_{\max}, g_{\max} = 7$, $j_{\max}, g_{\max} = 8$, and $j_{\max}, g_{\max} = 9$ are .882 (.027), .919 (.020), .927 (.014), .94 (.011), and .94 (.009) respectively. We chose .92 as our point estimate.[23] For Model A.1, we obtain .913 when we ignore (15) and freely estimate ρ_ω by I-I.

In model B we make use of an expression for the difference in the variance of wage growth conditional on $JC_{it} + E_{it}(1 - E_{i,t-1}) = 1$ and conditional on $JC_{it} + E_{it}(1 - E_{i,t-1}) = 0$ to express the wage innovation variance shift factor ϕ_2 in terms of other parameters of the model. See Appendix A2. We do not use moment conditions analogous to (15) to estimate ρ_ω, although the estimate we obtain, .95, is in the neighborhood of the above values.

Finally, when we include tenure in the wage equation (model A.3 in Table 2), we impose Altonji and Williams' (2005) estimates of tenure polynomial $P(TEN_{it})$ based on PSID data for the years 1975-2001 rather than attempting to estimate it by I-I, which would have required the addition of a number of additional variables to the auxiliary model.[24]

The parametric bootstrap standard errors reported below do not account for sampling error in the above sample moments of the PSID or in the tenure parameters.

[23]The number of moments varies from 2,789 when $j_{\max}$ and $g_{\max}$ are 5 to 7,906 when $j_{\max}$ and $g_{\max}$ are 9. The standard errors account for heteroskedasticity but not for correlation among the moments, which use overlapping data. They are probably understated. In the case of Model A.3, equation (15) is an approximation because Model A.3 states that employment transition probabilities depend on ω_t. This means that the evolution of ω_t depends on the number of periods of continous employment: $j + g$. We doubt that this is an important problem. Setting ρ_ω to .90 instead of .92 has little effect on the variance decompositions.

[24]The profile that we use corresponds to Table 6, Panel D, column 2 of their paper. It is $.0272563 * Ten - .0023283 * Ten^2 + .00815 Ten^3/100$
$-.000914 Ten^4/1000$. The implied return to 2, 5, 10 and 20 years of tenure are .046 (.0064), .008 (.0011), .112 (.016), and .119 (.029). It is obtained using an Altonji and Shakotko's (1987) instrumental variables approach, which treats t as exogenous and uses the within job j variation in Ten_{ijt}, Ten^2_{ijt}, Ten^3_{ijt}, and Ten^4_{ijt} to identify the effects of tenure. Our finding of a modest link between t and $v_{ij(t)}$ implies that Altonji and Williams' estimates are biased downward by a small amount.

4.4 Mechanics of Estimation

Our chosen values of $\lambda = 0.05$ and $M = 20$ yield a smooth objective function that allows the use of fast gradient-based optimization algorithms with little evidence of bias.[25] Not surprisingly given the size and complexity of our models, the objective function displays multiple local optima with respect to some of the parameters. We experimented extensively with different starting values to make sure that we are finding the global optimum. We began the process by obtaining estimates of a series of "reduced-form equations" that correspond to the equations in the structural model. As part of the process, we have used grid evaluations for the set of parameters that appear most problematic, and we have used the smaller models to help us find good initial guesses and then build up to more complex ones. The problem is more serious in the case of Model A.1-Model A.3 than for Model B. This is one of the reasons that we make use of the moments (15) to help distinguish ρ_ω from ρ_υ.

The fact that we have between 39 and 46 parameters estimated by indirect inference, the large size of the auxiliary model, and the number of simulations make computation very time-consuming even though we use a fast gradient based optimization algorithm. To reduce estimation time, we exploit the highly parallelizable structure of our estimation methodology.[26]

4.5 Local Identification and Analysis of Estimation Bias

We have chosen the auxiliary model with an eye toward distinguishing among state dependence, fixed heterogeneity, and transitory shocks and and eye toward establishing links across equations in the heterogeneity components. However, one cannot easily verify that the parameters of our model are identified by matching up the parameters against sample moments. In particular, the fact that the number of moments that play a role in the likelihood function of the auxiliary model is much larger than the number of structural model

[25] We use a standard quasi-Newton algorithm with line search, which can additionally handle simple bounds on the choice variables. The algorithm approximates the (inverse) Hessian by the BFGS formula, and uses an active set strategy to account for the bounds. Gradients are computed by finite differences.

[26] Specifically, for any given value of the structural parameters, the M simulations required to evaluate the objective function are essentially independent and can be conducted simultaneously by k different processors. Using our parallelized computer algorithm on $k \leq M + 1$ (balanced) processors reduces computation time by a factor of about $\frac{\lceil \frac{M}{k-1} \rceil}{M}$, where $\lceil\ \rceil$ is the ceiling function. All programs are written in FORTAN 90. The parallelization is implemented using the Message Passing Interface (MPI). We estimate the model using 21 CPUs supplied by 11 Intel Xeon 5150 dual core processors, which reduces estimation time by a factor of 20, to between 2 to 7 hours.

parameters does not establish identification of any particular parameter. Consequently, we use Monte Carlo experiments extensively to establish local identification and analyze the adequacy of our auxiliary model given the sample size and demographic structure of the available data and to check for bias. For a hypothesized vector of parameter values, we simulate data and then verify that the parameter values that maximize the likelihood function of the auxiliary model are close to the hypothesized values. Using a number of model specifications, including ones that differ somewhat from the ones presented in the paper, we informally experimented with varying parameter values to get a sense of how robust identification is to the particular values. We also used these experiments to investigate whether the objective function has flat regions near the solution, or multiple global optima.

In general we have found that identification of most of the parameters is quite robust. However, our Monte Carlo studies also indicate that a few of the parameters are poorly determined given the sample size. We also found local optima involving alternative combinations of subsets of the parameters. Bringing in additional information through the moment conditions described above solved the most serious problems. However, some of the parameters remain sensitive to changes in the auxiliary model, and starting values must be chosen carefully. This is particularly true of the coefficients of the experience profiles in the EE, UE and to a lesser extent, the JC equations, for reasons that we not fully understand. Overall, however, the relatively small values of the bootstrap standard errors in the tables below indicates that for the sample size and demographic structure of the PSID sample, our auxiliary model is quite informative about most of the model parameters. Furthermore, in almost all cases the means of the bootstrap replications are close to the point estimates, indicating that the degree of bias in the procedure is small for most of our parameters.

5 Empirical Results

First, we discuss the parameter estimates for Model A, with heavy emphasis on Model A.3. Model A.3 is basically the model presented above with the duration dependence term in the UE equation restricted to 0 and δ^{v}_{JC} set to 0. We also present simpler versions of Model A and briefly discuss results for Model B, with emphasis on the wage equation. Second, we evaluate the fit of the model by comparing means and standard deviations of the PSID data to the corresponding values based on simulated data from the model and by comparing simple regression relationships in actual and simulated data. Third, we present impulse response functions. Finally, we decompose the variance of wages, hours, and earnings into the contributions of the main types of shocks in our model. In the case of Model A.3, inference is based on 300 bootstrap replications. Because the bootstrap procedure is very computer intensive, we only use 100 replications for the other models.

5.1 Parameter Estimates for Model A

Columns 3a, 3b, and 3c of Table 2 report parameter estimates, the means of the parametric bootstrap estimates, and standard error estimates for Model A.3. Columns 2a-2c refer to Model A.2, which is identical to Model A.3 except that we exclude tenure terms from the wage equation. Columns 1a-1c refer to Model A.1, which in addition excludes $wage'_{it}$ from EE and $\upsilon_{ij(t-1)}$ and $\upsilon'_{ij'(t)}$ from JC, and allows δ^{υ}_{JC} to be nonzero. The row headings indicate the variable or error component that parameter estimates correspond to and also list the parameter names. The estimates are grouped by equation, beginning with EE.

5.1.1 Employment Transitions and Job Changes

The EE coefficients on t and t^2 imply that the latent variable determining E_{it} conditional on $E_{i,t-1} = 1$ declines slowly with t until t is about 12 and then rises slowly. However, the difference between $t = 30$ and $t = 0$ in the latent variable is only .21, and the effect on the odds of a transition is small because the EE probability is very high. Since the value of $\min(ED_{t-1}, 9)$ is rising rapidly over the first few years in the labor market, the overall relationship between EE and t is weak. The point estimates should be taken with a grain of salt, because the relative values of the constant, the coefficients on t and t^2, and the coefficient on $\min(ED_{t-1}, 9)$ are sensitive to the exact specification. Furthermore, the bootstrap replications provide evidence of bias.

The coefficient on $\min(ED_{t-1}, 9)$ is .044 (.018), indicating modest positive duration dependence in the odds of remaining employed. Below we show that a regression of E_{it} on $ED_{i,t-1}$ conditional on $E_{i,t-1} = 1$ gives similar results in data simulated from Model A.3 and in PSID data, which indicates that the combined effect of duration dependence and unobserved heterogeneity in the model does a good job of matching the weak positive state dependence found in the data.

The coefficient on $wage'_{it}$ is -.058 (.076). At first glance the negative sign would seem to be opposite the expected sign because there is only a substitution effect at 0 work hours.

However, it is probably better to think of employment at the survey date as a movement along the extensive margin when one views labor supply from the perspective of a year. In any event, the coefficient is not statistically significant and the implied effect on the EE probability is small.

The UE equation is the least satisfactory equation in the model. The estimated t profile implies that the exit probability declines with experience and then increases, but the standard errors are very large. As we document below, the model does a poor job of tracking the relationship between UE and t in the data. We experimented with models that included $UD_{i,t-1}$ but had difficulty estimating the duration coefficient, perhaps because the overall number of unemployment spells is small and relatively few individuals were unemployed for two or more surveys in a row. (Most work on duration dependence in unemployment spells uses weekly, monthly or quarterly data.) Simulations reported below show that the equation without state dependence is consistent with the negative link between UE and $UD_{i,t-1}$ found in the PSID, presumably because of the important role played by permanent heterogeneity. However it understates that relationship to some extent.

The latent variable for JC declines slightly with t over the first 20 years, but is strongly decreasing in job tenure. The coefficient on $TEN_{i,t-1}$ is $-.0673$ (.0156), indicating that 10 years of seniority shift the index determining JC by .673 standard deviations of the job change shock ε_{it}^{JC}. It is noteworthy that we obtain a large negative effect of tenure on JC even after accounting for unobserved person specific heterogeneity (μ and η) and for job match heterogeneity.

The job match components $\upsilon_{ij(t-1)}$ and $\upsilon'_{ij'(t)}$ play an important role in job mobility without unemployment, and they have signs and relative magnitudes that are consistent with the theoretical discussion above. The coefficient on $\upsilon_{ij(t-1)}$ is -.923 (.127). To get a sense of the magnitude, note that the standard deviation of $\upsilon_{ij(t-1)}$ is .273 for a person with 10 years of experience. Consequently, a one standard deviation increase in $\upsilon_{ij(t-1)}$ lowers the JC index by -.252. Since the coefficient on TEN is -.067, this is roughly equivalent to the effect of 3.8 years of seniority. An increase from 0 to .273 in the value of $\upsilon_{ij(t-1)}$ lowers the probability of a job change for a white individual with 12 years of education, 2 years of seniority, 10 years of experience, $\mu = 0$ and $\eta = 0$ from .130 to .084, keeping $\upsilon'_{ij'(t)}$ constant. The current value $\upsilon'_{ij'(t)}$ has a coefficient of .594.[27] A one standard deviation shock to $\upsilon'_{ij'(t)}$

[27] The total effect of $\upsilon_{ij(t-1)}$ on the JC probability is .029. It is smaller than the partial effect because a unit shift in $\upsilon_{ij(t-1)}$ shifts the distribution of $\upsilon_{ij'(t)}$ by .625.

raises the job change probability by 0.066.

Thus far, we have not discussed the role of race and education or the unobserved individual heterogeneity terms. $BLACK$ has a substantial negative effect on the latent variable for EE and a substantial negative effect on UE, while $EDUC$ has a substantial positive effect on both. In the JC equation $BLACK$ is positive, with an effect that is equivalent to having about 2.5 fewer years of tenure. $EDUC$ does not alter JC. The coefficient on the "productivity factor" μ is .443 (.094) in the EE equation and .637 (.128) in the UE equation. Since the standard deviations of μ, η, ε_{it}^{EE}, and ε_{it}^{UE} are all 1, the factor loadings imply that μ accounts for 15.6% of the error variance in the EE equation and 27.9% in the UE equation. The productivity factor μ has a coefficient of -.280 (.136) in the JC equation. All three results are sensible in light of the fact that μ has a positive sign in both the wage and hours equations. In thinking about the magnitudes, keep in mind that the factor loadings represent the partial effect of the heterogeneity components in a given period holding spell duration constant.

The mobility/hours preference component η is normalized to have a positive sign in the JC equation. It enters the EE, UE and JC indices with coefficients of -.237 (.107), .222 (.187), and .531 (.100), respectively, and accounts for 4.5%, 3.4%, and 19.7% of the error variances of these equations. The magnitude implies only a small effect on the EE transition probability and the sign in the EE equation is positive in the other models. The results suggest that η raises the probabilities of transiting out of unemployment and of moving from job to job without unemployment. It has a coefficient that is essentially zero in the hours equation in the case of Model A.3 but is larger in the other specifications. Across Models A.1, A.2, A.3 and B.1, the relative importance of μ and η varies somewhat.

5.1.2 The Wage Model

The estimates of the wage model are very interesting. We begin with the parameters of the autoregressive component,

$$\omega_{it} = \rho_\omega \omega_{i,t-1} + \gamma_{1-E_{it}}^{\omega}(1 - E_{it}) + \gamma_{1-E_{i,t-1}}^{\omega}(1 - E_{i,t-1}) + \varepsilon_{it}^{\omega}$$

The estimated standard deviation of the initial condition ω_{i1} varies with race and education but is between .248 and .306 in Model A.3. The autoregressive coefficient $\hat{\rho}_\omega$ is .92, which implies considerable persistence but is well below unity. The shocks $\varepsilon_{it}^{\omega}$ have a standard deviation of .095 (.003). This value strikes us as large given that we separately account for the effects of job specific error components, but we do not know of any other evidence in the literature with which to compare it. In the data, the standard deviation of wage changes for stayers is .136 after adjusting for measurement error.

The coefficients of -.1895 (.010) on $(1 - E_{it})$ and .1041 (.013) on $(1 - E_{i,t-1})$ imply that being unemployed at the survey date has a large effect on the mean of the wage that persists for some time, even when the value of lost tenure is held constant. As will become clear from the impulse response functions, unemployment also leads to a loss of tenure as well as to a reduction on average in the value of the job match component, which implies further reductions in wages.[28]

The coefficient δ^w_μ on μ is only .049 (.028), so the direct contribution of unobserved permanent heterogeneity to the variance of wages is small. Note, however, that μ also has an additional effect on wages through its connection to employment transitions and job changes. One should also keep in mind that μ is net of the effect of X_{it}, which contains the important permanent variables $EDUC$ and $BLACK$, and that the standard deviation of the initial condition ω_{i1} is large. Almost 30 percent of the effect of ω_{i1} is still present at $t = 15$ The component μ is much more important in Model A.1 and in Model B.1, perhaps because there is no selective quit behavior in A.1 and B.1 does not incorporate a job specific wage component at all. It is also more important in the SRC sample.

The parameters of the job match component $\upsilon_{ij(t)}$ are also quite interesting. The initial condition $\upsilon_{ij(1)}$ has a standard deviation of .197 (.021) in A.3. The autoregression parameter ρ_υ is .625 (.022) and the value of $\hat{\sigma}_\upsilon$ is large: .269 (.007). As we have already noted, the substantial persistence of $\upsilon_{ij(t)}$ across jobs suggests that wage offers are based in part on salary history, that demand shocks may reflect narrow occupation, industry and region and thus may not be entirely job specific, and/or that the search network available to workers depends on job quality. As we shall see below, the contribution of the job specific component to the variance of wages and earnings is substantial.

As we discuss in Appendix 2, one can use Model A.3 to decompose $E(wage_{it}|t)$, the experience profile of wages, into the contributions of general human capital $P(t)$, job mobility $E(\upsilon_{ij(t)})|t)$, and accumulated job seniority $\gamma^w_{TEN}E(P(T_{it})|t)$. Figure 1 graphs the components and thus addresses the fundamental question of what accounts for wage growth over a career. Most of the return to potential experience is due to general skill accumulation.

[28]In Models A.1 and A.2, which exclude tenure from wages, $\hat{\gamma}^\omega_{1-E_{it}}$ is about -.22 and $\hat{\gamma}^\omega_{1-E_{it-1}}$ is about .10.

Job shopping and the accumulation of tenure account for 14.1 percent and 13.1 percent, respectively, of the overall growth of wages over the first ten years. They account for 13.7 percent and 15.0 percent of growth over the first 30 years.[29] Using social security records for quarterly earnings (rather than hourly wage rates) Topel and Ward (1992) find larger gains from job mobility early in careers than we find. We suspect their estimates are overstated by the school to work transition and growth across jobs in hours worked, while ours are understated because we miss some job changes and we do not use the first three years of wages.[30]

5.1.3 Hours and Earnings

In the $hours_{it}$ equation $\hat{\gamma}_E^h$, the effect of E_{it}, is .413 (.008). This indicates that unemployment at the survey date is associated with relatively long completed spells of nonemployment. Short spells will tend to be missed given our point in time measure at annual frequencies. They will show up in the hours component ε_{it}^e. The wage elasticity is small and negative, which is consistent with a large literature on the behavior of male household heads. The coefficients on μ and η are .125 and .0145 respectively, suggesting only a modest role for individual heterogeneity (net of $EDUC$ and $BLACK$) in annual hours in a given year. However, permanent heterogeneity turns out to be quite important over the lifetime. The importance of μ relative to η varies across the specifications. The standard deviation σ_h of the iid hours error is .169, indicating substantial year to year variation in hours even when the job specific component ξ does not change. The standard deviation of ξ is large—.157 (.014).

Turning to earnings, recall that the coefficients γ_w^e and γ_h^e are constrained to equal 1. The earnings component e_{it} has an autoregression coefficient of .553 (.007) and the standard

[29] Using the PSID Buchinsky et al (2008) estimate a simultaneous model of employment, job mobility, and wage rates that incorporates tenure effects, general experience, and job specific error components. They find a large effect of human capital accumulation and returns to seniority that are more than double the values from Altonji and Williams (2005) that we impose but do not present estimates of the gains from job mobility. Bagger et al (2007) do not allow for a direct effect of seniority on wages such as would arise from shared investment in firm specific capital but obtain an indirect effect that arises through the response of firms to outside offers. They attribute average growth of wages within the firm and growth of wages across firms to job search. Using Danish matched employer/employee data they find that human capital accounts for about half of all wage growth for workers with more than 12 years of education that occurs after the first five years in the labor market. Human capital accumulation is neglibible for less educated groups.

[30] Model A.1 does not allow for selective quit behavior but includes the term $\gamma_0^{\upsilon} JC_{it}$. The coefficient $\hat{\gamma}_0^{\upsilon}$ is .042, which indicates that on average JC increases υ_{ij} by .042. Models A.1 and A.2 exclude tenure effects on wages.

deviation of the shock ε^e_{it} is .211 (.002).

5.2 Estimates of Model B

Columns 4a-4c of Table 2 reports estimates for Model B.1. As we noted earlier, Model B does not contain job specific wage or hours components. However, it allows the autoregression coefficient and the standard deviation of wage shocks to shift when individuals change jobs or leave unemployment. The specification is closest to that of Model A.1. We had relatively little difficulty estimating this model and estimate the autoregressive parameter ρ_ω by indirect inference rather than by using the moment conditions (15), which do not apply to Model B.1 without modification.

The results for Model B.1, in keeping with those for Model A, provide clear evidence that job changes, whether with or without unemployment, involve substantial wage risk. The coefficient on lagged wages is .958 for job stayers but only .723 if the job changes. At the same time, wage innovations have a standard deviation of .0934 (.002) for stayers but are about 3 times larger for persons who have changed jobs ((1+2.10)*0.0934). The estimate of δ^w_μ is .151 (.012), so permanent heterogeneity plays a more important role in Model B.1 than in A.3[31]

5.3 Evaluating the fit of the Model:

We simulate careers for 138,960 individuals using the parameter estimates for Model A.1, A.2, A.3 and Model B.1—30 for each individual in the PSID sample. From each simulated career we select data so that the temporal pattern, education level, and race matches that

[31] We experimented with adding the health status equation

$$H_{it} = I[\gamma^H_0 + \gamma^H_{PE}(t-1) + \rho_H H_{i,t-1} + \delta^H_\varsigma \varsigma_i + \varepsilon^H_{it} > 0] \tag{16}$$

to models that are very similar to A.1 and to model B, where H_{it} is an indicator for poor health, and is 1 for those who answer yes to the question, "Do you have any physical or nervous condition that limits the type of work or the amount of work you can do?", ς_i captures fixed heterogeneity in health, and ε^H_{it} is an iid health shock. We added H_{it} to the EE, UE, *wage*, and *hours* equations. We also added an equation for H_{it} to the auxiliary model and added two lags of H_{it} to all equations of the auxiliary model.

The heterogeneity term ς accounts for 67% of the variance of the composite error term for the health latent variable. We also find strong state dependence in health and (not surprisingly) that health status worsens with age. The effects of health are small for EE, UE, and for wages. Work hours are about 6% lower for people in poor health, everything else equal. The relatively small effects of health, at least as we have measured it, on employment, wages, and hours imply that health shocks account for little of the variance in career earnings. We focus on models without the health equation for this reason. See Vidangos (2007) for models with health and permanent disability.

of a corresponding PSID case. Note that in all cases the simulated variables incorporate measurement error. We examine the fit of the model in two ways. First, we compare the means and standard deviations of the key variables implied by the model with corresponding values from the PSID. We then turn to a comparison of the regression relationships among key variables that are implied by the model with those of the corresponding PSID estimates.

5.3.1 Predicted and Actual Mean and Standard Deviations of Key Variables, by Potential Experience

Figure 2 compares the actual standard deviations of $wage^*_{it}$, $earn^*_{it}$ and $hours^*_{it}$ in the PSID to the 95 percent confidence interval estimates of the standard deviations based on data simulated from Model A.3. The standard error bands in the figure reflect both sampling error in $\hat{\beta}$ and sampling error due to randomness in the careers of the individuals in a particular sample.[32] The bands are tight, which means that we sometimes conclude that the PSID values are statistically different from the model predictions even when the values are close in economic terms.

Across experience levels, the model overpredicts $SD(wage^*_{it})$ by about 6% (not shown) and implies less of an increase with experience than we find in the actual data. We suspect that most of the discrepancy is due to the fact that we have removed the effects of education and race from the PSID, but include them in EE, UE, and JC in the model. Consequently, the effects of race and education on wages, hours, and earnings that operate indirectly through EE, UE, and JC rather than directly through wages and hours influence the model predictions for $SD(wage^*_{it})$ but not the PSID estimates.[33] The values for Model A.2 (not reported) are reasonably similar to A.3, while the values for Model B.1 in Appendix Figure A1 shows a flatter profile for $SD(wage^*_{it})$.

The sample value for $SD(hours^*_t)$ is .285, which lies a bit below the model value .293.

[32] We obtain the distribution of $SD(wage^*_{it})$ implied by the model as follows. First,we obtain the point estimate $\hat{S}D(wage^*_{it})$ by using the point estimates of β for Model A.3 to simulate 30 careers for each member of the PSID. (That is, we preserve the race, education and experience mix of the available data.) To obtain the standard deviation of $\hat{S}D(wage^*_{it})$ given the PSID sample size and demographic properties, we repeat the simulation for each of the 300 parametric boostrap estimates of β using only 1 career for each member of the PSID and then compute the standard deviation of the 300 values of $SD(wage^*_{it})$. The bands we report are the point estimate $S\hat{D}(wage^*_{it})$ plus or minus 1.984 $SD(SD(wage^*_{it}))$. Other variables are handled in similar fashion. For each value of t in the table the results are the average over $t-1$, t, and $t+1$ with the exception of $t = 40$,which is the average for $t = 39$ and $t = 40$.

[33] We could not think of an easy way to check this, but in earlier work excluding education and race from these equations, the model closely matched the standard deviations for $SD(wage^*_{it})$. There were still some minor discrepancies in the experience profiles.

The model implies that $SD(hours_t^*)$ varies little with t and misses the increase when t is 40, which might reflect the effects of partial retirement not captured by the experience profiles in the model. The results for Model A2 and B.1 are similar. The actual and simulated SD for earnings are 0.567 and 0.585, respectively. However, there is an erratic pattern in the data that is not matched by the model predictions, which display a smooth hump-shape pattern with a peak around $t = 20$. The error at $t = 40$ for earnings mirrors the error for hours and probably reflects partial retirement.

The left panels of Figure 3 compare the PSID values and the model predictions for the mean of E_t and for the mean of JC_t conditional on $E_t = 1$ and $E_{t-1} = 1$. The PSID values lie within or close to the 1.984 standard errors of the model estimates The overall mean for E_t is .967 in the data and .960 based on the model. Overall, the model overstates JC_t by about .014 but tracks the experience profile fairly closely.

The upper right panel of Figure 3 reports the sample means and simulated means of EE transitions, which match reasonably well. However, the lower right panel shows that A.3 does poorly in explaining the experience profile of exits from unemployment (UE). The actual and simulated means of UE are .74 and .69, a substantial discrepancy. Furthermore, the model does not track the experience profile well.[34] Model B.1 does somewhat better than model A.3 despite the fact that the EE and UE equations of A.3 and B.1 are the same (Appendix Figure A2).

Figure 4 examines the behavior of the mean of TEN, ED, and UD. The fits for TEN and UD are reasonably close, although the two models overpredict UD by an average of about 0.4 years. The models overpredict ED by a substantial amount. This is probably

[34] As was mentioned earlier, we experienced difficulty in estimating the effect of t in the UE equation, by I-I, whether or not we allow for state dependence. We also had difficulty with the EE equation. We are puzzled as to why. It may be that longer lags than we use would be helpful in pinning down the parameters. For the UE specification reported in the paper we also estimated the UE equation by maximum likelihood, treating $(\delta_\eta^{UE})^2 + (\delta_\mu^{UE})^2$ as a single parameter. Even though UE does not depend on duration, this is only an approximation. The problem stems from the fact that η and μ appear in both the EE and UE equations. Consequently, the distributions of η and μ conditional on $E_{t-1} = 0$ depend on $t-1$, $BLACK$, and $EDUC$, and the stochastic components that influence the wage. They will not have a normal distribution. In any event, the MLE estimates (standard errors) are

$$E_{it}^* = \underset{(.439)}{-.838} \quad \underset{(.029)}{.121\text{EDUC}} \quad \underset{(.109)}{-.439\text{BLACK}} \quad \underset{(.273)}{+.33t/10} \quad \underset{(.066)}{-.048t^2/100} \qquad \widehat{[(\delta_\eta^{UE})^2 + (\delta_\eta^{EE})^2]^{.5}} = .630(.119)$$

(N=1065, number of individuals contributing spells=748.) The I-I estimates of the cofficients on BLACK and EDUC are somewhat larger. The experience profiles are also quite different. The I-I estimate of $\widehat{[(\delta_\eta^{UE})^2 + (\delta_\eta^{EE})^2]^{.5}}$ is .675, which is close to the MLE estimate. We cannot take the same approach with the EE equation due to state dependence in that equation.

attributable to our use of TEN as the initial value for ED when an individual first enters the sample (see the Data Section). The results for the models A.2 and B.1 are similar.

5.3.2 Comparison of Regression Relationships Among Key Variables

Tables 3a-3d report a series of regressions. The "a" columns are based on the PSID sample, and the "b" and "c" columns are based on data simulated using the estimates of Model A.3 and Model B.1. We also report robust panel standard errors for the PSID estimates.[35] Columns 1a and 1b of Table 3a report regressions of E_t on $(t_i - 1)/10$, $(t_i - 1)^2/100$, and ED_{t-1} conditional on $E_{t-1} = 1$. This is a stripped down version of the EE equation in the structural model. There are some differences in the experience profiles. The coefficient on ED_{t-1} is .0026 in the PSID and .0033 in the simulation, a close correspondence.

Columns 2a and 2b report results for a version of the UE equation. The differences in the coefficients on the experience profile are substantial, a fact that is reflected in the failure of the model to fit the experience profile. The model understates the degree of persistence in unemployment spells to some extent. The equations for JC in columns 3a and 3b match fairly closely, although the weights on the experience terms are somewhat different.

Table 3b examines the dynamics of wages. When only one lag is included, the coefficient on $wage^*_{t-1}$ is .885 in the PSID and .905 in the simulation. When two lags are included, the sums of the coefficients are very close but there is a substantial difference in the coefficient pattern. The coefficient on JC_t is small and negative in the PSID and in the simulated data for A.3. It is small and positive for B.1.

Table 3c examines hours. The results based on the simulated data and actual data match reasonably closely, although the sum of the coefficients on the lags of hours is about .6 in the simulated data and about .55 in the actual data (columns 1a and 1b). The wage coefficient is essentially 0 in the actual data and -.0191 in the simulated data—a close correspondence.

Finally, in Table 3d we report earnings regressions. Note that all of the dynamics in earnings stem from dynamics in the wage, hours, and the autoregressive earnings component

[35] For the simulated data the point estimates are based upon a sample 30 times as large as the PSID with the same demographic structure. The coefficient standard errors reported for the simulated data are based on this large sample and are intended to provide a sense of numerical accuracy rather than sampling error. We could provide the latter by estimating the regressions on each of a series of simulated samples that match PSID demographic structure with only one career per person. One sample in the series would be created using one of the 300 values of the model parameters obtained through our parametric bootstrap procedure. The PSID standard errors should provide a rough guide to whether the coefficients based on the simulated data are different from the PSID regression coefficients.

e_{it}. The sum of the coefficients on $earn^*_{i,t-1}$ and $earn^*_{i,t-2}$ is .873 in the PSID data and .800 in simulated data, so that the model understates the persistence of earnings by a small amount. There is also some difference between the data and the model in the coefficients on $wage^*_t$ and $hours^*_t$ (Column 2a and 2b).

Overall, the match between the model and the data is good, although there is room for improvement, particularly in the case of UE.

5.4 Mean and Variance Impulse Response Functions

Figures 5a-c report impulse responses to shocks that occur when $t = 10$. The point estimates are constructed as follows. First, using the parameter estimates for Model A.3, we simulate a large number of cases up to $t = 10$. Then we impose the shock indicated in the figures in Period 10. After that, we continue the simulation in accordance with the model parameters. The figures show the mean paths of wages, hours and earnings relative to the base case. The base case represents the mean of the simulated paths in the absence of the specified intervention in period 10.[36]

Since wages and hours are reflected in earnings with coefficients of 1, we focus on earnings to save space. The diamond line in Figure 5a reports the response of the mean of $earn_{it}$ to a one standard deviation shock to $\varepsilon^{\omega}_{it}$, the error term in the autoregressive component of wages. Earnings rise by about .088, and the effect slowly decays, governed by the value .92 for ρ^{ω}_{w}. The pattern for earnings closely mirrors the response to wages because the coefficient on the wage is 1 and the response of hours to the wage is small.

The line with circles shows the effect of becoming unemployed when $t = 10$. The pattern is very interesting. The log of earnings drops by about -.62, recovers by more than two thirds after one year, and then slowly returns to the base case. The initial drop is the combination of a drop of about -.38 in log hours and a drop of about -.25 in the wage. Hours recover almost completely after one period. The wage increases by about .08 in the first year and recovers slowly after that.

The drop in wages is due to three main factors. First, the distributed lag coefficients on unemployment in the wage equation and $\hat{\rho}_{\omega}$ indicate that unemployment reduces ω_{it} by -.190 (0.010) followed by an increase a year later of .104 (0.013) plus .190*(1-.92) if the

[36] 1.984 standard error bands were obtained by computing impulse responses using each of the 300 values of the model parameters obtained by parametric bootstrap (100 values in the case of B.1). The bands are quite narrow, and we omit them to avoid cluttering the figures. They are available upon request.

person leaves unemployment. After that, the response of ω_{it} to unemployment is governed solely by $\hat{\rho}_\omega$. Second, the loss of tenure lowers the wage by an average of .064 relative to the baseline average for persons at $t = 10$. Third, since there is no selectivity in the job change induced by the unemployment spell, on average workers suffer a decline in $\upsilon_{ij(t)}$ equal to $(1-\rho_\upsilon)E(\upsilon_{ij(t)}|t=10)$, or .027. On average, endogenous mobility following the unemployment spell leads $\upsilon_{ij(t)}$ to back up toward the base case mean for a given value of t.

The pattern of a long-lasting impact of unemployment on earnings is broadly consistent with a number of previous studies, including Jacobson, Lalonde, and Sullivan (1993), who use establishment earnings records. Using the PSID and a fixed effects strategy, Stevens (1997) finds a 30% drop in earnings and a 14% drop in wages in the year of a layoff. Earnings recover substantially in the first year, but wages recover very slowly. Her estimate of the initial earnings loss is smaller than ours, perhaps because those who are laid off do not necessarily become unemployed, and those who are unemployed at the survey date tend to be in a long spell. Our model and the PSID data permit us to examine effects that operate through wages and hours separately, as well as to identify the specific channels of influence.[37]

Finally, the figures report the response of wages, hours, and earnings to an exogenous job change. In this case, JC_{it} is set equal to one in period 10 for individuals with $E_t = E_{t-1} = 1$ which one should think of as resulting from a large positive realization of the iid component ε_{it}^{JC} that negatively affects the relative attractiveness of the current job rather than a large draw of $\upsilon'_{ij'(t)}$. The line marked with "x" shows the average response. The negative effect on earnings reflects the value of lost tenure (.063). Since the job change is not selective on υ_{ij}, $\upsilon_{ij(t)}$ declines by $(1-\rho_\upsilon)E(\upsilon_{ij(t)}|t=10)$ or .027. The line with triangles is the effect of an exogenous job change that is accompanied by a value of $\varepsilon_{ij(t)}^{\upsilon}$ that is one standard deviation above its mean, or .269. The net positive effect is large and highly persistent. These results are mirrored in wages (Figure 5b). In addition, we show the effect of an exogenous job change that is accompanied by a 1 standard deviation increase of .157 in the job specific hours component $\xi_{ij(t)}$. This is associated with a positive increase in hours worked and

[37]Kletzer (1998) surveys the literature on job loss and wages. A number of studies examine how employer and industry tenure affects the size of the loss. When the problem of unobserved worker heterogeneity (but not job heterogeneity) is addressed there appear to be modest tenure effects of the loss that are consistent with Altonji and William's (2005) estimates used here. Neal (1995), Carrington (1993) and Parent (2000) argue that industry tenure is more important than firm tenure. Manovskii and Kambourov (2008) argue that occupational tenure is more important than firm or industry tenure. As we noted earlier, one could extend the model we consider to include industry and occupation transition equations, but leave this to future research.

in earnings that decays in half in the first few years but slowly thereafter. Since $\xi_{ij(t)}$ is independent across jobs, the persistence stems from the fact that when t is greater than 10, job changes with or without unemployment are infrequent.[38]

We also use the model to estimate the effects of an exogenous job loss and an exogenous job change on earnings risk using the methodology described above. The circle line in Figure 6a graphs the ratio of $var(earn_{it} - earn_{i,t-1})$ following an exogenous unemployment shock when $t = 10$ to the baseline variance for the model. The variance ratio is slightly below 1 when $t = 10$, it is 1.46 when $t = 11$, declines to 1.17 when $t = 12$, and then slowly declines to 1 over the next ten years. The pattern is mirrored in Figure 6b, which graphs the corresponding ratio for $var(earn_{it})$. The variance of earnings slowly rises following the shock, presumably because in some cases the initial job change induces additional ones. We have produced corresponding figures for shocks at $t = 3$ (not shown). The impact on the variance is somewhat smaller and less persistent.

5.5 Variance Decompositions

We have used our model to measure the relative importance of the initial condition and shocks to the autoregressive wage component, the iid hours shocks, the iid earnings shocks, job changes and employment spells and the associated shocks, the permanent heterogeneity components μ and η, and the effects of education and race. However, because the sample overrepresents blacks, we report variance decompositions using the white sample. To do this, we first compute the variance in the sum of the annual values of $earn_{it}$, $wage_{it}$, and $hours_{it}$ over a 40 year career. We then repeat the simulation after setting the variance of the particular random component in the model to 0. We use the drop in the variance relative to the base case as the estimated contribution of the particular type of shock. Since the model is nonlinear, the contributions do not sum to 100% and may be negative. We have normalized them to sum to 100. We report results for the levels of variables, accounting for the experience profile in all variables. The decompositions of the sums of the annual values of logs of earnings, hours, and wages are similar (not reported). We use the parametric bootstrap distribution of the $\hat{\beta}$ to estimate the standard deviation of variance contributions,

[38] We also computed, but do not report, the effects of shocks that occur when $t = 3$. The immediate effect of unemployment on earnings and wages is somewhat smaller than when $t = 10$ because the decline in tenure and in υ is smaller. The effects are also less persistent. Job changes accompanied by shocks to υ and to ξ also have less persistent effects.

which are reported in parentheses. We continue to focus on Model A.3, but also briefly discuss results for Model A.2 and B.1.

The results for Model A.3 are in Table 4a. The first row refers to the sum of lifetime earnings. The earnings shocks ε_{it}^{e} account for 6.6% of the variance in lifetime earnings even though they account for about 17% of $var(earn_{it})$ in a given year (Table 4b). The reason for the relatively small contribution is that the shocks are not very persistent. Similarly, the value in column II indicates that iid hours shocks ε_{it}^{h} contribute only 2.4% of the variance in lifetime earnings but account for between 7.7% and 9.3% in annual earnings (Table 4b). One can easily self-insure against these shock categories. In contrast, in column III, the initial condition $\varepsilon_{i1}^{\omega}$ and the iid shocks to ω_{it} are together responsible for 12.4% of the variance in lifetime earnings. The earnings results reflect the fact that these shocks account for 20.6% of the variance in lifetime wages. They contribute little to the variance in hours because the response of hours to wages is small.

The most striking result is in Column IV, which shows the collective impact of job specific hours and wage components, unemployment spells, and job changes. Altogether, mobility and unemployment related shocks account for 36.7%, 48.2%, and 46.8% of the variance in lifetime earnings, wages, and hours, respectively. Given the interactions among the job change and employment related factors, we break down their relative contributions by first turning off the job specific hours shocks, then turning off both hours and job specific wage shocks, then turning off hours, wage, and unemployment shocks, and finally turning off hours, wage, and unemployment and the idiosyncratic job change shocks (ε_{it}^{JC}). The estimates are reported in columns VIII, IX, X, and XI. For earnings, job specific wage shocks are more important than hours shocks. Job specific wage shocks dominate for wages, while job specific hours shocks dominate for hours.[39]

Finally, we turn to the three permanent heterogeneity components for whites: η, μ, and $EDUC$. Surprisingly, the estimates in column V indicate that the mobility preference component η does not play much of a role. The point estimate is actually negative. However, μ accounts for 11.4% of the variance in lifetime earnings and 46.2% of the variance in work hours but explains none of the variance in wages.[40] The positive direct effect that μ has on the wage variance is offset by its role in the reducing transitions into unemployment and

[39] A few of the estimated variances contributions are negative. We have verified that this reflects nonlinearity in the model. Variance in one shock can reduce the influence other shocks.

[40] Note that because the lifetime variance of log hours is lower than the lifetime variance of log wages, the 46.2% impact of μ on work hours translates into only an 11.4% impact on the variance in lifetime earnings.

job changes.. Education is very important, contributing 31.4% of the variance in lifetime earnings and 34.6% of the variance in lifetime wages but only 4.9% of the variance in lifetime hours.

The results for Model A2 (not reported) are basically similar to those for Model A3, except that μ plays a somewhat larger role in the variance of earnings and wages. The results for Model B.1 are in Tables 5a and 5b. The model does not include job specific wage or hours components. Without these features, the interpretation of the results in terms of underlying economic factors is less transparent than those for A3. However, job changes with and without unemployment are associated with reduced persistence and large innovations in ω_{it}. This is reflected in the fact that the initial condition for ω_{it} and the $\varepsilon_{it}^{\omega}$ shocks account for 24.6% of the variance in lifetime earnings and 45.59% of the variance in lifetime wages, respectively. The two unobserved heterogeneity components account for about 30% of the earnings variance, 12% of the wage variance, and 85% of the hours variance. In the case of hours, $EDUC$ also plays an important role. These factors are much less important for variance in a given year. Note that one can use the model to examine the contributions of the shocks between t and t+5 to the variance in earnings over the same period or subsequent periods, but we exclude these computations.

6 Results for Other Samples

In this section we briefly summarize results using the full SRC sample, the SRC sample of whites with some college or more, and the SRC sample of whites with a high school degree or less and no post secondary vocational education. We estimate wage, hours, and earnings residuals separately for each sample, removing the race dummies from the models for whites. We use the tenure profile from Altonji and Williams (2005) for all subsamples. In the case of Model A.3 we continue to use $\rho_{\omega} = .92$ for the full SRC sample and the SRC sample of whites. For the SRC samples of whites by educational attainment we use $\rho_{\omega} = .90$ because the evidence based on (15) pointed to a slightly lower value. Column 1b and 2b of Table 6 reports point estimates of Model A.3 and Model B.1 for the full SRC sample. Because of the computational burden, we did not compute standard errors. For ease of comparison, we report estimates for the combined SRC-SEO sample in Columns 1a and 2a.

Overall, the point estimates for the SRC sample are very similar to those for the SRC-SEO sample. The coefficient on $BLACK$ is smaller in the SRC sample, which may reflect that fact that it was SEO sample was drawn from households in low income areas. Individual heterogeneity plays a somewhat more important role in the wage equation.

Figure 7a reports the average response of earnings to shocks at t=10 and may be compared to Figure 5a for the SRC-SEO sample. The results are very similar to those for the full sample. Panel A of Table 7 reports decompositions of lifetime earnings, wages, and hours. The results are also quite similar to those for the SRC-SEO sample in Table 4a.

Columns 1d and 1e in Table 6 reports model estimates for SRC subsamples of whites with a high school degree or less and whites with some college or more. (Individuals with a high degree and some postsecondary vocational education are excluded from both samples.) For comparison, we also report estimates for the full SRC sample of whites in column 1c. The point estimates are quite similar overall. However, a few differences are worth noting. First, mobility is less sensitive to seniority for the high education sample than the low education sample. Second, JC is more responsive to outside offers in the case of the high education sample. Third, unemployment is less common for the high education sample sample. These facts are reflected in the decompositions of the experience profile of wages in Appendix Figures B1 and B2, which show little growth in υ with t for the less educated sample. We also find that σ_{ω} is considerably larger for the high education sample—.100 versus .075. The standard deviation of the iid component of hours is much larger for the less educated sample, which probably reflects greater variation in overtime hours and in unemployment spells between surveys.

The variance decompositions in Panel B and Panel C of Table 7 indicate that the persistent productivity component ω_{it} is more important for the high education sample than the low education sample for wages and earnings. Employment shocks, iid hours shocks, and within group heterogeneity in education and μ are more important for the low education sample. The job specific hours component, ω and υ are more important for the high education sample.

7 Conclusion

In this paper we study the sources of variation across individuals and over careers in earnings. To this end, we build up a model of earnings dynamics from equations governing employment transitions, job changes without unemployment, wages, and work hours. Since both state dependence and heterogeneity are important and one cannot determine the role of one without accounting for the other, our models incorporate state dependence in employment, job changes, and wages, while also including multiple sources of unobserved heterogeneity, as well as job-specific error components in both wages and hours. These turn out to play an important role in the variance of lifetime earnings. A big advantage of our approach to estimation compared to previous studies that have used maximum likelihood or a conventional method-of-moments approach is our ability to handle a highly unbalanced sample in the context of a model that mixes discrete and continuous variables and allows for both state dependence and multifactor heterogeneity and measurement error. Vidangos (2007) shows the potential for using models of the type we develop by studying the implications of a related multi-equation model of family income for precautionary behavior and welfare within the context of a lifecycle consumption model.[41]

We will not attempt to summarize every result here. Particularly noteworthy is our finding that the short-term earnings losses from unemployment are dominated by hours and the long-term costs by wages, our decomposition of the effect of unemployment on earnings, and our finding that job changes and unemployment spells contribute in a major way to earnings variance primarily by leading to large changes in job-specific components of wages and hours rather than through their direct effects on wages and hours. We find that human capital accumulation is the dominant source of career wage growth, although accumulation of job tenure and job mobility also play significant roles.

There are number of extensions to the model that would be worth exploring. It would be natural to add sources of aggregate risk to the model. Thus far, we simply remove year effects from wages, hours, and earnings. It would also be natural to use the model to build on work by Gottschalk and Moffitt (1994, 2006), Haider (2001), Shin and Solon (2008) and others to explore changes in the stability of earnings, although this would require a very different auxiliary model. With matched employer-employee data such as that used by Abowd et al (1999) and Bagger et al (2007), one could distinguish firm specific risk associated with observed as well as unobserved variables from job match specific risk. A much more ambitious extension is to construct a model of the household income of an individal that incorporates marriage, divorce, and death of a spouse. This will be pursued in separate

[41] He allows for additional sources of variation in family income such as health and disability shocks. The consumption model is used to quantify the welfare effects of uncertainty generated by each source of variation and to measure the contribution of each source to the accumulation of precautionary savings. Perhaps surprisingly, he finds that for plausible values of the coefficient of relative risk aversion, consumers would be willing to give up only a small percentage of consumption to eliminate risks.

work.

Given the large number of issues that the paper already addresses, we decided not to embark on the formidable task of seeking to identify how much of the stochastic variation in earnings that we analyze is anticipated by agents, how far in advance they anticipate it, or how much is insured. Adding a family income model (with private and public transfers) as in Vidangos (2007) gets partially at the question of insurance. Dealing with expectations is more difficult. One needs either data on expectations or an expanded model that incorporates decisions (such as consumption choices) that depend on –and reveal– the information set of the agent. Work by Blundell and Preston (1998), Blundell, Pistaferri, and Preston (2008), Cunha, Heckman, and Navarro (2005), and Cunha and Heckman (2006) illustrate the latter approach.

8 References

Abowd, J.M. and Card, D.E. (1987). "Intertemporal labor supply and long-term employment contracts", *American Economic Review*, 77(1), 50-68.

Abowd, J.M. and Card, D.E. (1989). "On the covariance structure of hours and earnings changes", *Econometrica*, 57(2), 411-445.

Abowd, J.M., F. Kramarz, D.N. Margolis (1999) "High Wage Workers and High Wage Firms", *Econometrica* 67 (2) , 251–333.

Aiyagari, S.R. (1994). "Uninsured idiosyncratic risk and aggregate saving", *Quarterly Journal of Economics* 109, 659-684.

Altonji, J. G., A.P. Martins and A. Siow (2002). "Dynamic Factor Models of Wages, Hours, and Earnings", *Research in Economics* 56(1), 3-59.

Altonji, J. G., and C. H. Paxson (1986). "Job Characteristics and Hours of Work", in *Research in Labor Economics*, Vol. 8, Part A, ed. by R. G. Ehrenberg, Greenwich: Westview Press, 1-55.

Altonji, J. G. and C. R. Pierret (2001). "Employer Learning and Statistical Discrimination", *Quarterly Journal of Economics*, 116, 313-350.

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

Altonji, J. G. and N. Williams (1998). "The Effects of Labor Market Experience, Job Seniority, and Mobility on Wage Growth", *Research in Labor Economics*, 17, 233-276.

Altonji, J. G. and N. Williams (2005). "Do Wages Rise With Job Seniority? A Reassessment", Industrial and Labor Relations Review, 58(3), 370-397.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.M. Robin (2007). "A Tractable Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation", unpublished paper.

Baker, M. (1997). "Growth-rate heterogeneity and the covariance structure of life cycle earnings", *Journal of Labour Economics*, 15(2), 338-375.

Baker, Michael and Gary Solon (2003)"Earnings Dynamics and Inequality Among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records", *Journal of Labor Economics* 21 (2003), 289–321.

Barlevy, Gadi, (2008) "Identification of Search Models Using Record Statistics". *Review of Economic Studies*, 75(1):29-64

Blundell, R. and I. Preston (1998), “Consumption inequality and income uncertainty”, Quarterly Journal of Economics 113, 603-640.

Blundell, R., L. Pistaferri, and I. Preston (2008). “Consumption Inequality and Partial Insurance”, *American Economic Review* (forthcoming).

Blundell, R. and T. MaCurdy (1999). “Labor Supply: A Review of Alternative Approaches”, in Handbook of Labor Economics, Vol. 3A.

Blundell, R. and I. Preston (1998). “Consumption Inequality and Income Uncertainty”. *Quarterly Journal of Economics* 113(2), 603-640.

Bound, J., Brown, C., and Mathiowetz, N. (2001). “Measurement Error in Survey Data.” in *Handbook of Econometrics, V. 5*, eds. E. E. Leamer and J. J. Heckman, pp 3705-3843.

Buchinsky,M. , Fougère, D., Kramarz, F. and Tchernis, R. (2008). ”Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.” (March). IZA Discussion Paper No. 1521.

Carrington, W. J. (1993). “Wage Losses for Displaced Workers.” *Journal of Human Resources*, 28 (3) (Summer), pp. 435–62.

Castañeda, A., Díaz-Giménez, J., and Ríos-Rull V. (2003). ”Accounting for the U.S. Earnings and Wealth Inequality”, *Journal of Political Economy*, 111(4), 818-857.

Cunha, F., J. J. Heckman, and S. Navarro (2005). “Separating Uncertainty from Heterogeneity in Life Cycle Earnings, The 2004 Hicks Lecture”. *Oxford Economic Papers* 57(2), 191—261.

Cunha, F., J. J. Heckman (2006). “Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling: A Survey of Recent Developments”, unpublished paper, University of Chicago.

Deaton, A. (1991). ”Saving and liquidity constraints”, *Econometrica*, 59(5), 1221-1248.

Farber, H. (1999), “Mobility and stability: The dynamics of job change in labor markets”, in O. Ashenfelter and D. Card Editors, *Handbook of Labor Economics* Volume 3, Part 2, 2439-2483.

Fitzgerald, Gottschalk, and Moffit (1998). ”An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics”, *Journal of Human Resources* 33(2):251-299.

Geweke, J. and Keane, M. (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*, IX, 351-80.

Gottschalk, P. and R. Moffitt (1994), “The Growth of Earnings Instability in the U.S. Labor Market.”, *Brookings Papers on Economic Activity*, Issue 2, p217-272.

Gottschalk, P., and R. Moffitt. (2006) “Trends in Earnings Volatility in the U.S.: 1970-2002.”.

Gourieroux, C., Monfort, A., and Renault, E. (1993). ”Indirect Inference”, *Journal of Applied Econometrics* 8, S85-S118.

Gourinchas, P.O., and Parker, J. (2002). ”Consumption over the Life Cycle”, *Econometrica* 70(1) 47-89.

Guvenen, F. (2007). “Learning Your Earning: Are Labor Income Shocks Really Very Persistent?”, *American Economic Review*, 97(3), 687-712.

Haider, S.J. (2001). “Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991”, *Journal of Labor Economics*, 19(4), 799-836.

Ham, J. and Reilly, K. (2002). “Testing Intertemporal Substitution, Implicit Contract, and Hours Restriction Models of the Labor Market Using Micro Data”, *American Economic Review* 92(4), 905-927.

Hause, J.C. (1980). ”The fine structure of earnings and the on-the-job training hypothesis”, *Econometrica*, 48(4), 1013-1029.

Heaton, J. and Lucas, D.J. (1996). ”Evaluating the effects of incomplete markets on risk sharing and asset pricing”, *Journal of Political Economy*, 104(3), 443-487.

Hubbard, G., Skinner, J., and Zeldes, S. (1994). ”Expanding the Life-Cycle Model: Precautionary Saving and Public Policy”, *American Economic Review (Papers and Proceedings),* 84(2), 174-179.

Huggett, M. (1996). ”Wealth Distribution in Life-Cycle Economies”, *Journal of Monetary Economics,* 38(3), 469-494.

Imrohoroglu, A. (1989). ”Cost of Business Cycles with Indivisibilities and Liquidity Constraints”, *Journal of Political Economy*, 97(6), 1364-1383.

Jacobson, L., LaLonde, R., and Sullivan, D. (1993). ”Earnings Losses of Displaced Workers”, *American Economic Review*, 83(4), 685-709.

Kambourov, G. and Manovskii, I., ”Occupation Specificity of Human Capital”, forthcoming, *International Economic Review*

Keane, M. and Smith Jr., A . A. (2003). ”Generalized Indirect Inference for Discrete Choice Models”, unpublished manuscript, Yale University.

Kletzer, L. G. (1998), “Job Displacement”, *The Journal of Economic Perspectives*, Vol. 12, No. 1 (Winter), pp. 115-136.

Krusell, P. and Smith Jr., A . A. (1997). ”Income and Wealth Heterogeneity, Portfolio Selection, and Equilibrium Asset Returns”, *Macroeconomic Dynamics*, 1, 387-422.

Krusell, P. and Smith Jr., A . A. (1998). ”Income and Wealth Heterogeneity in the Macroeconomy”, *Journal of Political Economy*, 106(5), 867-896.

Krusell, P. and Smith Jr., A . A. (1999). ”On the welfare effects of eliminating business cycles”, *Review of Economic Dynamics* 2, 254-272.

Lillard, L. and Weiss, Y. (1979). ”Components of variation in panel earnings data: American scientists 1960-1970”, *Econometrica* 47(2), 437-454.

Lillard, L. and Willis, R. (1978). ”Dynamic aspects of earning mobility”, *Econometrica* 46(5), 985-1012.

Low, H., Meghir, C., and Pistaferri, L. (2006). “Wage Risk and Employment Risk over the Life Cycle”, unpublished manuscript, Cambridge University, University College London, and Stanford University.

MaCurdy, T.E. (1982). ”The use of time series processes to model the error structure of earnings in a longitudinal data analysis”, *Journal of Econometrics*, 18, 83-114.

Meghir, C. and Pistaferri, L. (2004). ”Income variance dynamics and heterogeneity”, *Econometrica*, 72(1), 1-32.

Nagypal, E. (2007).“Learning-by-Doing versus Learning About Match Quality: Can We Tell Them Apart?”, *Review of Economic Studies*, 74 (2), 537-566.

Neal, D. (1995). “Industry-Specific Human Capital: Evidence from Displaced Workers,”*Journal of Labor Economics*, 13(4), 653–677.

Neal, D. (1999). “The Complexity of Job Mobility Among Young Men,” *Journal of Labor Economics*, 17(2), 237–261.

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(2), 306–323.

Postel-Vinay, F. and Robin, J.-M. (2002). “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity”, *Econometrica* 70(6), 2295-2350.

Postel-Vinay, F. and Turon, H. (2005). “On-the-job Search, Productivity Shocks, and the Individual Earnings Process”, unpublished manuscript, University of Bristol.

Schoenberg, U. (2005). “Testing for Asymmetric Employer Learning”, unpublished paper, University of Rochester.

Senesky, S. (2005). “Testing the Intertemporal Labor Supply Model: Are Jobs Important?”, *Labour Economics,* 12, 749-772.

Shin, D. and G. Solon. (2008) “Trends in Men’s Earnings Volatility: What Does the Panel Study of Income Dynamics Show?”, NBER Working Paper W14075

Smith, A.A., Jr. (1990). ”Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models”, Ph.D. thesis (Duke University).

Smith, A.A., Jr. (1993). ”Estimating Nonlinear Time-Series Models using Simulated Vector Autoregressions”, *Journal of Applied Econometrics* 8, S63-S84.

Stevens, A.H. (1997) ”Persistent Effects of Job Displacement: The Importance of Multiple Job Losses” *Journal of Labor Economics*, 15(1) Part 1, 165-188.

Storesletten, K., Telmer, C., and Yaron, A. (2001a). ”The Welfare Costs of Business Cycles Revisited: Finite Lives and Cyclical Variation in Idiosyncratic Risk”, *European Economic Review*, 45, 1311-1339.

Storesletten, K., Telmer, C., and Yaron, A. (2001b). ”How Important are Idiosyncratic Shocks? Evidence from Labor Supply”, *American Economic Review (Papers and Proceedings)*, 91, 413-417.

Storesletten, K., Telmer, C., and Yaron, A. (2004a). ”Consumption and Risk Sharing Over the Life Cycle”, *Journal of Monetary Economics*, 51(3), 609-633.

Storesletten, K., Telmer, C., and Yaron, A. (2004b). ”Cyclical Dynamics in Idiosyncratic Labor Market Risk”, *Journal of Political Economy*, 112(3), 695-717.

Storesletten, K., Telmer, C., and Yaron, A. (2007). “Asset Pricing with Idiosyncratic Risk and Overlapping Generations”, *Review of Economic Dynamics*, forthcoming.

Tartari, M. (2006). “Divorce and the Cognitive Achievement of Children”, unpublished paper, Department of Economics, University of Pennsylvania.

Telmer, C. (1993). ”Asset-Pricing Puzzles and Incomplete Markets”, *Journal of Finance*, 48(5), 1803-1832.

Topel, R. (1991). ”Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority”, *Journal of Political Economy*, 99(1), 145-176.

Topel, R. and Ward, M. (1992). "Job Mobility and the Careers of Young Men", *Quarterly Journal of Economics*, 107(2), 439-479.

Vidangos, I. (2004). "Understanding Fluctuations in Labor Income: A Panel VAR Analysis", unpublished manuscript, Yale University.

Vidangos, I. (2007). "Household Welfare, Precautionary Saving, and Social Insurance under Multiple Sources of Risk", unpublished manuscript, Yale University.

Wolpin, K. (1992). "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth", *Journal of Political Economy*, 100(3), 535-560.

9 Appendix 1: A Moment Condition for ϕ_1 and ϕ_2 in Model B

[This appendix needs to be revised/updated]

Recall that the autoregressive wage component in equation (**??**), is:

$$\begin{aligned}\omega_{it} &= \rho_\omega(1+\phi_1[JC_{it}+E_{it}\cdot(1-E_{i,t-1})])\omega_{i,t-1}+\gamma^\omega_{JC}JC_{it}+\gamma^\omega_{1-E}\frac{1-E_{it}}{1+UD_{it}} \\ &\quad +\delta^\omega_\mu\mu_i+(1+\phi_2[JC_{it}+E_{it}\cdot(1-E_{i,t-1})])\varepsilon^\omega_{it}\end{aligned}$$

Define $\omega^{obs}_{it} \equiv \omega_{it} + m^w_{it}$ and $S_{it} \equiv JC_{it} + E_{it}\cdot(1-E_{i,t-1})$, and rewrite the above equation as:

$$\begin{aligned}\omega^{obs}_{it} - \rho_\omega(1+\phi_1 S_{it})\omega^{obs}_{i,t-1} - \gamma^\omega_{JC}JC_{it} - \gamma^\omega_{1-E}\frac{1-E_{it}}{1+UD_{it}} \\ = \delta^\omega_\mu\mu_i + (1+\phi_2 S_{it})\varepsilon^\omega_{it} + m^w_{it} - \rho_\omega(1+\phi_1 S_{it})m^w_{i,t-1}\end{aligned} \tag{17}$$

Denote the left-hand side of 17 by $L(\rho_\omega, \phi_1, \gamma^\omega_{JC}, \gamma^\omega_{1-E}, DATA_{it})$, and the right-hand side by $R(\delta^\omega_\mu, \rho_\omega, \phi_1, \phi_2, \sigma_\omega, \sigma_{mw})$.

Now, consider the following conditional variances of $R(\delta^\omega_\mu, \rho_\omega, \phi_1, \phi_2, \sigma_\omega, \sigma_{mw})$:

$$V^R_1 \equiv Var[R(.)|S_{it}=1] = (\delta^\omega_\mu)^2 var(\mu_i|S_{it}=1) + (1+\phi_2)^2\sigma^2_\omega + \sigma^2_{mw} + \rho^2_\omega(1+\phi_1)^2\sigma^2_{mw}$$

and

$$V^R_0 \equiv Var[R(.)|S_{it}=0] = (\delta^\omega_\mu)^2 var(\mu_i|S_{it}=0) + \sigma^2_\omega + \sigma^2_{mw} + \rho^2_\omega\sigma^2_{mw}$$

Assuming that $(\delta^\omega_\mu)^2[var(\mu_i|S_{it}=1) - var(\mu_i|S_{it}=0)]$ is small, their difference is:
(A2.2)

$$D^R(\rho_\omega, \phi_1, \phi_2, \sigma_\omega, \sigma_{mw}) \equiv V^R_1 - V^R_0 \simeq [(1+\phi_2)^2 - 1]\sigma^2_\omega + \rho^2_\omega[(1+\phi_1)^2 - 1]\sigma^2_{mw}.$$

The corresponding conditional variances of $L(\rho_\omega, \phi_1, \gamma^\omega_{JC}, \gamma^\omega_{1-E}, DATA_{it})$ are:

$$
\begin{aligned}
V_1^L &\equiv Var[L(.)|S_{it}=1] \\
&= var(\omega_{it}^{obs}|S_{it}=1) \ + \rho_\omega^2(1+\phi_1)^2 var(\omega_{i,t-1}^{obs}|S_{it}=1) \\
&-2\rho_\omega(1+\phi_1)cov(\omega_{it}^{obs},\omega_{i,t-1}^{obs}|S_{it}=1) \ -2\gamma_{JC}^{\omega}cov(\omega_{it}^{obs},JC_{it}|S_{it}=1) \\
&-2\rho_\omega(1+\phi_1)\gamma_{JC}^{\omega}cov(\omega_{i,t-1}^{obs},JC_{it}|S_{it}=1)
\end{aligned}
$$

and

$$
\begin{aligned}
V_0^L &\equiv Var[L(.)|S_{it}=0] \\
&= var(\omega_{it}^{obs}|S_{it}=0) + \rho_\omega^2 var(\omega_{i,t-1}^{obs}|S_{it}=0) \\
&-2\rho_\omega cov(\omega_{it}^{obs},\omega_{i,t-1}^{obs}|S_{it}=0) - 2\gamma_{1-E}^{\omega}cov(\omega_{it}^{obs},\frac{1-E_{it}}{1+UD_{it}}|S_{it}=0) \\
&-2\rho_\omega(1+\phi_1)\gamma_{1-E}^{\omega}cov(\omega_{i,t-1}^{obs},\frac{1-E_{it}}{1+UD_{it}}|S_{it}=0)
\end{aligned}
$$

However, the wage is not observed when $E_{it}=0$. Consequently, in the observed data we can only estimate

$$
\begin{aligned}
V_1^L &= var(\omega_{it}^{obs}|S_{it}=1) + \rho_\omega^2(1+\phi_1)^2 var(\omega_{i,t-1}^{obs}|S_{it}=1,E_{i,t-1}=1) \\
&-2\rho_\omega(1+\phi_1)cov(\omega_{it}^{obs},\omega_{i,t-1}^{obs}|S_{it}=1,E_{i,t-1}=1) - 2\gamma_{JC}^{\omega}cov(\omega_{it}^{obs},JC_{it}|S_{it}=1) \\
&-2\rho_\omega(1+\phi_1)\gamma_{JC}^{\omega}cov(\omega_{i,t-1}^{obs},JC_{it}|S_{it}=1,E_{i,t-1}=1)
\end{aligned}
$$

and

$$
\begin{aligned}
V_0^L &= var(\omega_{it}^{obs}|S_{it}=0,E_{it}=1) + \rho_\omega^2 var(\omega_{i,t-1}^{obs}|S_{it}=0,E_{i,t-1}=1) \\
&-2\rho_\omega cov(\omega_{it}^{obs},\omega_{i,t-1}^{obs}|S_{it}=0,E_{it}=1,E_{i,t-1}=1) \\
&-2\gamma_{1-E}^{\omega}cov(\omega_{it}^{obs},\frac{1-E_{it}}{1+UD_{it}}|S_{it}=0,E_{it}=1) \\
&-2\rho_\omega(1+\phi_1)\gamma_{1-E}^{\omega}cov(\omega_{i,t-1}^{obs},\frac{1-E_{it}}{1+UD_{it}}|S_{it}=0,E_{it}=1,E_{i,t-1}=1)
\end{aligned}
$$

Let $D^L(\rho_\omega,\phi_1,\gamma_{JC}^{\omega},\gamma_{1-E}^{\omega},DATA_{it}) \equiv V_1^L - V_0^L$. We assume that the effects of variation in $var(\mu_i)$ induced by selection in JC_{it} and E_{it} is small. In our basic specification, JC and

E do not depend on ω_{it} so σ^2_ω does not dependent on I_{it} and E_{it}. σ^2_ω will depend on them when $\omega_{i,t-1}$ enters the JC and EE equations, as in model B.3. We ignore this.

At each stage of iteration, given estimates $\widehat{\rho}_\omega, \widehat{\phi}_1, \widehat{\gamma}^\omega_{JC}$ and$\widehat{\gamma}^\omega_{1-E}$, and the moments from the PSID data, we compute $\widehat{D}^L = D^L(\widehat{\rho}_\omega, \widehat{\phi}_1, \widehat{\gamma}^\omega_{JC}, \widehat{\gamma}^\omega_{1-E}, DATA_{it})$. We set this equal to the expression for $D^R(\rho_\omega, \phi_1, \phi_2, \sigma_\omega, \sigma_{mw})$ in (A2.2), evaluated at $\widehat{\rho}_\omega, \widehat{\phi}_1, \widehat{\sigma}_\omega, \widehat{\sigma}_{mw}$, and solve for $\hat{\phi}_2$.

This yields:

(A2.3) $$\widehat{\phi}_2(\widehat{\rho}_\omega, \widehat{\phi}_1, \widehat{\sigma}_\omega, \widehat{\sigma}_{mw}) = \sqrt{1 + \frac{\widehat{D}^L - \widehat{\rho}^2_\omega \widehat{\sigma}^2_{mw} \widehat{\phi}_1(\widehat{\phi}_1 + 2)}{\widehat{\sigma}^2_\omega}} - 1.$$

10 Appendix 2: Decomposing Career Wage Growth into the Effects of General Human Capital, Tenure, and Job Shopping

The experience profile of wages $E(wage_{it}|t)$ is the sum of the effect of general human capital accumulation, the accumulation of job tenure and the gains from job shopping. That is,

$$E(wage_{it}|t) = hc(t) + \gamma^w_{TEN} E(P(TEN_{it})|t) + E(\upsilon_{ij(t)}|t)$$

where , $hc(t)$, is the value of general human capital, $\gamma^w_{TEN} E(P(TEN\ _{it})|t)$ is the expected value of the terms of the tenure polynomial, and $E(\upsilon_{ij(t)}|t)$ is the expected value of the job match component. We approximate $E(wage_{it}|t)$ using a cubic polynomial in t and obtain estimates from the regression of $wage^*_{it}$ on a cubic in t, education, race, and a set of year dummies. The coefficients of the experience polynomial are reported in Table 2. In the estimation of Model A.3 but not the other models, we account for the fact that $\gamma^w_{TEN} E(P(TEN\ _{it})|t) + E(\upsilon_{ij(t)}|t)$ is removed from the wage residuals used with our I-I estimator by adding a quadratic in t to the wage equation. This polynomial is a quadratic approximation to $-[\gamma^w_{TEN} E(P(TEN\ _{it})|t) + E(\upsilon_{ij(t)}|t)]$.Failure to include t and t^2 when estimating the other model parametes is likely to bias the parameters involving the job match coefficient. We simulate data from the model to compute the values of $E(\upsilon_{ij(t)}|t)$ and $\hat{\gamma}^w_{TEN} E(P(TEN\ _{it})|t)$, where $\hat{\gamma}^w_{TEN}$ is taken from Altonji and Williams (2005). In figure 1 we graph $E(wage_{it}|t), hc(t), \gamma^w_{TEN} E(P(TEN\ _{it})|t)$ and $E(\upsilon_{ij(t)}|t)$. As one can see, most of the return to potential experience is due to general skill accumulation or the effect of age. Job shopping and the accumulation of tenure account for 14.6 percent and 13.5 percent, respectively, of the overall growth of wages over the first ten years. They account for 12.1 percent and 15.8 percent of growth over the first 35 years. In thinking about this, one should keep in mind that job losses counter the effects of selective mobility on growth in $E(v_{ij(t)}|t)$. The fact that we exclude the first three years of labor market experience in the I-I estimator and miss job changes probably leads to an understatement of the return to job shopping.

Table 1a
Descriptive Statistics - PSID Sample

Variable	Obs.	Mean	StDev	Min	Max
E_t	33,933	0.97	0.18	0	1
$E_t \mid E_{t-1} = 1$	32,868	0.98	0.15	0	1
$E_t \mid E_{t-1} = 0$	1,065	0.71	0.45	0	1
JC_t	33,933	0.08	0.28	0	1
ED_t	33,933	11.58	7.45	0	42.25
UD_t	33,933	0.05	0.31	0	8
TEN_t	33,933	9.34	7.81	0	42.25
$wage^*_t$	32,889	2.73	0.49	1.25	4.98
$hours^*_t$	33,933	7.73	0.29	5.30	8.34
$earn^*_t$	33,933	3.53	0.67	-5.19	6.49
$w_t^{(a)}$	32,828	0.03	0.39	-2.00	2.22
$h_t^{(a)}$	33,933	0.04	0.28	-2.51	0.87
$e_t^{(a)}$	33,933	0.06	0.57	-8.91	2.44

The table presents descriptive statistics for variables used in the structural and auxiliary models. All variables are constructed from the PSID. Lead values are excluded for sample statistics.

[(a)] Variable is the residual from a 1-st stage least-squares regression against race, years of education, a cubic in potential experience, and year indicators.

Table 1b
Additional Descriptive Statistics - PSID sample

Variable	Obs.	Mean	StDev	Min	Max
Potential Experience	33,933	19.34	8.80	4	40
Education (years)	33,933	12.94	2.38	6	17
Black	33,933	0.29	0.45	0	1
Calendar Year	33,933	1987.5	5.25	1978	1996

The table presents descriptive statistics for additional variables describing the PSID sample. Lead values are excluded.

Table 2
Point Estimates - Various Specifications

		Model A.1			Model A.2			Model A.3			Model B.1		
Column		1a	1b	1c	2a	2b	2c	3a	3b	3c	4a	4b	4c
Equation / Variable	Parameter	Point Est.	MC Mean	S.E.	Point Est.	MC Mean	S.E.	Point Est.	MC Mean	S.E.	Point Est.	MC Mean	S.E.
E-E Equation													
(cons)	γ^{EE}_{0}	**0.9360**	0.6855	(0.1878)	**0.9366**	0.6624	(0.2186)	**1.0141**	0.7833	(0.1853)	**1.0309**	0.8374	(0.1359)
$(PE_{t-1})/10$	γ^{EE}_{PE}	**-0.6208**	-0.3675	(0.1453)	**-0.8330**	-0.5520	(0.1639)	**-0.3707**	-0.1664	(0.0976)	**-0.5654**	-0.3432	(0.1225)
$(PE^2_{t-1})/100$	γ^{EE}_{PEsq}	**0.2242**	0.1590	(0.0408)	**0.2714**	0.1963	(0.0414)	**0.1465**	0.0927	(0.0231)	**0.1908**	0.1289	(0.0342)
(ED_{t-1})	γ^{EE}_{ED}	**0.0211**	0.0186	(0.0217)	**0.0298**	0.0310	(0.0246)	**0.0440**	0.0439	(0.0176)	**0.0711**	0.0740	(0.0086)
BLACK	γ^{EE}_{BLACK}	**-0.2910**	-0.2670	(0.0529)	**-0.3047**	-0.2671	(0.0646)	**-0.3608**	-0.3261	(0.0571)	**-0.3117**	-0.2724	(0.0429)
EDUC	γ^{EE}_{EDUC}	**0.1096**	0.1063	(0.0190)	**0.1264**	0.1186	(0.0246)	**0.0801**	0.0758	(0.0159)	**0.0694**	0.0621	(0.0102)
$(what_t)$	γ^{EE}_{what}				**-0.2169**	-0.1877	(0.0941)	**-0.0582**	-0.0217	(0.0763)			
(μ)	δ^{EE}_{μ}	**0.4574**	0.4275	(0.0678)	**0.5833**	0.5117	(0.1049)	**0.4426**	0.3921	(0.0936)	**0.3427**	0.3076	(0.0333)
(η)	δ^{EE}_{η}	**0.1949**	0.1809	(0.0930)	**0.1102**	0.1140	(0.1062)	**-0.2370**	-0.1868	(0.1070)	**0.1005**	0.0750	(0.0366)
U-E Equation													
(cons)	γ^{UE}_{0}	**-0.0264**	-1.0104	(0.8326)	**-0.6039**	-1.4430	(0.7368)	**0.0771**	-0.9389	(0.8416)	**-0.1514**	-0.9070	(0.7311)
$(PE_{t-1})/10$	γ^{UE}_{PE}	**-0.8677**	-0.5467	(0.4987)	**-0.0218**	0.0824	(0.4277)	**-1.0505**	-0.5641	(0.4892)	**-0.5021**	-0.2730	(0.3896)
$(PE^2_{t-1})/100$	γ^{UE}_{PEsq}	**0.2608**	0.2115	(0.1350)	**0.0699**	0.0702	(0.1195)	**0.3330**	0.2339	(0.1358)	**0.1696**	0.1322	(0.1040)
BLACK	γ^{UE}_{BLACK}	**-0.5375**	-0.4685	(0.1212)	**-0.6158**	-0.5488	(0.1493)	**-0.4860**	-0.4123	(0.1325)	**-0.4810**	-0.4424	(0.1253)
EDUC	γ^{UE}_{EDUC}	**0.1798**	0.2132	(0.0503)	**0.1600**	0.1985	(0.0619)	**0.1742**	0.2009	(0.0537)	**0.1510**	0.1746	(0.0490)
(μ)	δ^{UE}_{μ}	**0.2570**	0.2548	(0.1696)	**0.2118**	0.1879	(0.1546)	**0.6372**	0.5434	(0.1281)	**0.2276**	0.2053	(0.1591)
(η)	δ^{UE}_{η}	**0.5789**	0.4650	(0.1233)	**0.6204**	0.4967	(0.1140)	**0.2218**	0.1257	(0.1873)	**0.5889**	0.4830	(0.0943)
JC Equation													
(cons)	γ^{JC}_{0}	**-0.3114**	-0.2486	(0.1523)	**-0.3781**	-0.3414	(0.1429)	**-0.6264**	-0.5177	(0.1628)	**-0.5048**	-0.4706	(0.1725)
$(PE_{t-1})/10$	γ^{JC}_{PE}	**-0.2132**	-0.2615	(0.1116)	**0.0018**	-0.0775	(0.1261)	**-0.0983**	-0.2112	(0.1062)	**-0.2125**	-0.3026	(0.1579)
$(PE^2_{t-1})/100$	γ^{JC}_{PEsq}	**-0.0134**	0.0049	(0.0271)	**-0.0637**	-0.0368	(0.0316)	**-0.0455**	-0.0113	(0.0247)	**-0.0137**	0.0144	(0.0388)
(TEN_{t-1})	γ^{JC}_{TEN}	**-0.0786**	-0.0705	(0.0149)	**-0.1138**	-0.1065	(0.0159)	**-0.0673**	-0.0544	(0.0156)	**-0.0767**	-0.0612	(0.0166)
BLACK	γ^{JC}_{BLACK}	**0.0885**	0.0924	(0.0390)	**0.0538**	0.0629	(0.0333)	**0.1658**	0.1839	(0.0554)	**0.1033**	0.1165	(0.0407)
EDUC	γ^{JC}_{EDUC}	**-0.0325**	-0.0328	(0.0104)	**-0.0222**	-0.0206	(0.0087)	**-0.0184**	-0.0205	(0.0108)	**-0.0189**	-0.0174	(0.0102)
(υ_{t-1})	$\delta^{JC}_{\upsilon-1}$				**-0.5082**	-0.4801	(0.0697)	**-0.9230**	-0.9187	(0.1274)			
(υ_t)	δ^{JC}_{υ}				**0.2101**	0.2333	(0.0659)	**0.5936**	0.6155	(0.1410)			
(μ)	δ^{JC}_{μ}	**-0.5491**	-0.5357	(0.0651)	**-0.3522**	-0.3503	(0.0681)	**-0.2796**	-0.2935	(0.1362)	**-0.5449**	-0.5587	(0.0707)
(η)	δ^{JC}_{η}	**0.0650**	0.0827	(0.0890)	**0.1446**	0.1414	(0.0845)	**0.5308**	0.5209	(0.0995)	**0.1270**	0.1409	(0.0692)

The table presents estimates and standard errors for models A.1, A.2, A.3, and B.1. Estimates were obtained by Indirect Inference, unless indicated otherwise. The sample contains 4.632 individuals and 33,933 person-year observations. Parametric bootstrap standard errors are in parentheses. Bootstraps are based on 100 replications, except for model A.3 which uses 300 replications.

[i] Estimate obtained in first-stage least-squares regression.

[ii] Estimate obtained using additional moment conditions. See discussion in Section 4.

[iii] Imposed.

Table 2 (cont.)
Point Estimates - Various Specifications

		Model A.1			Model A.2			Model A.3			Model B.1		
Column		1a	1b	1c	2a	2b	2c	3a	3b	3c	4a	4b	4c
Equation / Variable	Parameter	Point Est.	MC Mean	S.E.	Point Est.	MC Mean	S.E.	Point Est.	MC Mean	S.E.	Point Est.	MC Mean	S.E.
Wage Equation													
BLACK	γ^{w}_{BLACK} (i)	**-0.2048**		(0.0038)	**-0.2048**		(0.0038)	**-0.2048**		(0.0038)	**-0.2048**		(0.0038)
EDUC	γ^{w}_{EDUC} (i)	**0.1049**		(0.0008)	**0.1049**		(0.0008)	**0.1049**		(0.0008)	**0.1049**		(0.0008)
Tenure polynomial		**no**			**no**			**yes**			**no**		
$(PE_{t-1})/10$ (*)	γ^{w}_{PE} (i)	**0.7514**		(0.0211)	**0.7514**		(0.0211)	**0.7514**		(0.0211)	**0.7514**		(0.0211)
$(PE^2_{t-1})/100$	γ^{w}_{PEsq} (i)	**-0.2430**		(0.0118)	**-0.2430**		(0.0118)	**-0.2430**		(0.0118)	**-0.2430**		(0.0118)
$(PE^3_{t-1})/1000$	γ^{w}_{PEcub} (i)	**0.0278**		(0.0019)	**0.0278**		(0.0019)	**0.0278**		(0.0019)	**0.0278**		(0.0019)
(μ)	δ^{w}_{μ}	**0.1264**	0.1297	(0.0124)	**0.0633**	0.0653	(0.0218)	**0.0490**	0.0477	(0.0278)	**0.1505**	0.1512	(0.0118)
(JC_t)	γ^{υ}_{0}	**0.0420**	0.0442	(0.0074)							**0.0355**	0.0359	(0.0068)
(υ_{t-1})	ρ_{υ}	**0.5651**	0.5867	(0.0215)	**0.5962**	0.6142	(0.0227)	**0.6252**	0.6399	(0.0224)			
(ε^{υ})	σ_{υ}	**0.2656**	0.2765	(0.0064)	**0.2710**	0.2820	(0.0063)	**0.2686**	0.2775	(0.0070)			
$(\varepsilon^{\upsilon}_{1})$	$\sigma_{\upsilon 1}$	**0.1443**	0.1528	(0.0315)	**0.1833**	0.1943	(0.0162)	**0.1967**	0.2103	(0.0205)			
(ω_{t-1})	ρ_{ω}	**0.9200** (ii)			**0.9200** (ii)			**0.9200** (ii)			**0.9577**	0.9603	(0.0027)
(ω_{t-1})	φ_1										**-0.2379**	-0.2339	(0.0115)
$(1-E_t)$	γ^{ω}_{1-Et}	**-0.2016**	-0.2016	(0.0100)	**-0.2317**	-0.2298	(0.0104)	**-0.1895**	-0.1877	(0.0102)	**-0.1858**	-0.1866	(0.0099)
$(1-E_{t-1})$	γ^{ω}_{1-Et-1}	**0.0978**	0.0997	(0.0121)	**0.1010**	0.1017	(0.0121)	**0.1041**	0.1052	(0.0132)	**0.0626**	0.0630	(0.0110)
(ε^{ω})	σ_{ω}	**0.0954**	0.0916	(0.0026)	**0.0929**	0.0891	(0.0025)	**0.0950**	0.0922	(0.0029)	**0.0934**	0.0904	(0.0017)
(ε^{ω})	φ_2										**2.1000**	2.1969	(0.0553)
$(\varepsilon^{\omega}_{1})$ (Black, Low Educ)	$\sigma_{\omega 1}$ (ii)	**0.2572**	0.2477	(0.0177)	**0.2557**	0.2450	(0.0142)	**0.2488**	0.2343	(0.0195)	**0.2834**	0.2827	(0.0063)
$(\varepsilon^{\omega}_{1})$ (Black, High Educ)	$\sigma_{\omega 1}$ (ii)	**0.2836**	0.2751	(0.0159)	**0.2822**	0.2727	(0.0127)	**0.2760**	0.2632	(0.0171)	**0.3076**	0.3070	(0.0058)
$(\varepsilon^{\omega}_{1})$ (White, Low Educ)	$\sigma_{\omega 1}$ (ii)	**0.2563**	0.2467	(0.0178)	**0.2547**	0.2440	(0.0142)	**0.2478**	0.2333	(0.0196)	**0.2826**	0.2819	(0.0063)
$(\varepsilon^{\omega}_{1})$ (White, High Educ)	$\sigma_{\omega 1}$ (ii)	**0.3133**	0.3057	(0.0142)	**0.3120**	0.3034	(0.0114)	**0.3064**	0.2950	(0.0152)	**0.3351**	0.3345	(0.0053)
Hours Equation													
(cons)	γ^{h}_{0}	**-0.3630**	-0.3737	(0.0084)	**-0.3609**	-0.3710	(0.0081)	**-0.3632**	-0.3747	(0.0076)	**-0.3633**	-0.3745	(0.0081)
BLACK	γ^{h}_{BLACK} (i)	**-0.1055**		(0.0043)	**-0.1055**		(0.0043)	**-0.1055**		(0.0043)	**-0.1055**		(0.0043)
EDUC	γ^{h}_{EDUC} (i)	**0.0178**		(0.0007)	**0.0178**		(0.0007)	**0.0178**		(0.0007)	**0.0178**		(0.0007)
(E_t)	γ^{h}_{E}	**0.4104**	0.4114	(0.0082)	**0.4110**	0.4122	(0.0070)	**0.4129**	0.4142	(0.0075)	**0.4157**	0.4168	(0.0074)
	σ_{ξ}	**0.1631**	0.1802	(0.0143)	**0.1611**	0.1762	(0.0131)	**0.1574**	0.1726	(0.0136)			
(w_t)	γ^{h}_{w}	**-0.0680**	-0.0681	(0.0128)	**-0.0698**	-0.0682	(0.0139)	**-0.0692**	-0.0670	(0.0148)	**-0.0929**	-0.0921	(0.0148)
(μ)	δ^{h}_{μ}	**0.0707**	0.0714	(0.0170)	**0.0894**	0.0846	(0.0188)	**0.1248**	0.1204	(0.0135)	**0.0929**	0.0935	(0.0188)
(η)	δ^{h}_{η}	**0.0991**	0.0953	(0.0198)	**0.0848**	0.0888	(0.0224)	**0.0145**	0.0200	(0.0304)	**0.1545**	0.1539	(0.0102)
(ε^{h})	σ_{h}	**0.1676**	0.1654	(0.0026)	**0.1679**	0.1659	(0.0024)	**0.1686**	0.1667	(0.0023)	**0.1800**	0.1802	(0.0009)
Earnings Equation													
(cons)	γ^{e}_{0}	**-0.0043**	-0.0059	(0.0026)	**-0.0047**	-0.0061	(0.0026)	**-0.0061**	-0.0071	(0.0024)	**-0.0044**	-0.0060	(0.0022)
(w_t)	γ^{e}_{w} (iii)	**1.0000**			**1.0000**			**1.0000**			**1.0000**		
(h_t)	γ^{e}_{h} (iii)	**1.0000**			**1.0000**			**1.0000**			**1.0000**		
	ρ_{e}	**0.5510**	0.5508	(0.0084)	**0.5498**	0.5506	(0.0073)	**0.5527**	0.5528	(0.0068)	**0.5481**	0.5476	(0.0072)
(ε^{e})	σ_{e}	**0.2110**	0.2108	(0.0015)	**0.2105**	0.2103	(0.0015)	**0.2109**	0.2107	(0.0016)	**0.2109**	0.2106	(0.0014)

The table presents estimates and standard errors for models A.1, A.2, A.3, and B.1. Estimates were obtained by Indirect Inference, unless indicated otherwise. The sample contains 4.632 individuals and 33,933 person-year observations. Parametric bootstrap standard errors are in parentheses. Bootstraps are based on 100 replications, except for model A.3 which uses 300 replications.

(*) The potential-experience profile estimated in the first-stage regression reflects the effects of general human capital accumulation, job tenure accumulation, and job shopping. Since the effects of job shopping are endogenously accounted for in model A.3, by the inclusion of a job-specific wage component that affects job mobility, model A.3 includes a "correction term" in the wage model, estimated by indirect inference. The correction term is a quadratic in potential experience. The estimated correction term is: $-0.0343 - 0.0753*(PE_{t-1})/10 + 0.0072*(PE^2_{t-1})/100$, with corresponding standard errors (0.0340), (0.0344), and (0.0075).

(i) Estimate obtained in first-stage least-squares regression.

(ii) Estimate obtained using additional moment conditions. See discussion in Section 4.

(iii) Imposed.

Table 3a
Regressions Comparing PSID Sample and Data Simulated from Models A.3 and B.1
Employment and Job Change Regressions

	PSID			Model A.3			Model B.1		
	1a [(1)]	2a [(2)]	3a [(3)]	1b [(1)]	2b [(2)]	3b [(3)]	1c [(1)]	2c [(2)]	3c [(3)]
Variable	E_t	E_t	JC_t	E_t	E_t	JC_t	E_t	E_t	JC_t
$(t_i-1)/10$	-0.0069	0.1446	-0.0702	-0.0348	-0.2304	-0.0420	-0.0415	-0.1189	-0.0541
	(0.0050)	(0.0901)	(0.0090)	(0.0010)	(0.0154)	(0.0018)	(0.0010)	(0.0156)	(0.0018)
$(t_i-1)^2/100$	-0.0003	-0.0319	0.0169	0.0052	0.0632	0.0107	0.0070	0.0296	0.0135
	(0.0011)	(0.0217)	(0.0020)	(0.0002)	(0.0038)	(0.0004)	(0.0002)	(0.0039)	(0.0004)
ED_{t-1}	0.0025			0.0031			0.0030		
	(0.0002)			(0.0000)			(0.0000)		
UD_{t-1}		-0.1071			-0.0559			-0.0498	
		(0.0181)			(0.0011)			(0.0011)	
$TEN_{t-1}/10$			-0.0803			-0.0912			-0.0922
			(0.0026)			(0.0004)			(0.0004)
Constant	0.9638	0.7453	0.2173	0.9691	0.9647	0.2159	0.9759	0.8808	0.2266
	(0.0047)	(0.0866)	(0.0086)	(0.0010)	(0.0142)	(0.0017)	(0.0010)	(0.0145)	(0.0017)
Observations	27651	708	27055	816079	34691	793445	816281	34489	793549
R-squared	0.01	0.05	0.05	0.02	0.08	0.07	0.02	0.06	0.07
RMSE	0.14	0.44	0.26	0.16	0.45	0.28	0.16	0.45	0.28

The table presents least-squares regression results comparing PSID data and data simulated from estimated models A.3 and B.1. Regressions on simulated data are based on a simulated sample which is 30 times as large as the PSID sample, but has the same demographic structure (by potential experience) as the PSID sample. Standard errors are in parentheses.

(1) Sample restricted to observations where $E_{t-1}=1$.

(2) Sample restricted to observations where $E_{t-1}=0$.

(3) Sample restricted to observations where $E_t=1$ and $E_{t-1}=1$.

Table 3b
Regressions Comparing PSID Sample and Data Simulated from Models A.3 and B.1 - Wage Regressions

	PSID				Model A.3				Model B.1			
	1a	2a	3a	4a	1b	2b	3b	4b	1c	2c	3c	4c
Variable	w_t	w_t	w_t	w_t	w_t	w_t	w_t	w_t	w_t	w_t	w_t	w_t
w_{t-1}	0.8850	0.6173	0.6142		0.9051	0.7261	0.7256		0.9010	0.7240	0.7244	
	(0.0029)	(0.0063)	(0.0063)		(0.0005)	(0.0012)	(0.0012)		(0.0005)	(0.0012)	(0.0012)	
w_{t-2}		0.3164	0.3172			0.1984	0.1979			0.1964	0.1973	
		(0.0063)	(0.0063)			(0.0012)	(0.0012)			(0.0012)	(0.0012)	
JC_t			-0.0389				-0.0129				0.0185	
			(0.0044)				(0.0007)				(0.0007)	
$(t_i-1)/10$				(0.2209)				(0.1395)				(0.1002)
				(0.0378)				(0.0074)				(0.0072)
$(t_i-1)^2/100$				0.0615				0.0456				0.0179
				(0.0188)				(0.0037)				(0.0036)
$(t_i-1)^3/1000$				(0.0074)				(0.0061)				(0.0028)
				(0.0028)				(0.0006)				(0.0005)
$TEN_{t-1}/10$				0.3944				0.4107				0.2519
				(0.0175)				(0.0029)				(0.0028)
$TEN^2_{t-1}/100$				(0.1600)				(0.2207)				(0.0882)
				(0.0138)				(0.0023)				(0.0022)
$TEN^3_{t-1}/100$				0.0249				0.0342				0.0136
				(0.0029)				(0.0005)				(0.0005)
Constant	0.0132	0.0147	0.0175	0.0591	0.0058	0.0059	0.0070	-0.0010	0.0043	0.0041	0.0025	0.0200
	(0.0011)	(0.0011)	(0.0012)	(0.0223)	(0.0002)	(0.0002)	(0.0002)	(0.0044)	(0.0002)	(0.0002)	(0.0002)	(0.0043)
Observations	27055	22587	22587	32828	793445	652618	652618	976539	793549	652765	652765	976949
R-squared	0.77	0.8	0.8	0.07	0.83	0.84	0.84	0.04	0.82	0.83	0.83	0.05
RMSE	0.18	0.17	0.17	0.37	0.17	0.16	0.16	0.4	0.17	0.16	0.16	0.39

The table presents least-squares regression results comparing PSID data and data simulated from estimated models A.3 and B.1. Regressions on simulated data are based on a simulated sample which is 30 times as large as the PSID sample, but has the same demographic structure (by potential experience) as the PSID sample. Standard errors are in parentheses.

Table 3c
Regressions Comparing PSID Sample and Data Simulated from Models A.3 and B.1 - Hours Regressions

	PSID	Model A.3	Model B.1
	1a	1b	1c
Variable	h_t	h_t	h_t
$(t_i-1)/10$	-0.0081	0.0084	0.0069
	(0.0086)	(0.0018)	(0.0018)
$(t_i-1)^2/100$	0.0006	-0.0021	-0.0019
	(0.0018)	(0.0004)	(0.0004)
h_{t-1}	0.3697	0.3236	0.2758
	(0.0067)	(0.0011)	(0.0011)
h_{t-2}	0.1826	0.2741	0.2748
	(0.0065)	(0.0011)	(0.0011)
w_t	-0.0005	-0.0191	-0.0113
	(0.0036)	(0.0007)	(0.0007)
Constant	0.0368	0.0219	0.0276
	(0.0091)	(0.0019)	(0.0019)
Observations	23322	689672	689749
R-squared	0.23	0.28	0.24
RMSE	0.21	0.24	0.23

The table presents least-squares regression results comparing PSID data and data simulated from estimated models A.3 and B.1. Regressions on simulated data are based on a simulated sample which is 30 times as large as the PSID sample, but has the same demographic structure (by potential experience) as the PSID sample. Standard errors are in parentheses.

Table 3d
Regressions Comparing PSID Sample and Data Simulated from Models A.3 and B.1 - Earnings Regressions

	PSID		Model A.3		Model B.1	
	1a	2a	1b	2b	1c	2c
Variable	e_t	e_t	e_t	e_t	e_t	e_t
$(t_i-1)/10$	0.0304		-0.0105		-0.0113	
	(0.0130)		(0.0029)		(0.0029)	
$(t_i-1)^2/100$	-0.0078		0.0023		0.0025	
	(0.0028)		(0.0006)		(0.0006)	
e_{t-1}	0.6873		0.5488		0.5330	
	(0.0069)		(0.0011)		(0.0011)	
e_{t-2}	0.1859		0.2513		0.2684	
	(0.0070)		(0.0011)		(0.0011)	
w_t		0.9232		0.9601		0.9624
		(0.0043)		(0.0008)		(0.0008)
h_t		0.7701		0.8757		0.8703
		(0.0068)		(0.0011)		(0.0012)
Constant	-0.0214	0.0214	0.0137	0.0007	0.0169	0.0039
	(0.0137)	(0.0017)	(0.0030)	(0.0003)	(0.0030)	(0.0003)
Observations	23915	32828	717450	976539	717450	976949
R-squared	0.65	0.65	0.57	0.69	0.57	0.69
RMSE	0.32	0.3	0.39	0.31	0.38	0.3

The table presents least-squares regression results comparing PSID data and data simulated from estimated models A.3 and B.1. Regressions on simulated data are based on a simulated sample which is 30 times as large as the PSID sample, but has the same demographic structure (by potential experience) as the PSID sample. Standard errors are in parentheses.

Table 4a
Decomposition of Cross-Sectional Variance in Lifetime Earnings, Wage, and Hours - Model A.3.
Shocks turned off one at a time (for all t).

	Column										
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
	Shock							Breakdown of Composite 'Shock'			
Variable	ε^e	ε^h	ε^w	Composite	η	μ	Educ	ξ	υ	E	JC
Lifetime Earnings	**6.56**	**2.35**	**12.41**	**36.65**	**-0.78**	**11.42**	**31.38**	**7.80**	**27.58**	**1.71**	**-0.44**
(SE)	(0.18)	(0.07)	(0.93)	(2.42)	(2.18)	(4.04)	(0.91)	(1.36)	(2.29)	(0.28)	(0.17)
Lifetime Wage	**0**	**0**	**20.61**	**48.18**	**-3.06**	**-0.29**	**34.56**	**0**	**47.14**	**1.69**	**-0.64**
(SE)	(0.00)	(0.00)	(1.29)	(2.28)	(1.32)	(2.77)	(1.01)	(0.00)	(2.43)	(0.37)	(0.28)
Lifetime Hours	**0**	**4.49**	**0.45**	**46.81**	**-2.78**	**46.17**	**4.86**	**42.65**	**0.61**	**3.73**	**-0.18**
(SE)	(0.00)	(0.13)	(0.17)	(7.48)	(4.22)	(7.73)	(0.42)	(7.27)	(0.29)	(0.46)	(0.06)

Entries in columns I to VII display the contribution of a given type of shock to the variance of lifetime earnings, wage, and hours, and are expressed as a percentage of the lifetime variance in the basecase. In the basecase we simulate of the full estimated model. To compute the contribution of a particular shock, we simulate the model again, setting the variance of a given shock to zero for all t. We then compute the variance of the appropriate variables. The difference relative to the basecase is the contribution of the given shock. Since the model is nonlinear, the contributions don't sum up to 100%. We normalize columns I to VII to sum to 100. Column IV is the combined contribution of the job match wage and hours components, employment and unemployment shocks, and job change shocks. In columns VIII through XI we decompose Column IV. Column VIII shows the marginal contribution of ξ, IX the marginal contribution of υ with var(ξ) set to 0, X the marginal contribution of unemployment spells with Var(ξ) and Var(υ) set to 0, and column XI displays the marginal contribution of job changes with Var(ξ) and Var(υ) set to 0, and no unemployment.

Table 4b
Decomposition of Cross-Sectional Variance in Earnings, Wage, and Hours at different t - Model A.3.
Shocks turned off one at a time (for all t).

	Column										
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
	Shock							Breakdown of Composite 'Shock'			
Variable/Horizon	ε^e	ε^h	ε^w	Composite	η	μ	Educ	ξ	υ	E	JC
Earnings											
t = 1	**14.0**	**9.3**	**23.4**	**19.0**	**0.1**	**10.0**	**24.2**	**8.4**	**10.5**	**0.2**	**0**
	(0.28)	(0.28)	(2.44)	(2.59)	(0.57)	(2.64)	(0.37)	(1.50)	(2.32)	(0.07)	(0.00)
t = 5	**17.7**	**7.7**	**17.9**	**26.1**	**0.3**	**8.5**	**21.9**	**7.8**	**18.8**	**0.9**	**-1.5**
	(0.38)	(0.27)	(1.54)	(2.07)	(1.05)	(2.59)	(0.37)	(1.35)	(1.95)	(0.22)	(0.19)
t = 10	**17.9**	**8.5**	**15.8**	**28.6**	**0.4**	**8.0**	**20.8**	**7.6**	**21.9**	**0.9**	**-1.8**
	(0.37)	(0.27)	(1.15)	(1.98)	(1.33)	(2.89)	(0.50)	(1.28)	(1.89)	(0.22)	(0.20)
t = 20	**17.0**	**8.3**	**14.7**	**32.2**	**-0.1**	**7.2**	**20.7**	**6.5**	**25.6**	**1.1**	**-0.9**
	(0.45)	(0.26)	(0.98)	(2.03)	(1.39)	(3.03)	(0.56)	(1.26)	(1.96)	(0.19)	(0.15)
t = 30	**16.8**	**8.7**	**13.1**	**32.6**	**2.5**	**6.6**	**19.7**	**6.9**	**25.4**	**0.7**	**-0.4**
	(0.42)	(0.34)	(0.91)	(1.99)	(1.74)	(3.19)	(0.59)	(1.22)	(1.91)	(0.14)	(0.12)
t = 40	**17.5**	**7.9**	**14.5**	**34.0**	**1.5**	**5.6**	**19.0**	**7.0**	**26.9**	**0.3**	**-0.2**
	(0.46)	(0.32)	(0.89)	(1.84)	(2.02)	(2.70)	(0.54)	(1.19)	(1.88)	(0.08)	(0.09)
Wage											
t = 1	**0**	**0**	**45.1**	**22.7**	**0**	**1.7**	**30.4**	**0**	**22.7**	**0**	**0**
	(0.00)	(0.00)	(4.50)	(4.60)	(0.00)	(1.77)	(0.11)	(0.00)	(4.60)	(0.00)	(0.00)
t = 5	**0**	**0**	**35.7**	**36.5**	**-0.1**	**1.0**	**27.0**	**0**	**37.4**	**0.7**	**-1.5**
	(0.00)	(0.00)	(2.52)	(2.77)	(0.50)	(1.55)	(0.47)	(0.00)	(2.72)	(0.24)	(0.26)
t = 10	**0**	**0**	**31.4**	**42.5**	**-0.8**	**0.7**	**26.2**	**0**	**43.2**	**1.0**	**-1.7**
	(0.00)	(0.00)	(1.84)	(2.29)	(0.80)	(1.77)	(0.61)	(0.00)	(2.31)	(0.31)	(0.31)
t = 20	**0**	**0**	**27.8**	**48.1**	**-1.6**	**0.0**	**25.8**	**0**	**47.3**	**1.6**	**-0.8**
	(0.00)	(0.00)	(1.55)	(2.14)	(1.05)	(2.09)	(0.73)	(0.00)	(2.22)	(0.32)	(0.24)
t = 30	**0**	**0**	**26.8**	**49.7**	**-0.7**	**-0.9**	**25.1**	**0**	**48.8**	**1.2**	**-0.3**
	(0.00)	(0.00)	(1.56)	(2.02)	(1.08)	(1.90)	(0.74)	(0.00)	(2.08)	(0.26)	(0.19)
t = 40	**0**	**0**	**27.8**	**49.8**	**0.2**	**-1.5**	**23.8**	**0**	**49.4**	**0.6**	**-0.1**
	(0.00)	(0.00)	(1.58)	(1.94)	(1.04)	(1.84)	(0.87)	(0.00)	(1.96)	(0.16)	(0.12)
Hours											
t = 1	**0**	**39.5**	**0.5**	**35.8**	**0.3**	**21.8**	**2.0**	**33.9**	**0.3**	**1.7**	**0**
	(0.00)	(1.31)	(0.19)	(5.27)	(1.97)	(4.33)	(0.07)	(5.54)	(0.12)	(0.64)	(0.00)
t = 5	**0**	**36.5**	**0.5**	**37.4**	**0.3**	**23.1**	**2.3**	**32.3**	**0.1**	**4.9**	**0.1**
	(0.00)	(1.31)	(0.19)	(4.93)	(1.97)	(3.94)	(0.38)	(5.07)	(0.14)	(0.83)	(0.03)
t = 10	**0**	**37.1**	**0.4**	**36.9**	**0.5**	**22.5**	**2.6**	**32.3**	**0.0**	**4.5**	**0.1**
	(0.00)	(1.42)	(0.16)	(4.88)	(2.03)	(3.86)	(0.35)	(4.84)	(0.17)	(0.52)	(0.04)
t = 20	**0**	**37.1**	**0.4**	**36.7**	**0.6**	**22.5**	**2.7**	**31.9**	**0.0**	**4.8**	**0.0**
	(0.00)	(1.31)	(0.16)	(5.10)	(1.98)	(4.00)	(0.38)	(5.11)	(0.21)	(0.49)	(0.03)
t = 30	**0**	**39.5**	**0.5**	**36.0**	**-0.5**	**22.4**	**2.1**	**34.1**	**-0.1**	**2.1**	**0.0**
	(0.00)	(1.44)	(0.18)	(5.46)	(2.23)	(4.24)	(0.30)	(5.47)	(0.20)	(0.32)	(0.02)
t = 40	**0**	**40.3**	**0.3**	**35.1**	**-0.6**	**23.0**	**1.8**	**35.0**	**-0.1**	**0.2**	**0.0**
	(0.00)	(1.50)	(0.17)	(5.59)	(2.37)	(4.40)	(0.37)	(5.55)	(0.19)	(0.22)	(0.02)

Entries in columns I to VII display the contribution of a given type of shock to the variance in earnings, wage, and hours for a cross section of simulated individuals with potential experience t. The contribution is expressed as a percentage of the variance in the basecase. In the basecase we simulate the full estimated model. To compute the contribution of a particular shock, we simulate the model again, setting the variance of the given shock to zero for all t. We then compute the variance of the appropriate variables at the specified value of t. The difference relative to the basecase is the contribution of the given shock. Since the model is nonlinear, the contributions don't sum up to 100%. We have normalized columns I to VII to sum to 100. Column IV is the combined contribution of the job match wage and hours components, unemployment shocks, and job change shocks. In columns VIII through XI we decompose Column IV. Column VIII is the marginal contribution of ξ, IX is the marginal contribution of υ with var(ξ) set to 0, X is the marginal contribution of eliminating unemployment spells with Var(ξ) and Var(υ) set to 0, and column XI is the marginal contribution of job changes with Var(ξ) and Var(υ) set to 0, and no unemployment.

Table 5a
Decomposition of Cross-Sectional Variance in Lifetime Earnings, Wage, and Hours - Model B.1
Shocks turned off one at a time (for all t).

	Column							
	I	II	III	IV	V	VI	VII	VIII
	Shock							
Variable	ε^e	ε^h	ε^w	J	E	η	μ	Educ
Lifetime Earnings	**6.85**	**2.66**	**24.62**	**2.35**	**1.88**	**11.53**	**18.76**	**31.35**
(SE)	(0.13)	(0.03)	(1.61)	(0.56)	(0.29)	(1.36)	(3.13)	(0.76)
Lifetime Wage	**0**	**0**	**45.59**	**3.46**	**1.83**	**0.75**	**11.46**	**36.90**
(SE)	(0.00)	(0.00)	(2.09)	(0.88)	(0.50)	(0.25)	(1.89)	(0.83)
Lifetime Hours	**0**	**4.98**	**1.87**	**0.23**	**3.03**	**64.61**	**20.39**	**4.91**
(SE)	(0.00)	(0.09)	(0.52)	(0.09)	(0.40)	(8.28)	(7.96)	(0.29)

Entries in columns I to VIII display the contribution of a given type of shock to the variance of lifetime earnings, wage, and hours, and are expressed as a percentage of the lifetime variance in the basecase. In the basecase we simulate of the full estimated model. To compute the contribution of a particular shock, we simulate the model again, setting the variance of a given shock to zero for all t. We then compute the variance of the appropriate variables. The difference relative to the basecase is the contribution of the given shock. Since the model is nonlinear, the contributions don't sum up to 100%. We normalize columns I to VIII to sum to 100.

Table 5b
Decomposition of Cross-Sectional Variance in Earnings, Wage, and Hours at different t - Model B.1
Shocks turned off one at a time (for all t).

	Column							
	I	II	III	IV	V	VI	VII	VIII
	Shock							
Variable/Horizon	ε^e	ε^h	ε^w	J	E	η	μ	Educ
Earnings								
t = 1	**14.8**	**10.9**	**25.4**	**0.0**	**0.2**	**8.3**	**17.3**	**23.1**
	(0.23)	(0.15)	(1.40)	0.00	(0.06)	(1.07)	(2.52)	(0.30)
t = 5	**17.6**	**9.9**	**26.3**	**6.2**	**1.0**	**8.1**	**11.8**	**19.1**
	(0.31)	(0.20)	(1.13)	(0.69)	(0.22)	(0.71)	(2.09)	(0.30)
t = 10	**17.1**	**8.7**	**26.7**	**7.6**	**1.4**	**7.1**	**12.0**	**19.3**
	(0.27)	(0.16)	(1.12)	(0.81)	(0.20)	(0.69)	(2.01)	(0.30)
t = 20	**17.0**	**9.8**	**25.9**	**6.1**	**1.2**	**7.6**	**11.7**	**20.6**
	(0.31)	(0.15)	(1.19)	(0.68)	(0.24)	(0.72)	(2.04)	(0.35)
t = 30	**17.3**	**9.4**	**25.3**	**4.4**	**1.1**	**8.2**	**13.9**	**20.4**
	(0.33)	(0.15)	(1.31)	(0.53)	(0.21)	(0.77)	(2.04)	(0.36)
t = 40	**18.1**	**10.3**	**25.2**	**2.3**	**0.5**	**7.8**	**14.0**	**21.9**
	(0.36)	(0.17)	(1.35)	(0.45)	(0.15)	(0.76)	(2.01)	(0.35)
Wage								
t = 1	**0**	**0**	**56.0**	**0**	**0**	**0**	**13.5**	**30.6**
	(0.00)	(0.00)	(1.99)	(0.00)	(0.00)	(0.00)	(2.11)	(0.12)
t = 5	**0**	**0**	**54.5**	**11.6**	**0.8**	**0.5**	**8.1**	**24.5**
	(0.00)	(0.00)	(1.48)	(1.01)	(0.29)	(0.27)	(1.35)	(0.42)
t = 10	**0**	**0**	**53.2**	**13.6**	**1.5**	**0.6**	**7.5**	**23.5**
	(0.00)	(0.00)	(1.45)	(1.04)	(0.38)	(0.25)	(1.14)	(0.41)
t = 20	**0**	**0**	**52.5**	**11.3**	**2.2**	**0.6**	**8.2**	**25.2**
	(0.00)	(0.00)	(1.46)	(0.84)	(0.41)	(0.27)	(1.02)	(0.48)
t = 30	**0**	**0**	**51.3**	**7.6**	**2.6**	**1.7**	**11.0**	**25.8**
	(0.00)	(0.00)	(1.51)	(0.70)	(0.42)	(0.35)	(1.17)	(0.54)
t = 40	**0**	**0**	**53.9**	**4.0**	**1.6**	**0.6**	**11.8**	**28.2**
	(0.00)	(0.00)	(1.67)	(0.77)	(0.35)	(0.23)	(1.40)	(0.65)
Hours								
t = 1	**0**	**47.8**	**1.5**	**0**	**2.0**	**35.6**	**11.1**	**2.1**
	(0.00)	(0.92)	(0.44)	(0.00)	(0.69)	(4.68)	(4.42)	(0.08)
t = 5	**0**	**44.7**	**1.6**	**0.0**	**5.1**	**34.5**	**11.3**	**2.6**
	(0.00)	(1.05)	(0.48)	(0.08)	(0.90)	(4.39)	(4.24)	(0.35)
t = 10	**0**	**44.3**	**1.8**	**0.0**	**5.1**	**34.4**	**11.6**	**2.8**
	(0.00)	(0.81)	(0.53)	(0.11)	(0.45)	(4.51)	(4.26)	(0.31)
t = 20	**0**	**44.1**	**1.7**	**0.0**	**5.1**	**34.1**	**12.0**	**3.0**
	(0.00)	(0.86)	(0.53)	(0.12)	(0.53)	(4.63)	(4.39)	(0.28)
t = 30	**0**	**47.2**	**1.7**	**0.0**	**2.1**	**35.9**	**10.7**	**2.5**
	(0.00)	(0.91)	(0.50)	(0.09)	(0.40)	(4.57)	(4.27)	(0.14)
t = 40	**0**	**49.0**	**1.8**	**0.1**	**0.0**	**36.0**	**10.7**	**2.4**
	(0.00)	(0.89)	(0.52)	(0.09)	(0.29)	(4.65)	(4.36)	(0.11)

Entries in columns I to VIII display the contribution of a given type of shock to the variance in earnings, wage, and hours for a cross section of simulated individuals with potential experience t. The contribution is expressed as a percentage of the variance in the basecase. In the basecase we simulate the full estimated model. To compute the contribution of a particular shock, we simulate the model again, setting the variance of the given shock to zero for all t. We then compute the variance of the appropriate variables at the specified value of t. The difference relative to the basecase is the contribution of the given shock. Since the model is nonlinear, the contributions don't sum up to 100%. We have normalized columns I to VIII to sum to 100.

Table 6
Point Estimates - Models A.3 and B.1 on SRC sample

		Model A.3					Model B.1	
Column		1a	1b	1c	1d	1e	2a	2b
Equation / Variable	Parameter	**SRC+SEO**	**All SRC**	**SRC Whites**			**SRC+SEO**	**All SRC**
				All	**Low Educ**	**High Educ**		
E-E Equation								
(cons)	γ^{EE}_{0}	1.0141	0.8305	0.6175	0.5553	0.7246	1.0309	0.8617
$(PE_{t-1})/10$	γ^{EE}_{PE}	-0.3707	-0.5002	-0.5523	-0.4733	-0.8043	-0.5654	-0.5832
$(PE^2_{t-1})/100$	γ^{EE}_{PEsq}	0.1465	0.1953	0.2080	0.1633	0.2691	0.1908	0.1893
(ED_{t-1})	γ^{EE}_{ED}	0.0440	0.0746	0.0893	0.0758	0.0966	0.0711	0.0795
BLACK	γ^{EE}_{BLACK}	-0.3608	-0.1691				-0.3117	-0.3176
EDUC	γ^{EE}_{EDUC}	0.0801	0.0742	0.0856	0.1065	0.0809	0.0694	0.0777
$(what_t)$	γ^{EE}_{what}	-0.0582	-0.0796	-0.1007	-0.2954	-0.2639		
(μ)	δ^{EE}_{μ}	0.4426	0.3816	0.3582	0.4245	0.3805	0.3427	0.3537
(η)	δ^{EE}_{η}	-0.2370	-0.2110	-0.1944	-0.2066	-0.1539	0.1005	0.0693
U-E Equation								
(cons)	γ^{UE}_{0}	0.0771	0.9653	1.8163	2.8077	0.5927	-0.1514	0.7095
$(PE_{t-1})/10$	γ^{UE}_{PE}	-1.0505	-1.3641	-0.7142	-1.7540	-0.5310	-0.5021	-0.9452
$(PE^2_{t-1})/100$	γ^{UE}_{PEsq}	0.3330	0.3695	0.2603	0.6474	0.1059	0.1696	0.2968
BLACK	γ^{UE}_{BLACK}	-0.4860	-0.1473				-0.4810	-0.1374
EDUC	γ^{UE}_{EDUC}	0.1742	0.0946	-0.0239	-0.0809	0.1000	0.1510	0.0805
(μ)	δ^{UE}_{μ}	0.6372	0.2948	0.3484	0.3444	0.4068	0.2276	0.1685
(η)	δ^{UE}_{η}	0.2218	0.1227	0.2701	0.0448	-0.0836	0.5889	0.3212
JC Equation								
(cons)	γ^{JC}_{0}	-0.6264	-0.3423	-0.6065	-0.2481	-1.8496	-0.5048	-0.3078
$(PE_{t-1})/10$	γ^{JC}_{PE}	-0.0983	-0.1509	0.2697	-0.1849	0.7347	-0.2125	-0.1783
$(PE^2_{t-1})/100$	γ^{JC}_{PEsq}	-0.0455	-0.0445	-0.1415	-0.0028	-0.2748	-0.0137	-0.0178
(TEN_{t-1})	γ^{JC}_{TEN}	-0.0673	-0.0528	-0.0863	-0.1237	-0.0570	-0.0767	-0.0605
BLACK	γ^{JC}_{BLACK}	0.1658	-0.0665				0.1033	-0.0796
EDUC	γ^{JC}_{EDUC}	-0.0184	-0.0368	-0.0262	-0.0201	0.0204	-0.0189	-0.0383
(υ_{t-1})	$\delta^{JC}_{\upsilon-1}$	-0.9230	-0.8088	-0.5495	-0.3932	-0.4633		
(υ_t)	δ^{JC}_{υ}	0.5936	0.7846	0.5035	0.1717	0.8330		
(μ)	δ^{JC}_{μ}	-0.2796	-0.3175	-0.1943	-0.2007	-0.2534	-0.5449	-0.5133
(η)	δ^{JC}_{η}	0.5308	0.5071	0.3593	0.3599	0.3607	0.1270	0.1834

The table presents estimates for models A.3 and B.1 restricting the PSID to the SRC sample and subsamples. Estimates were obtained by Indirect Inference, unless indicated otherwise.

(i) Estimate obtained in first-stage least-squares regression.

(ii) Estimate obtained using additional moment conditions. See discussion in Section 4.

(iii) Imposed.

Table 6 (cont.)
Point Estimates - Models A.3 and B.1 on SRC sample

		Model A.3					Model B.1	
Column		1a	1b	1c	1d	1e	2a	2b
Equation / Variable								
	Parameter	**SRC+SEO**	**All SRC**	**SRC Whites**			**SRC+SEO**	**All SRC**
				All	**Low Educ**	**High Educ**		
Wage Equation								
BLACK	γ^{w}_{BLACK} (i)	-0.2048	-0.2350				-0.2048	-0.2350
EDUC	γ^{w}_{EDUC} (i)	0.1049	0.1083	0.1069	0.0948	0.1271	0.1049	0.1083
Tenure polynomial		yes	yes	yes	yes	yes	no	no
$(PE_{t-1})/10$	γ^{w}_{PE} (i)	0.7514	0.8028	0.8182	0.8038	0.8027	0.7514	0.8028
$(PE^2_{t-1})/100$	γ^{w}_{PEsq} (i)	-0.2430	-0.2644	-0.2714	-0.2768	-0.2478	-0.2430	-0.2644
$(PE^3_{t-1})/1000$	γ^{w}_{PEcub} (i)	0.0278	0.0305	0.0312	0.0334	0.0248	0.0278	0.0305
cons. (*)	a0	-0.0343	-0.0542	-0.0514	-0.1077	0.0486		
$(PE_{t-1})/10$ (*)	a1	-0.0753	-0.0816	-0.0505	0.0671	-0.1910		
$(PE^2_{t-1})/100$ (*)	a2	0.0072	0.0092	0.0052	-0.0194	0.0334		
(μ)	δ^{w}_{μ}	0.0490	0.1015	0.0796	0.1006	0.1827	0.1505	0.1420
(JC_t)	γ^{υ}_{0}						0.0355	0.0327
(υ_{t-1})	ρ_{υ}	0.6252	0.6041	0.6136	0.5902	0.6513		
(ε^{υ})	σ_{υ}	0.2686	0.2739	0.2776	0.2690	0.2942		
$(\varepsilon^{\upsilon}_{1})$	$\sigma_{\upsilon 1}$	0.1967	0.1048	0.1318	0.1696	0.0947		
(ω_{t-1})	ρ_{ω}	0.9200 (ii)	0.9200 (ii)	0.9200 (ii)	0.9000 (ii)	0.9000 (ii)	0.9577	0.9567
(ω_{t-1})	φ_1						-0.2379	-0.2015
$(1-E_t)$	γ^{ω}_{1-Et}	-0.1895	-0.1485	-0.1370	-0.1866	-0.1740	-0.1858	-0.1561
$(1-E_{t-1})$	γ^{ω}_{1-Et-1}	0.1041	0.0744	0.0372	0.0737	0.0548	0.0626	0.0246
(ε^{ω})	σ_{ω}	0.0950	0.0937	0.0937	0.0753	0.1004	0.0934	0.0929
(ε^{ω})	φ_2						2.1000	2.1303
$(\varepsilon^{\omega}_{1})$ (Black, Low Educ)	$\sigma_{\omega 1}$ (ii)	0.2488	0.2858				0.2834	0.2878
$(\varepsilon^{\omega}_{1})$ (Black, High Educ)	$\sigma_{\omega 1}$ (ii)	0.2760	0.3098				0.3076	0.3116
$(\varepsilon^{\omega}_{1})$ (White, Low Educ)	$\sigma_{\omega 1}$ (ii)	0.2478	0.2850	0.2807	0.2522		0.2826	0.2869
$(\varepsilon^{\omega}_{1})$ (White, High Educ)	$\sigma_{\omega 1}$ (ii)	0.3064	0.3371	0.3335		0.3043	0.3351	0.3388
Hours Equation								
(cons)	γ^{h}_{0}	-0.3632	-0.4116	-0.4112	-0.3902	-0.4627	-0.3633	-0.4022
BLACK	γ^{h}_{BLACK} (i)	-0.1055	-0.0636				-0.1055	-0.0636
EDUC	γ^{h}_{EDUC} (i)	0.0178	0.0136	0.0139	0.0226	0.0197	0.0178	0.0136
(E_t)	γ^{h}_{E}	0.4129	0.4384	0.4417	0.4305	0.4698	0.4157	0.4362
	σ_{ξ}	0.1574	0.1632	0.1628	0.1426	0.1873		
(w_t)	γ^{h}_{w}	-0.0692	-0.1024	-0.0943	-0.1672	-0.1224	-0.0929	-0.1032
(μ)	δ^{h}_{μ}	0.1248	0.0947	0.0892	0.1195	0.0873	0.0929	0.0812
(η)	δ^{h}_{η}	0.0145	0.0290	0.0306	0.0251	0.0238	0.1545	0.1403
(ε^{h})	σ_{h}	0.1686	0.1402	0.1360	0.1664	0.0926	0.1800	0.1545
Earnings Equation								
(cons)	γ^{e}_{0}	-0.0061	-0.0005	-0.0053	0.0038	-0.0083	-0.0044	-0.0005
(w_t)	γ^{e}_{w} (iii)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
(h_t)	γ^{e}_{h} (iii)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	ρ_{e}	0.5527	0.6178	0.6251	0.5793	0.6849	0.5481	0.6132
(ε^{e})	σ_{e}	0.2109	0.1715	0.1662	0.1785	0.1419	0.2109	0.1720
Number of individuals		4,632	2,651	2,455	1,143	1,027	4,632	2,651
Number of person-year observations		33,933	20,502	19,131	8,446	8,305	33,933	20,502

The table presents estimates for models A.3 and B.1 restricting the PSID to the SRC sample and subsamples. Estimates were obtained by Indirect Inference, unless indicated otherwise.

(*) The potential-experience profile estimated in the first-stage regression reflects the effects of general human capital accumulation, job tenure accumulation, and job shopping. Since the effects of job shopping are endogenously accounted for in model A.3, by the inclusion of a job-specific wage component that affects job mobility, model A.3 includes a quadratic in potential experience as an adjustment in the wage model, which is estimated by indirect inference.

(i) Estimate obtained in first-stage least-squares regression.

(ii) Estimate obtained using additional moment conditions. See discussion in Section 4.

(iii) Imposed.

Table 7
Decomposition of Cross-Sectional Variance in Lifetime Earnings, Wage, and Hours - Model A.3
SRC Sample and SRC Whites Sample by Education

	Column										
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
	Shock							Breakdown of Composite 'Shock'			
Variable	ε^e	ε^h	ε^w	Composite	η	μ	Educ	ξ	υ	E	JC
	Panel A: SRC										
Lifetime Earnings	**5.3**	**1.6**	**11.7**	**35.4**	**-0.3**	**12.8**	**33.4**	**7.4**	**26.9**	**1.5**	**-0.4**
Lifetime Wage	**0**	**0**	**20.3**	**45.0**	**-2.8**	**3.5**	**34.0**	**0**	**43.5**	**1.6**	**-0.2**
Lifetime Hours	**0**	**3.5**	**1.2**	**57.8**	**0.0**	**31.1**	**6.4**	**53.2**	**1.2**	**3.5**	**-0.1**
	Panel B: SRC Whites, Low Education										
Lifetime Earnings	**5.9**	**2.5**	**5.1**	**41.3**	**-1.6**	**33.1**	**13.7**	**9.7**	**26.9**	**4.8**	**0.0**
Lifetime Wage	**0**	**0**	**12.2**	**63.7**	**-4.6**	**13.7**	**14.9**	**0**	**57.5**	**5.7**	**0.5**
Lifetime Hours	**0**	**5.0**	**1.4**	**47.5**	**2.2**	**41.1**	**2.8**	**40.5**	**4.3**	**2.9**	**-0.2**
	Panel C: SRC Whites, High Education										
Lifetime Earnings	**5.7**	**0.8**	**10.1**	**44.4**	**2.0**	**29.3**	**7.6**	**13.4**	**29.3**	**1.6**	**0.1**
Lifetime Wage	**0**	**0**	**18.3**	**55.8**	**-0.1**	**18.3**	**7.7**	**0**	**53.4**	**1.9**	**0.4**
Lifetime Hours	**0**	**1.7**	**0.8**	**77.2**	**-0.7**	**20.2**	**0.8**	**73.5**	**2.4**	**1.3**	**-0.1**

Entries in columns I to VII display the contribution of a given type of shock to the variance of lifetime earnings, wage, and hours, and are expressed as a percentage of the lifetime variance in the basecase. In the basecase we simulate of the full estimated model. To compute the contribution of a particular shock, we simulate the model again, setting the variance of a given shock to zero for all t. We then compute the variance of the appropriate variables. The difference relative to the basecase is the contribution of the given shock. Since the model is nonlinear, the contributions don't sum up to 100%. We normalize columns I to VII to sum to 100. Column IV is the combined contribution of the job match wage and hours components, employment and unemployment shocks, and job change shocks. In columns VIII through XI we decompose Column IV. Column VIII shows the marginal contribution of ξ, IX the marginal contribution of υ with var(ξ) set to 0, X the marginal contribution of unemployment spells with Var(ξ) and Var(υ) set to 0, and column XI displays the marginal contribution of job changes with Var(ξ) and Var(υ) set to 0, and no unemployment.

Table A1
Composition of PSID Sample before Sample Selection Based on Employment Status.

Emp. Status	Percentage
Working	87.98
Temp. Laidoff	1.48
Unemployed	5.9
Retired	0.87
Disabled	1.85
Housewife	0.19
Student	1.17
Other	0.56

The table presents the composition of the PSID sample, in terms of employment status, before we impose any sample restrictions based on employment status. The sample here meets all selection criteria which are not based on employment status.

Table A2
Percentage of Observations Excluded Based on Employment Status.

PE	Percentage	PE	Percentage	PE	Percentage	PE	Percentage
1	16.6 [(a)]	11	2.5	21	3.2	31	6.6
2	9.5	12	2.8	22	3.4	32	7.2
3	6.3	13	2.9	23	4.5	33	8.2
4	4.8	14	3.5	24	5.1	34	9.0
5	4.4	15	2.5	25	5.4	35	9.4
6	3.4	16	2.7	26	5.4	36	11.7
7	2.8	17	3.2	27	5.6	37	13.4
8	2.8	18	2.9	28	5.1	38	14.5
9	2.5	19	3.1	29	5.6	39	18.4
10	2.3	20	3.6	30	6.3	40	21.97 [(b)]

The table presents the percentage of observations excluded, based on employment status at the survey date, for each value of potential experience (PE).
(a) All students at PE=1.
(b) Of these, 13.5 are retired, 7.5 disabled.

Table A3
Distribution of Number of Observations Contributed Per Individual in PSID Sample.

Percentile	Min	5%	25%	50%	75%	95%	Max
Number of observations per individual	1	1	3	6	11	18	19

The table presents the cross-sectional distribution, across individuals, of the number of observations contributed to the sample by individual. Lead values are excluded.

Figure 1: Decomposing the Experience Profile of Wages

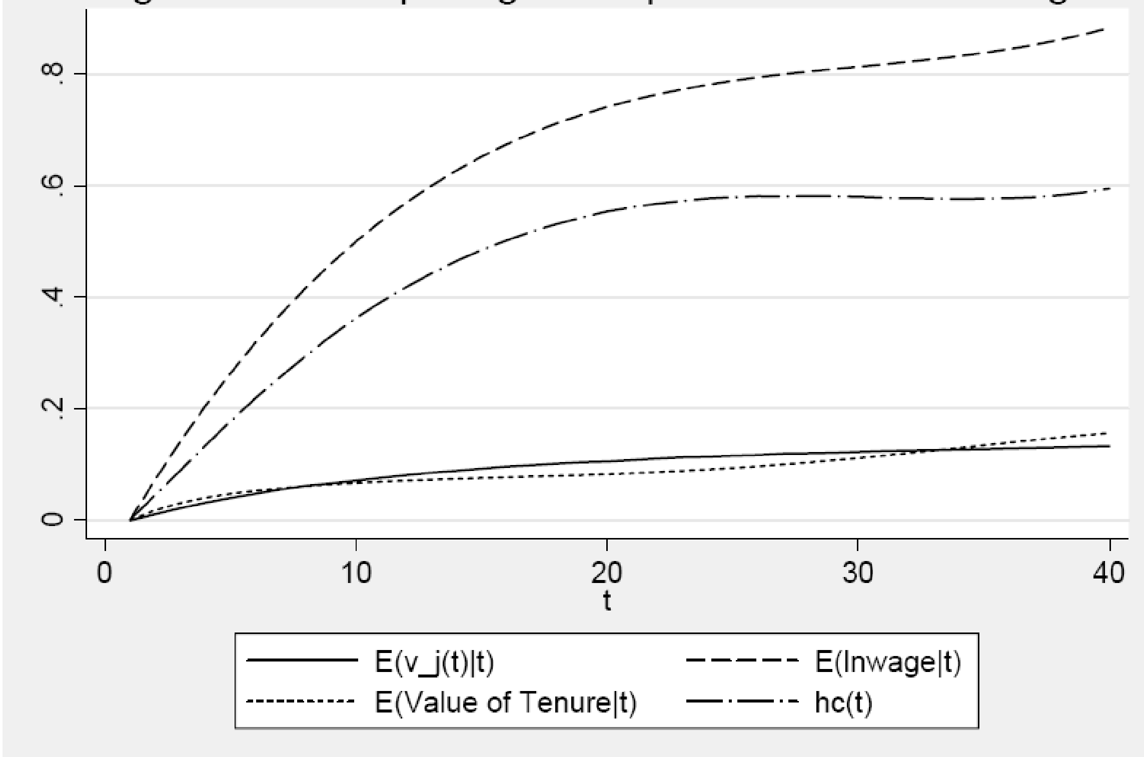


Figure 2
Comparisons of PSID and Simulated Data

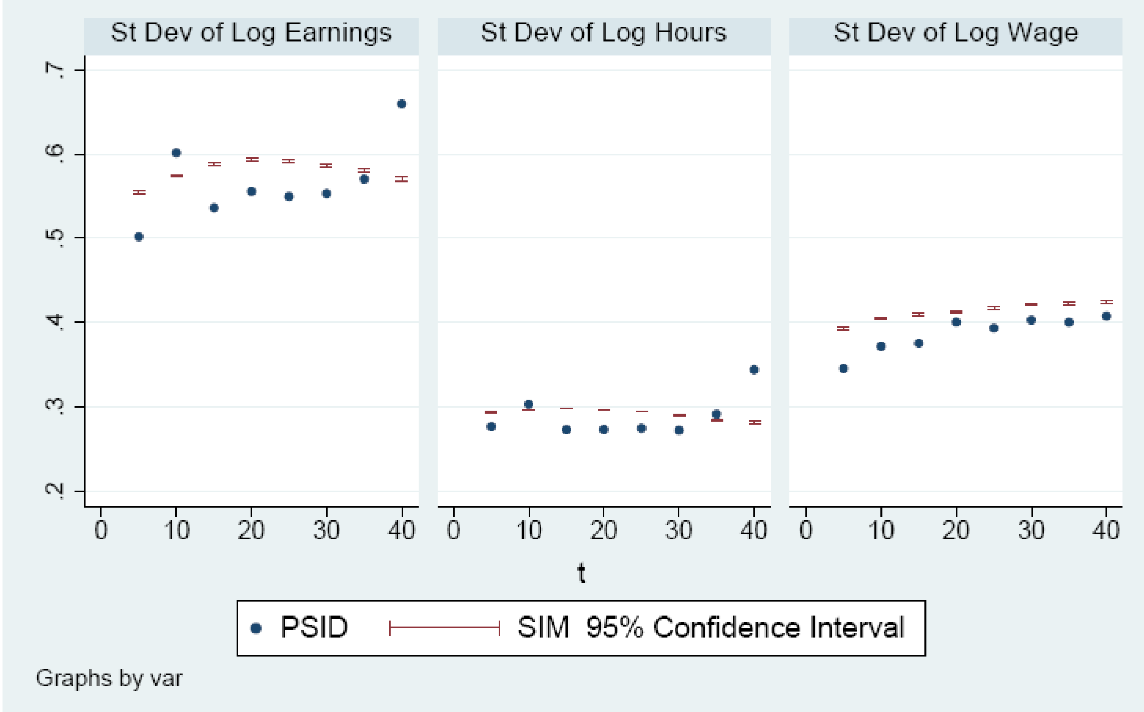


Figure 3
Comparisons of PSID and Simulated Data

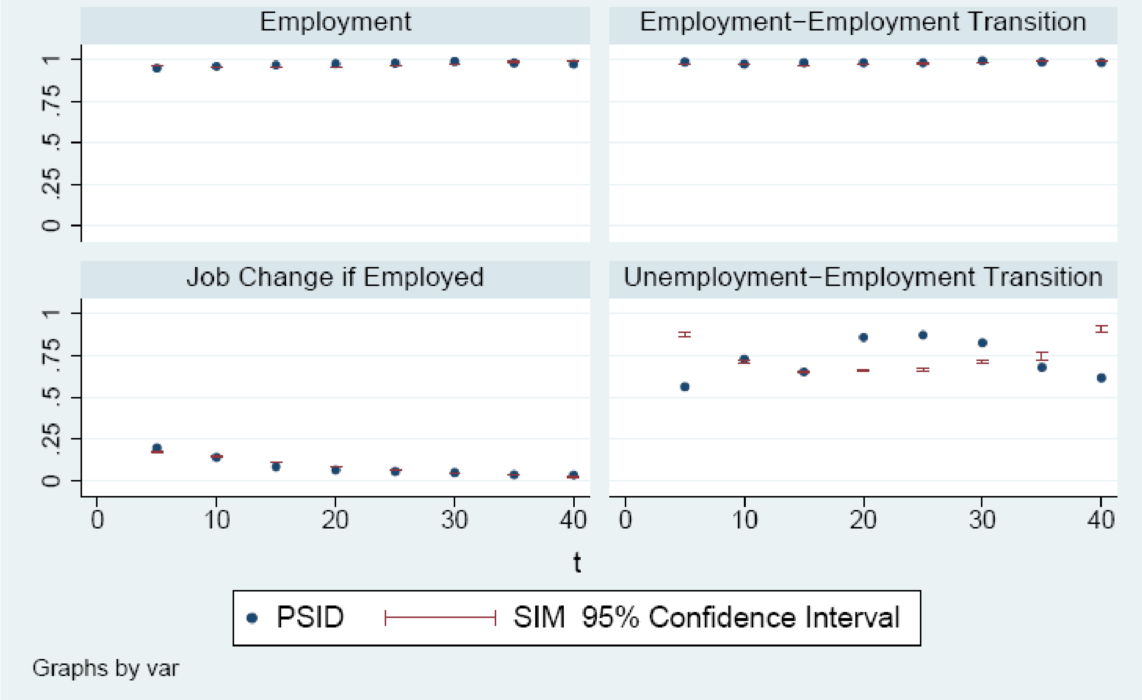
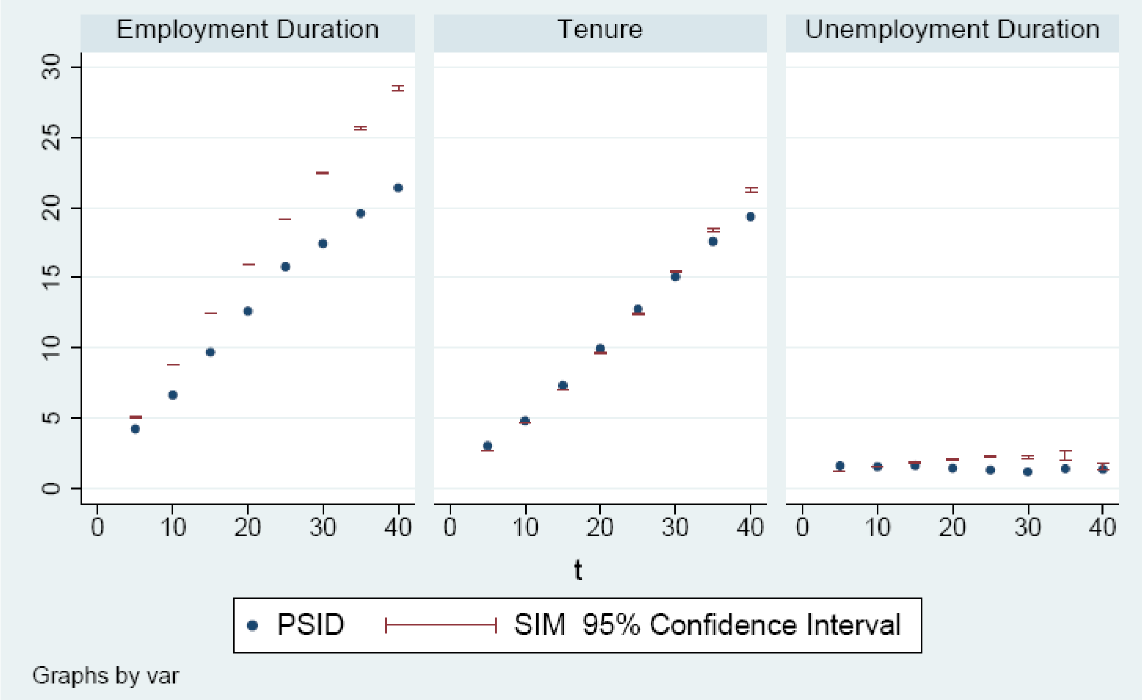


Figure 4
Comparisons of PSID and Simulated Data



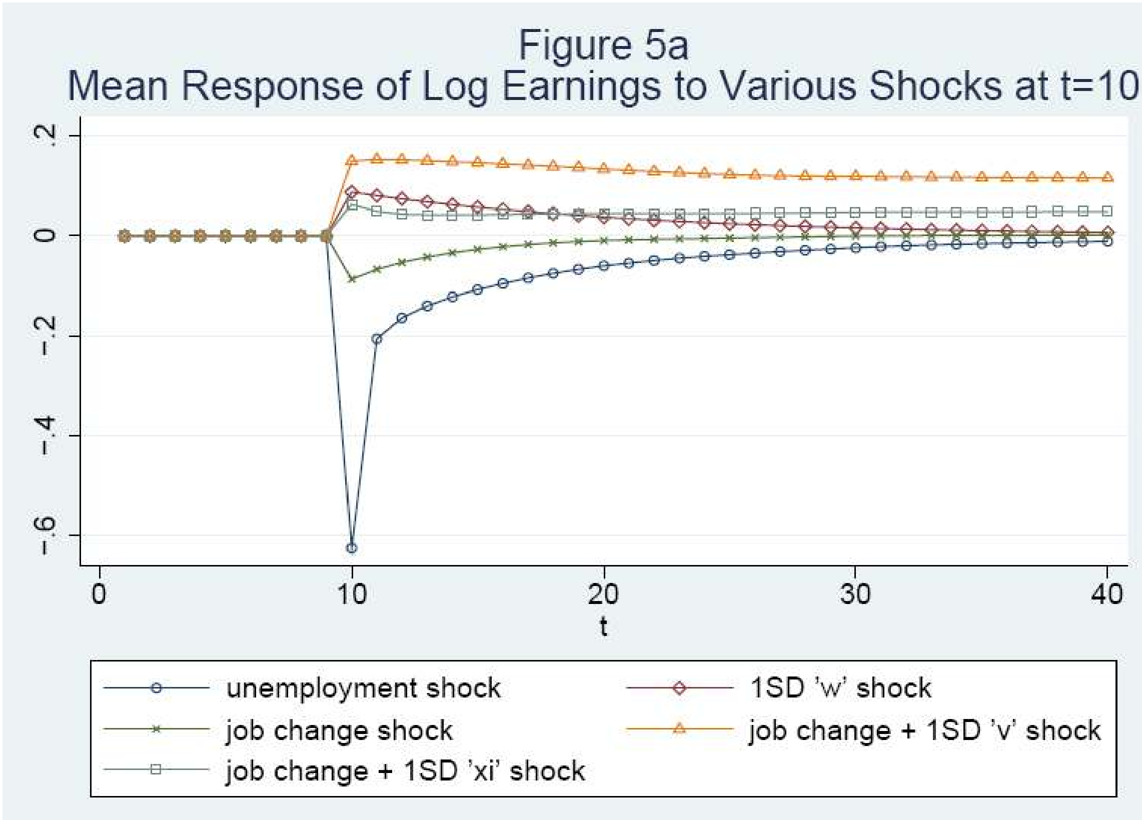


Figure 5a
Mean Response of Log Earnings to Various Shocks at t=10
.2
0
-.2
-.4
-.6
0
10
20
30
40
t
unemployment shock
1SD 'w' shock
job change shock
job change + 1SD 'v' shock
job change + 1SD 'xi' shock

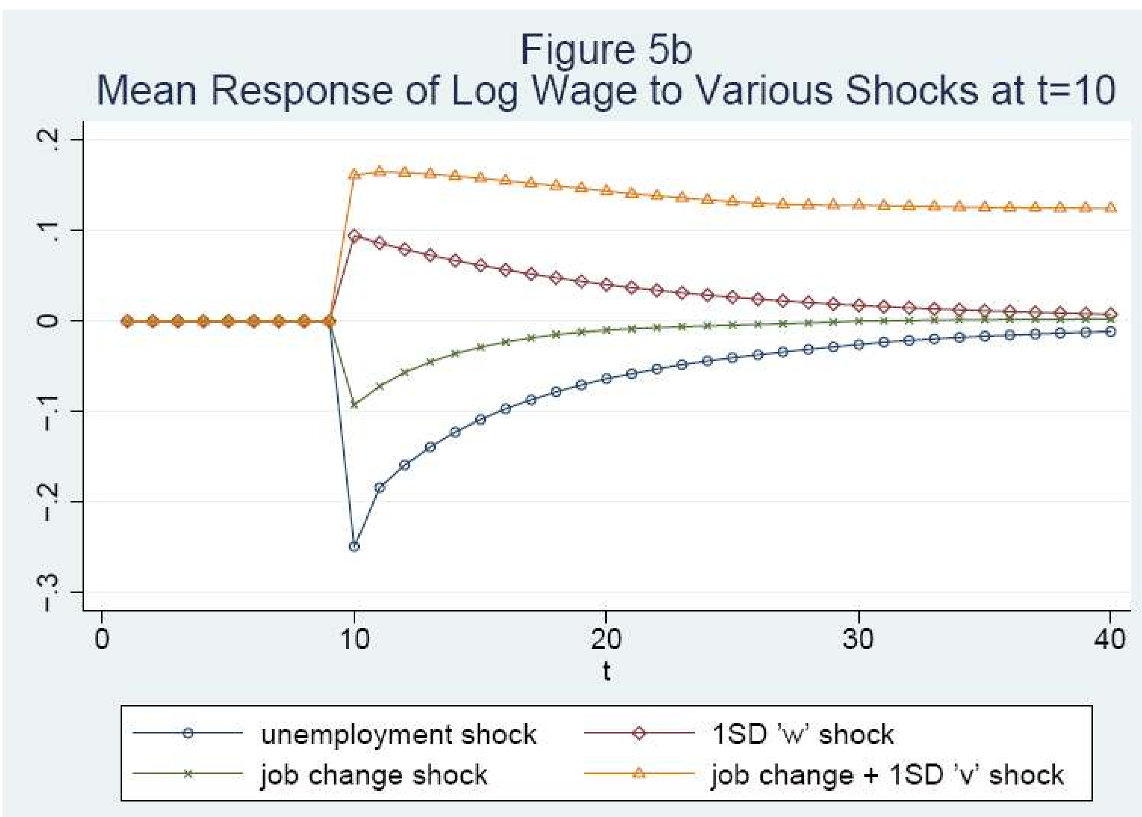


Figure 5b
Mean Response of Log Wage to Various Shocks at t=10
.2
.1
0
-.1
-.2
-.3
0
10
20
30
40
t
unemployment shock
1SD 'w' shock
job change shock
job change + 1SD 'v' shock

Figure 5c
Mean Response of Log Hours to Various Shocks at t=10

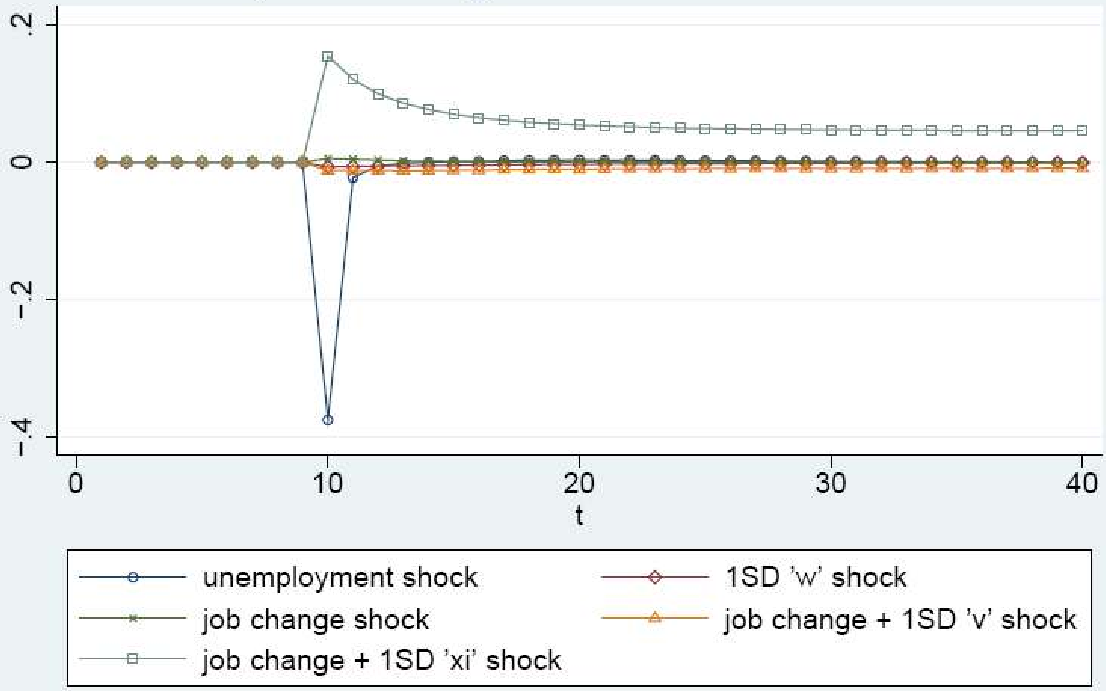
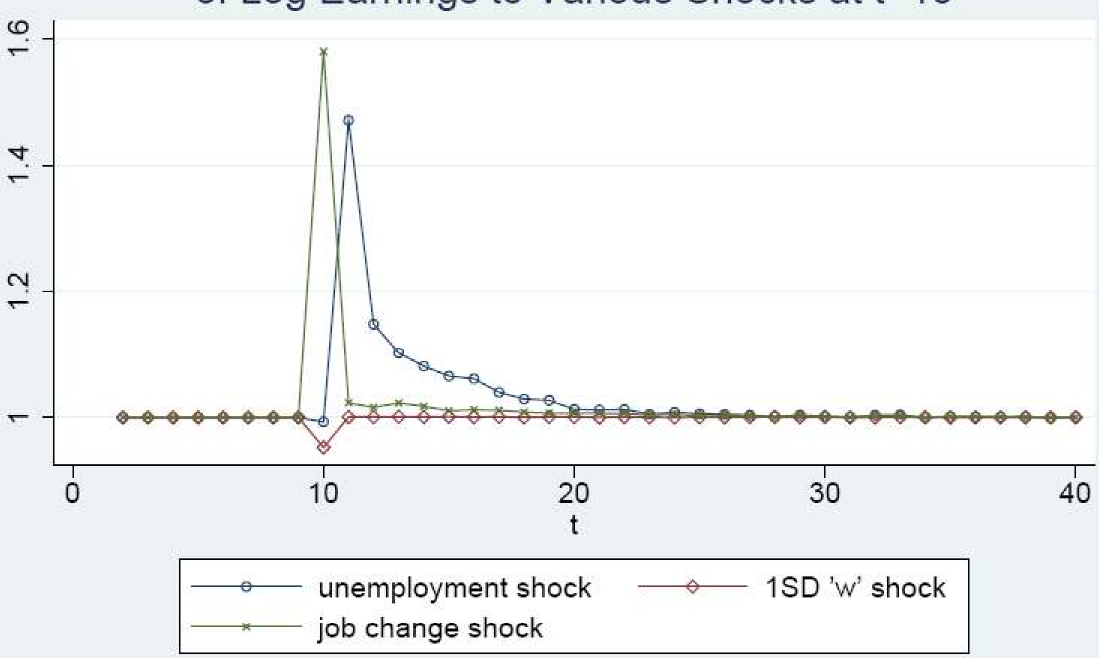


Figure 6a
Response of Cross-Sectional Variance of the First Difference
of Log Earnings to Various Shocks at t=10



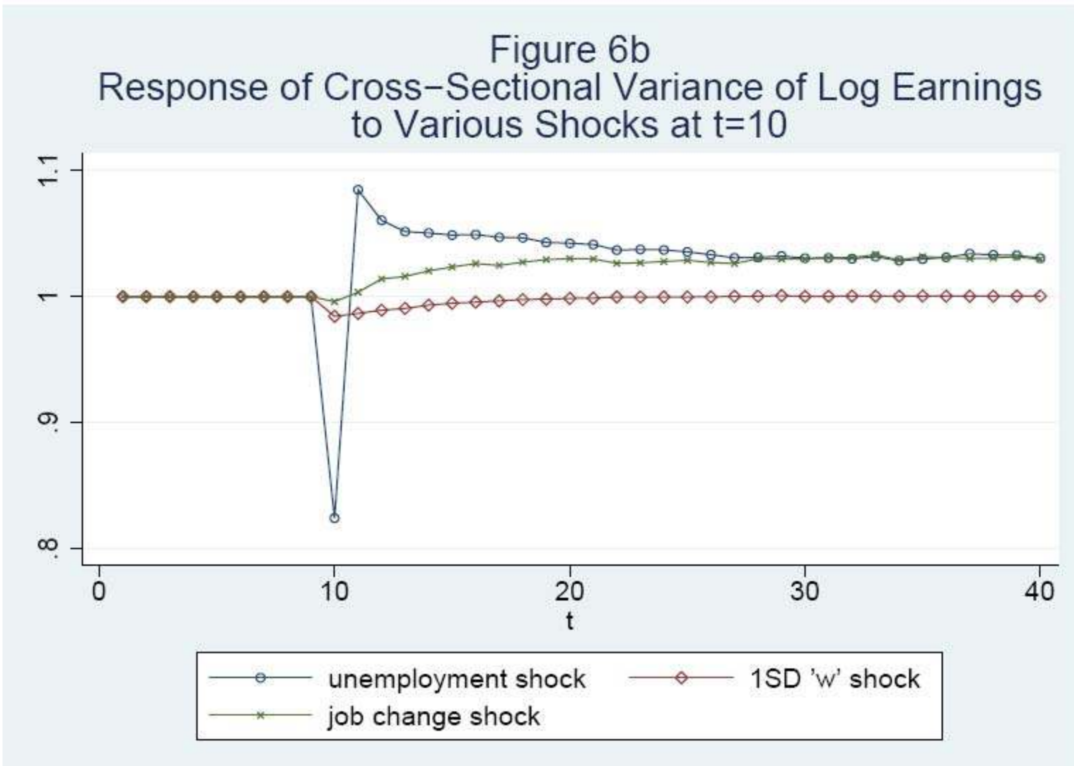


Figure 6b
Response of Cross−Sectional Variance of Log Earnings
to Various Shocks at t=10
1.1
1
.9
.8
0
10
20
30
40
t
unemployment shock
1SD 'w' shock
job change shock

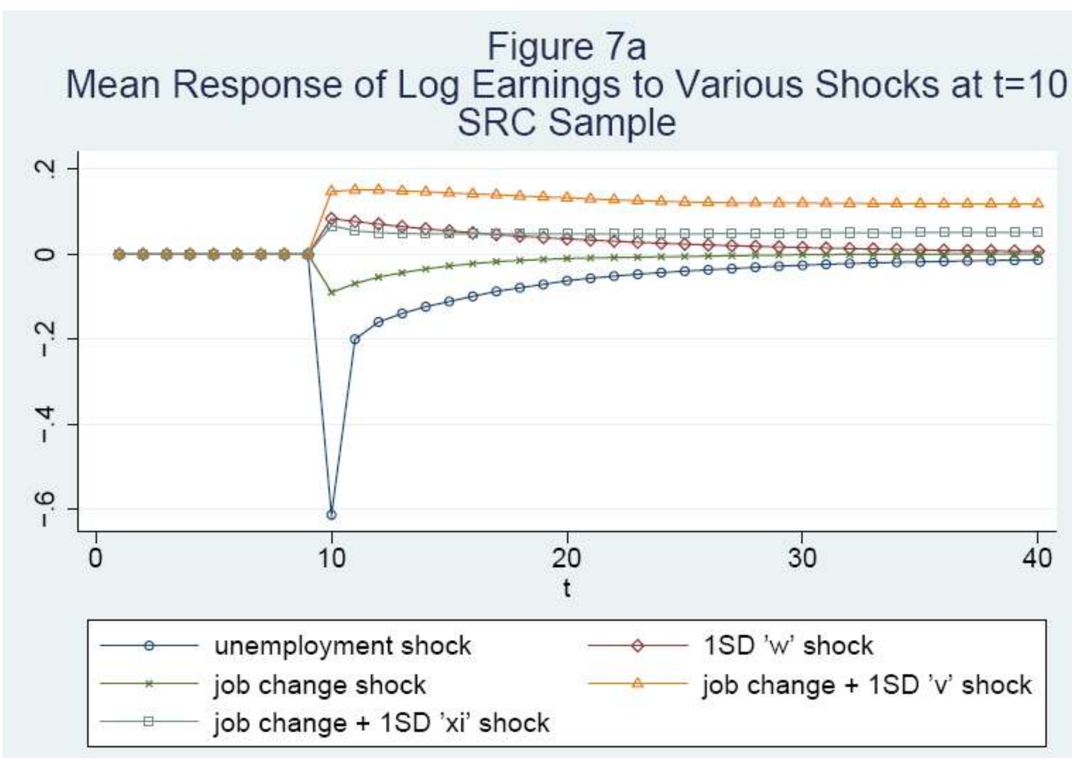


Figure 7a
Mean Response of Log Earnings to Various Shocks at t=10
SRC Sample
.2
0
-.2
-.4
-.6
0
10
20
30
40
t
unemployment shock
1SD 'w' shock
job change shock
job change + 1SD 'v' shock
job change + 1SD 'xi' shock

Figure 7b
Mean Response of Log Earnings to Various Shocks at t=10
SRC White, Low Education Sample

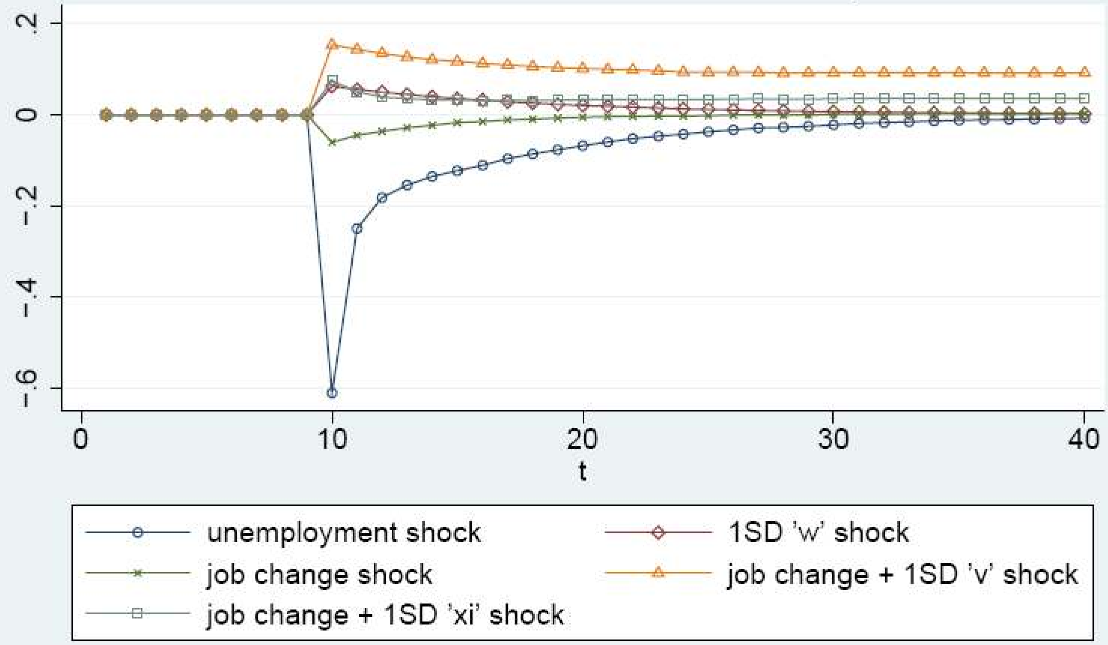
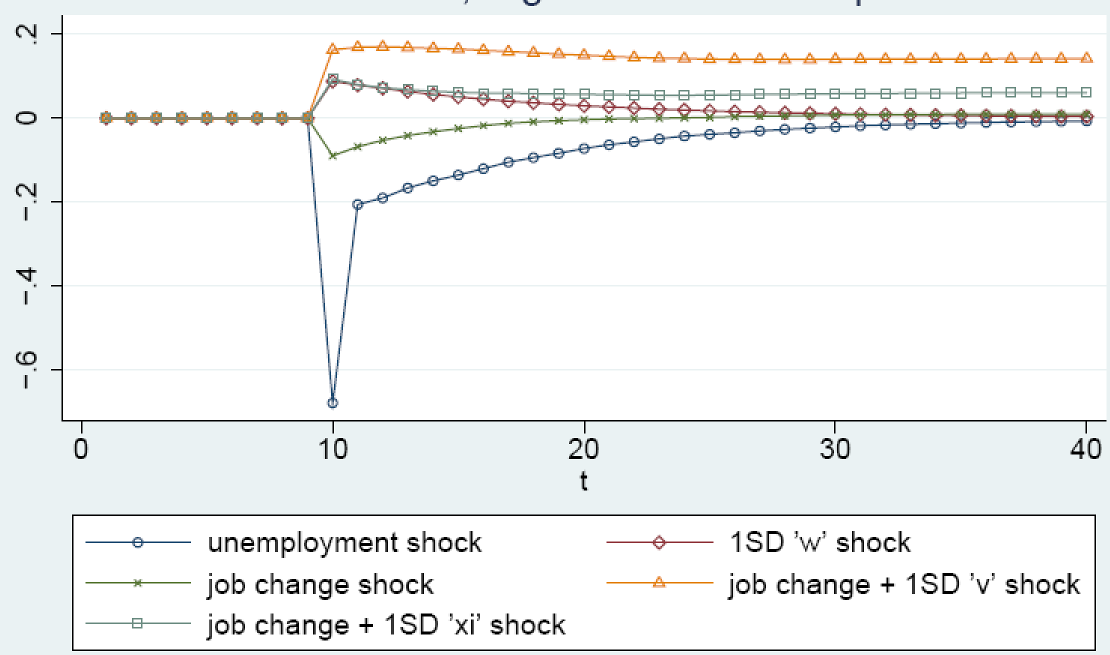


Figure 7c
Mean Response of Log Earnings to Various Shocks at t=10
SRC White, High Education Sample



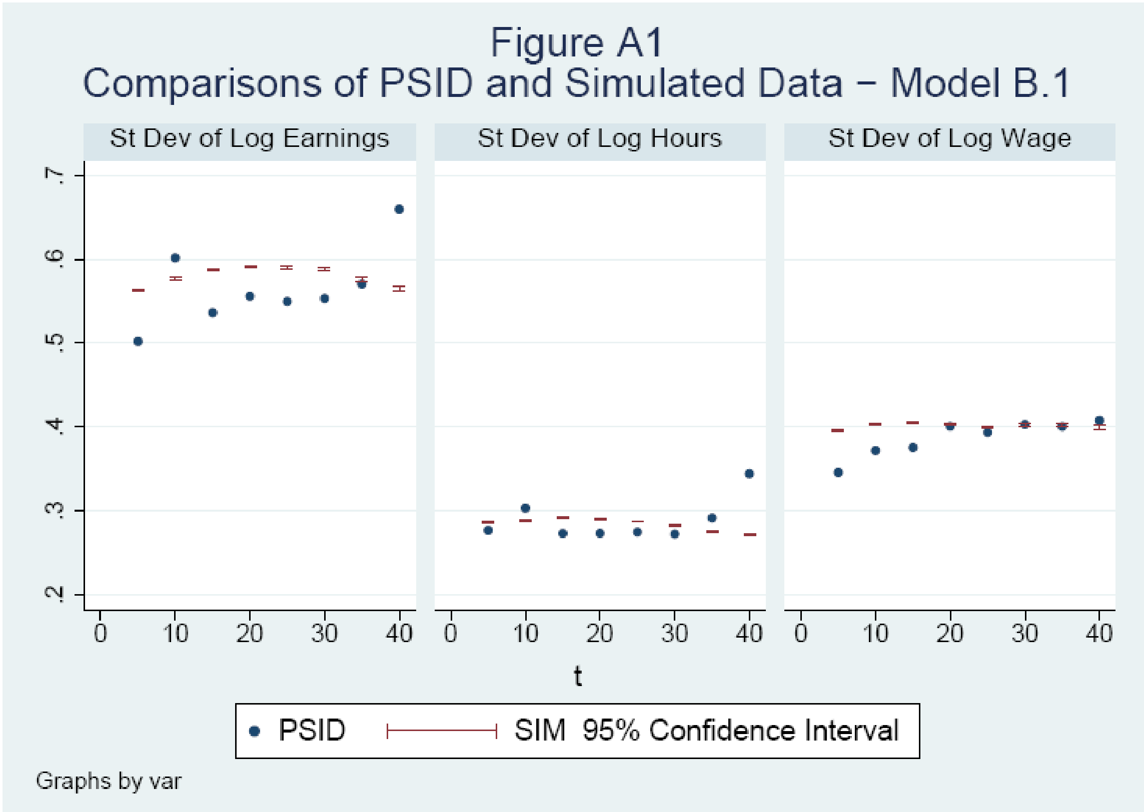


Figure A1
Comparisons of PSID and Simulated Data − Model B.1
St Dev of Log Earnings
St Dev of Log Hours
St Dev of Log Wage
.2
.3
.4
.5
.6
.7
0
10
20
30
40
t
PSID
SIM 95% Confidence Interval
Graphs by var

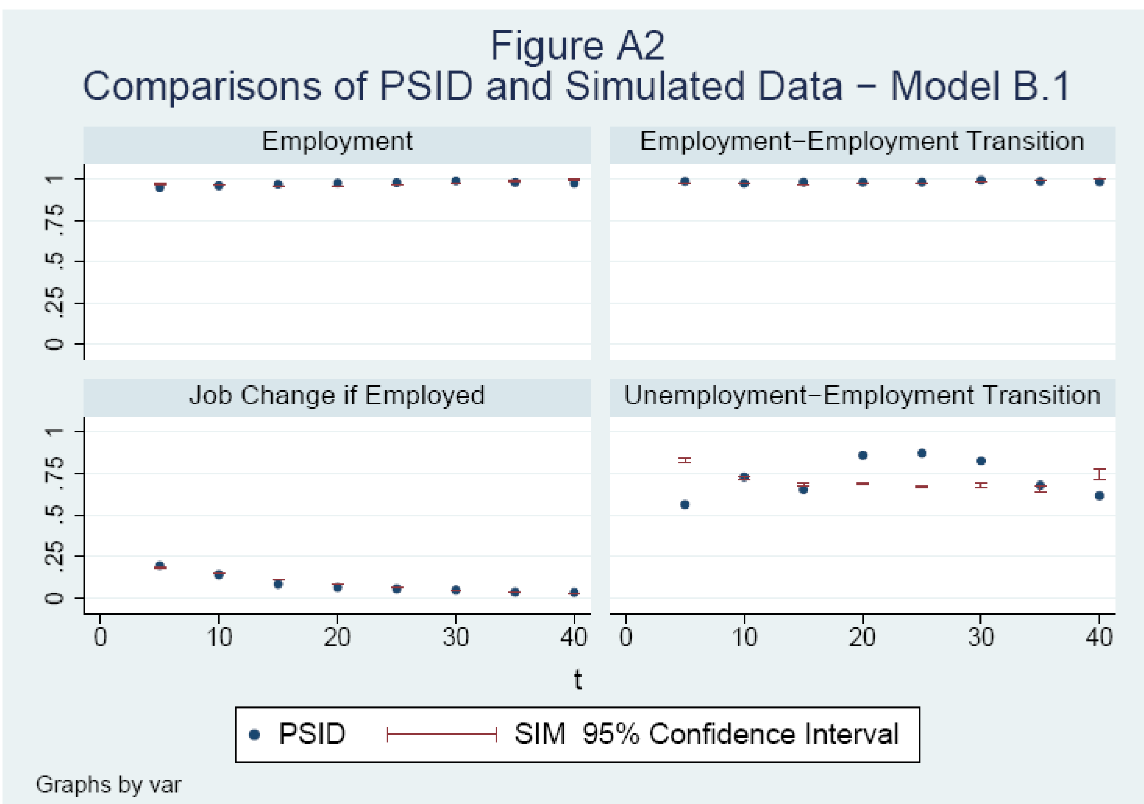


Figure A2
Comparisons of PSID and Simulated Data − Model B.1
Employment
Employment−Employment Transition
Job Change if Employed
Unemployment−Employment Transition
0
.25
.5
.75
1
0
10
20
30
40
t
PSID
SIM 95% Confidence Interval
Graphs by var

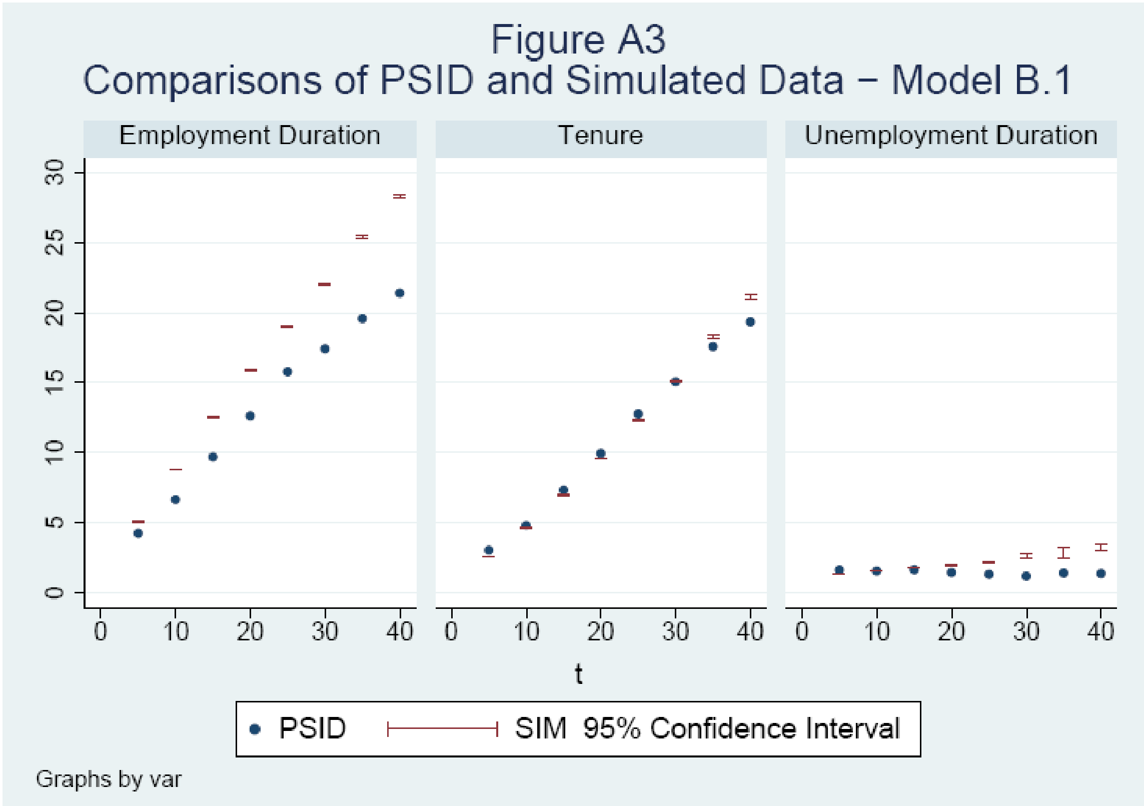


Figure A3
Comparisons of PSID and Simulated Data − Model B.1
Employment Duration
Tenure
Unemployment Duration
0
5
10
15
20
25
30
0
10
20
30
40
t
PSID
SIM 95% Confidence Interval
Graphs by var

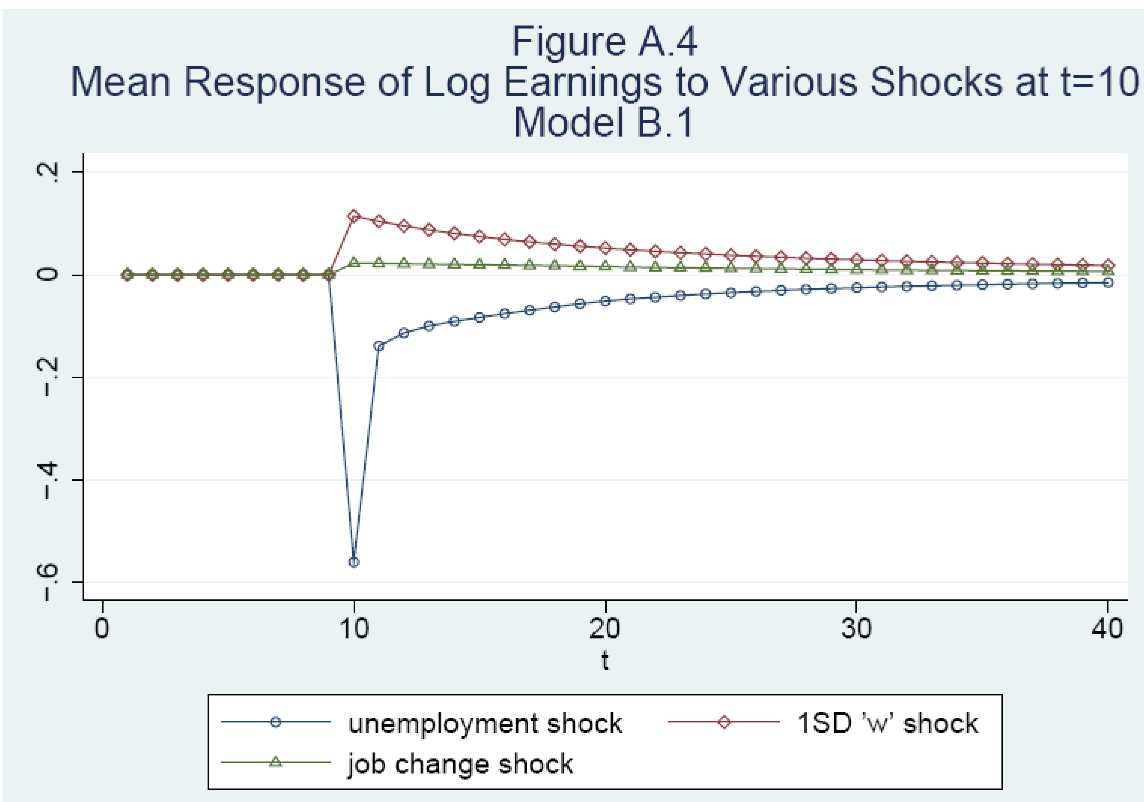


Figure A.4
Mean Response of Log Earnings to Various Shocks at t=10
Model B.1
.2
0
−.2
−.4
−.6
0
10
20
30
40
t
unemployment shock
1SD 'w' shock
job change shock

Figure A5
Mean Response of Log Wage to Various Shocks at t=10
Model B.1

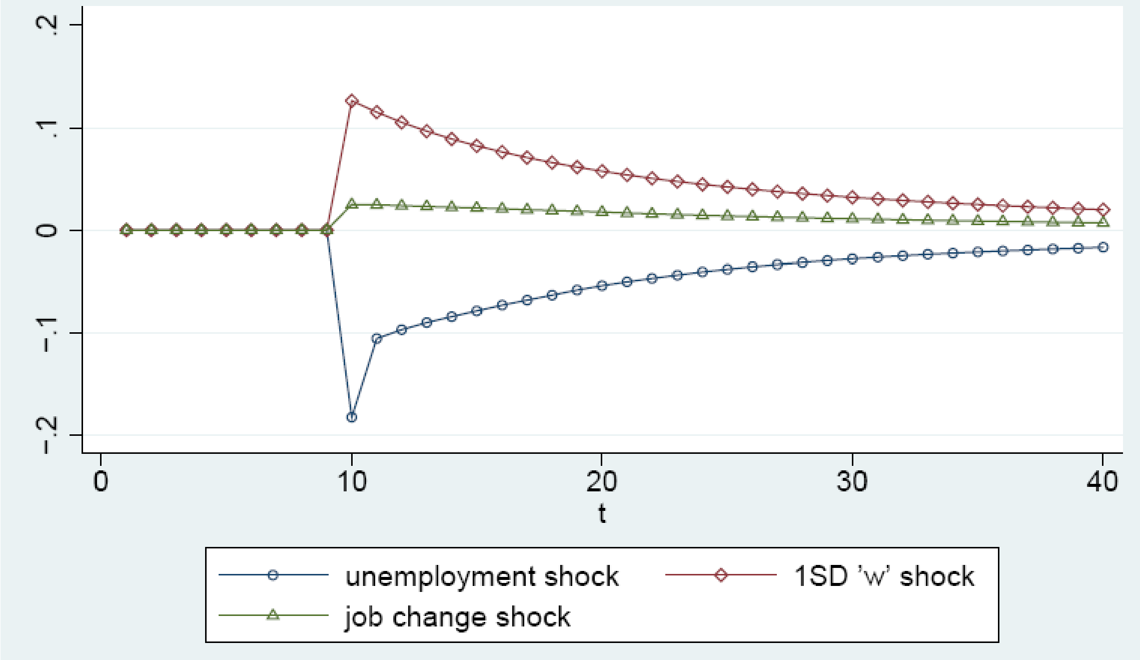
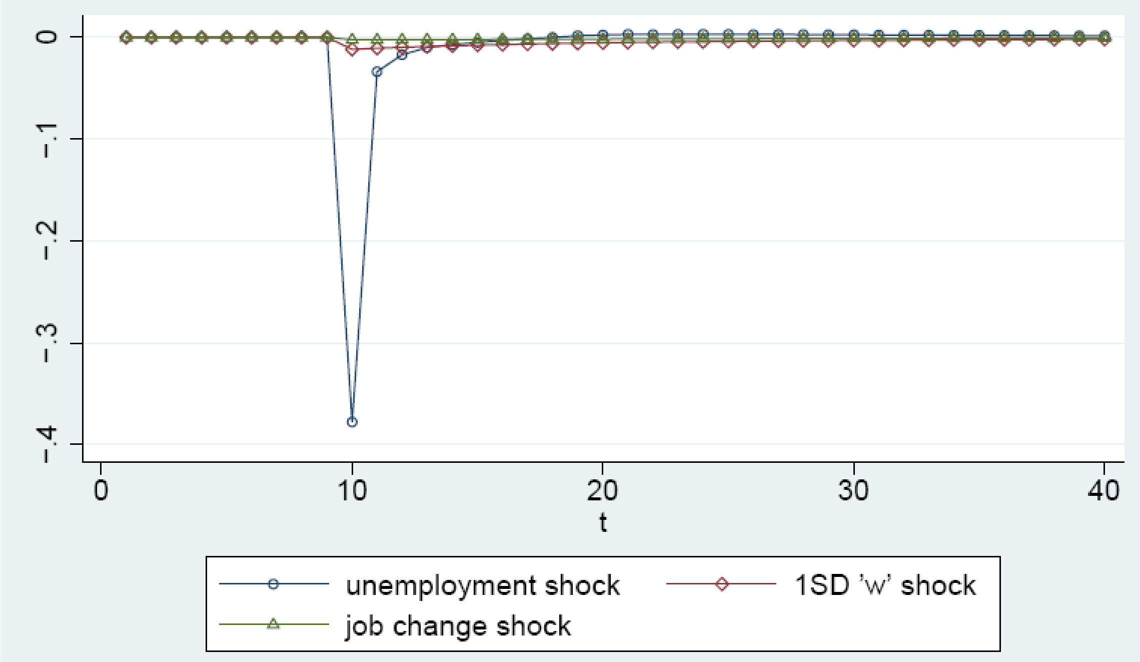


Figure A6
Mean Response of Log Hours to Various Shocks at t=10
Model B.1



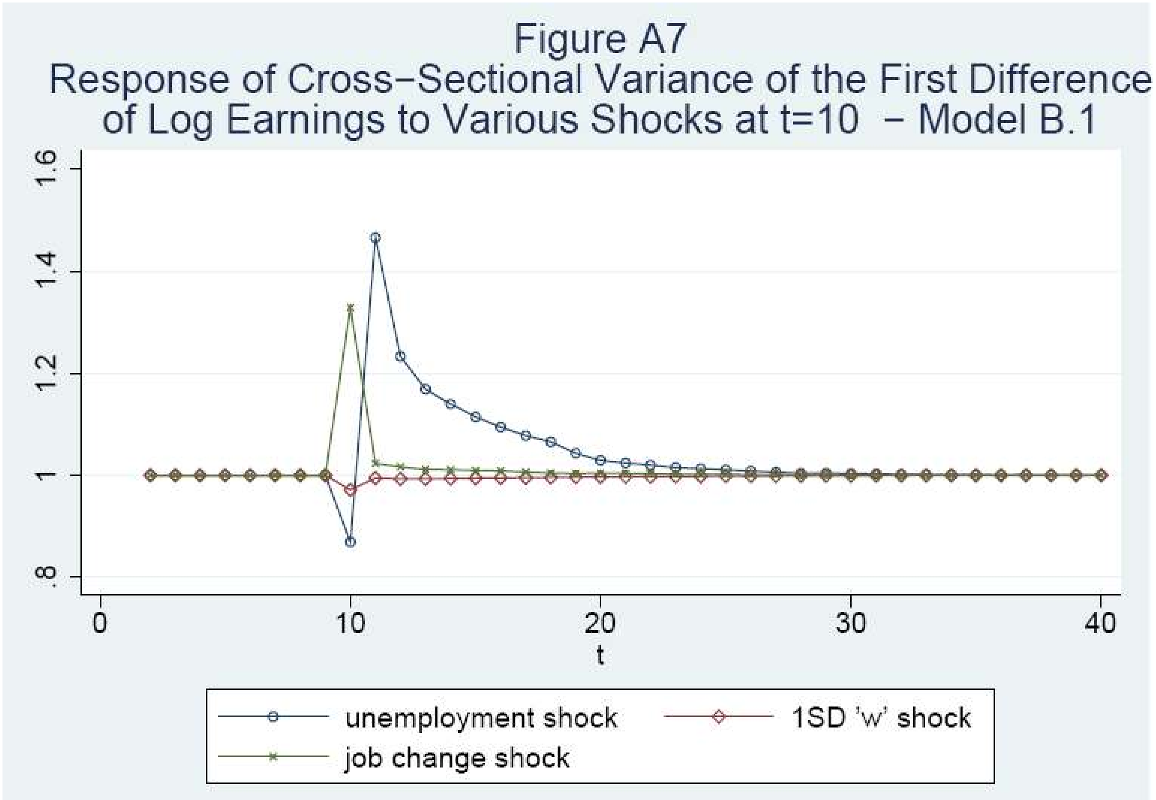


Figure A7
Response of Cross−Sectional Variance of the First Difference of Log Earnings to Various Shocks at t=10 − Model B.1
.8
1
1.2
1.4
1.6
0
10
20
30
40
t
unemployment shock
1SD 'w' shock
job change shock

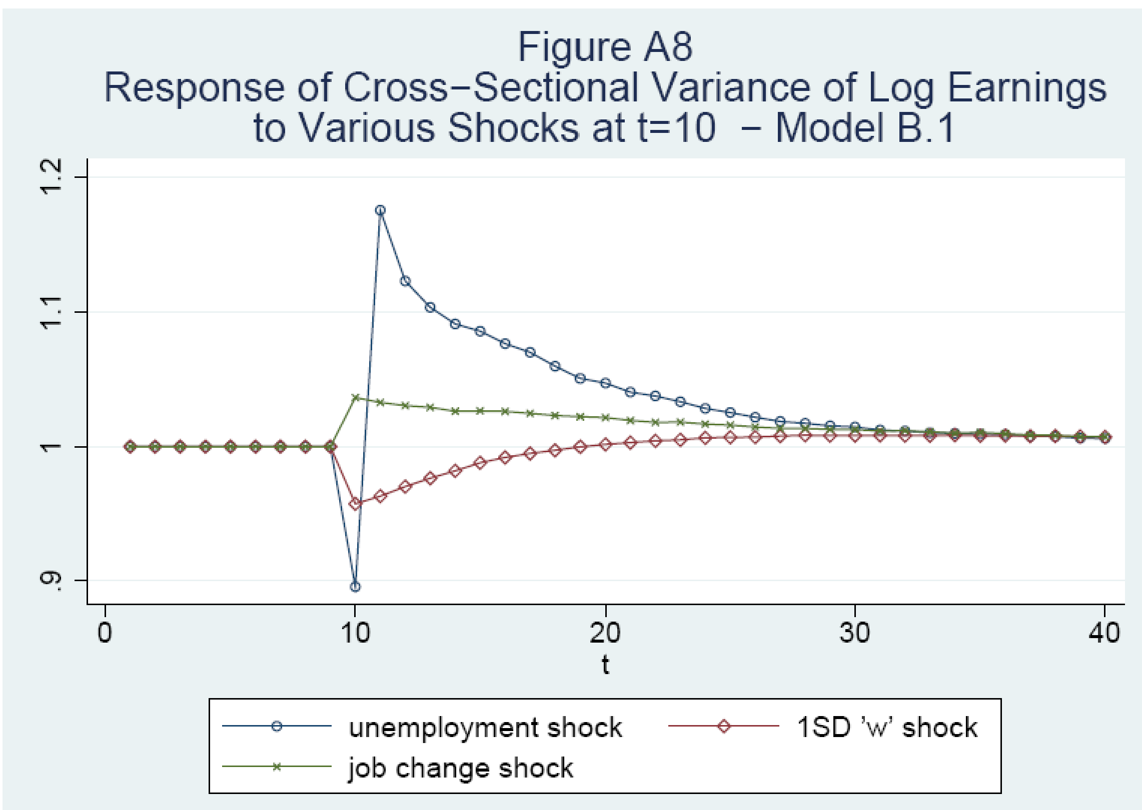


Figure A8
Response of Cross−Sectional Variance of Log Earnings to Various Shocks at t=10 − Model B.1
.9
1
1.1
1.2
0
10
20
30
40
t
unemployment shock
1SD 'w' shock
job change shock

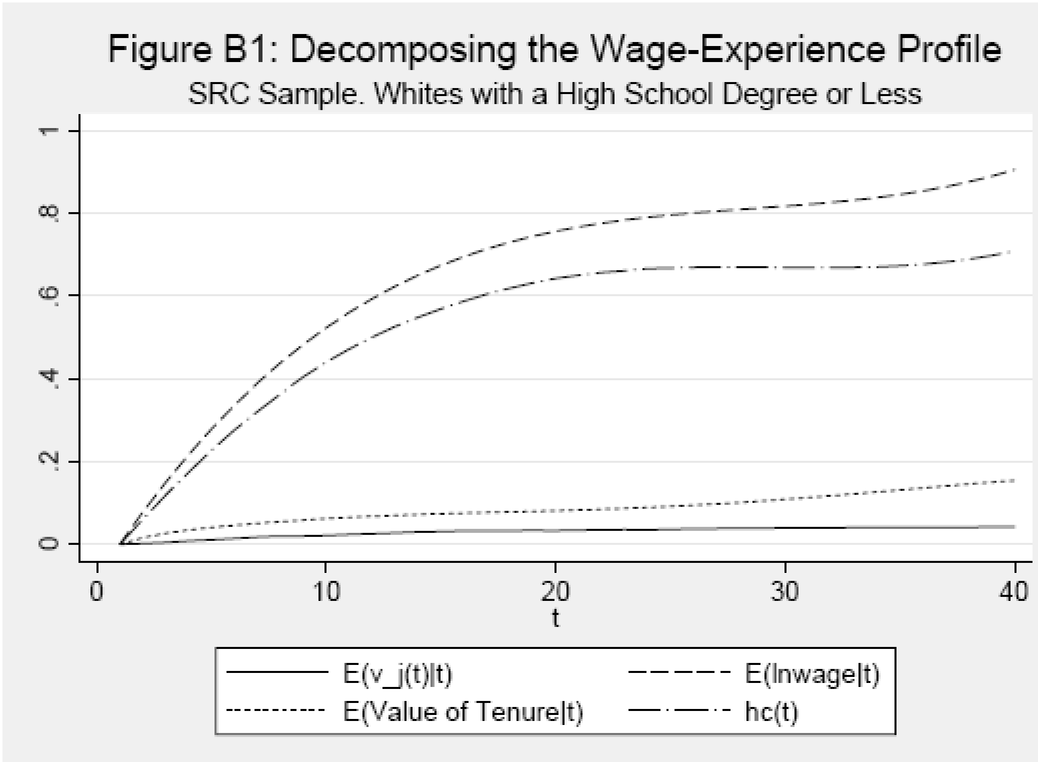


Figure B1: Decomposing the Wage-Experience Profile
SRC Sample. Whites with a High School Degree or Less
0
.2
.4
.6
.8
1
0
10
20
30
40
t
E(v_j(t)|t)
E(lnwage|t)
E(Value of Tenure|t)
hc(t)

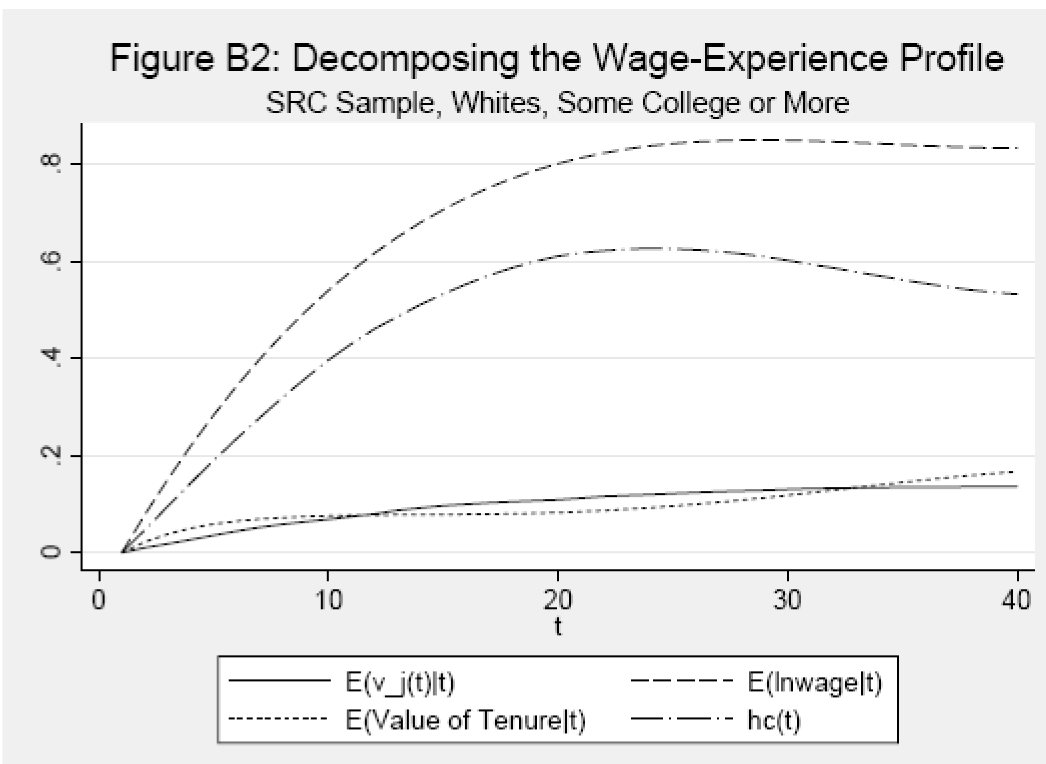


Figure B2: Decomposing the Wage-Experience Profile
SRC Sample, Whites, Some College or More
0
.2
.4
.6
.8
0
10
20
30
40
t
E(v_j(t)|t)
E(lnwage|t)
E(Value of Tenure|t)
hc(t)