Modeling Earnings Dynamics[1]

Joseph G. Altonji, Anthony A. Smith, and Ivan Vidangos
Yale University

November 2, 2006

[1]We are grateful to Richard Blundell, Mary Daly, and Costas Meghir for helpful discussions and to participants in seminars at UC Berkeley, U. of Chicago, the Federal Reserve Bank of San Francisco, U. of Rochester, Yale and conference sessions at the Society of Economic Dynamics (June 2005) and the World Congress of the Econometric Society (August 2005) 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). We are responsible for the remaining shortcomings of the paper.

Abstract

In this paper we use generalized indirect inference to estimate a joint model of earnings, employment, job changes, wage rates, and work hours over a career. Our model incorporates state and duration dependence in several variables, multiple sources of unobserved heterogeneity, job-specific error components in both wages and hours, and measurement error. 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. Shocks associated with job changes make a large contribution to the variance of career earnings and operate mostly through the job-specific error components in wages and hours. Unemployment shocks also make a large contribution and operate mostly through long-term effects on the wage rate.

1 Introduction

In this paper, we build and estimate a joint model of earnings, employment, health status, job changes, wage rates, and work hours over a career. We have three main goals. The first is to identify the key sources of earnings fluctuations over both short and long horizons, and to determine the channels through which they operate. To this end, we quantify the effects on earnings of health shocks, job changes, unemployment, wage shocks, and hours shocks, such as might arise from illnesses or accidents, changes in physical and cognitive skills, changes in the fortunes of employers, the luck of the draw in assignments within a firm, overtime possibilities, temporary layoffs, and so on. Our results are relevant for a number of literatures in labor economics. The second is estimate the contribution of each of these shocks and of permanent unobserved heterogeneity to the variance of lifetime earnings, and to uncertainty about future earnings over both short-term and long-term horizons. We examine differences by education in the variance of shocks to employment, hours, and wages.

Our third purpose is to provide a richer model of earnings for use in studies of consumption and saving, and in dynamic stochastic general-equilibrium models frequently used in macroeconomics and public finance. In macroeconomics, for instance, such models have been used to study the distribution of wealth and consumption in the macroeconomy (Huggett (1996), Krusell & Smith (1998), Castañeda, Díaz-Giménez, & Ríos-Rull (2003), Storesletten, Telmer, & Yaron (2004a)), the costs of business cycles (Imrohoroglu (1989), Krusell & Smith (1999), Storesletten, Telmer, & Yaron (2001a)), and asset pricing (Telmer (1993), Heaton & Lucas (1996), Krusell & Smith (1997), Storesletten, Telmer, & Yaron (2001c)). It is well-known that 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 of earnings uncertainty and the persistence, or predictability, of the earnings innovations (see Deaton 1991, Aiyagari 1994, and Krusell and Smith, 1997).

Almost all of the existing structural studies base their modeling and calibration choices for the earnings process on a substantial body of empirical research on earnings dynamics, including Lillard and Willis (1978), Lillard and Weiss (1979), Hause (1980), MaCurdy (1982), and many subsequent contributions. Some of the studies in this empirical literature use a relatively simple specification in which earnings depends on age, calendar time, a small set of control variables, an unobservable fixed error component, and a stochastic component of earnings that follows an AR(1) process. A more common specification for the stochastic component in the earnings literature is an ARMA (1,1) or ARMA (1,2), with an estimated autoregression coefficient that is usually close to 1. In recent years the models have been expanded to consider differences across education groups and over time in the persistence and variance of income shocks, and to richer stochastic structures that include both unobserved permanent and transitory components, and to explore the implications of fixed heterogeneity in earnings growth (Baker (1997), Haider (2001)).[1] Meghir and Pistaferri (2004) introduce ARCH shocks.

Unfortunately, univariate models do not provide sufficient information to identify and measure the various sources of earnings risk that households face, their relative importance, or their dynamic behavior. With only one indicator, it is very difficult to identify the causes of earnings fluctuations. Even models that incorporate transitory and permanent shocks typically do not provide information about what these shocks represent. Aggregating sources of variation in this way may mask important information regarding earnings uncertainty and risk, since different sources of earnings fluctuations may differ in their degree of predictability or in the dynamic nature of their effects. This type of aggregation also makes it very difficult to think about the potential welfare consequences of specific sources of risk, or about policies, such as unemployment insurance, wage subsidies, or earned income tax credits, that provide insurance against particular sources of risk. Furthermore, the innovation in the univariate representation of a multivariate time series process may be an aggregate of current and past shocks in the multivariate representation, leading to missestimation of 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 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. Recently, Vidangos (2004) has used a panel VAR model to study the dynamics of labor income, wages, hours, and unemployment. Below we discuss Low, Meghir, and Pistaferri's

[1] Guvenen (2004) provides evidence in support of the random growth specification and explores its implications for behavior of consumption and wealth over the life cyle.

(2003) study of earnings risk in the context of a decision model of consumption, employment participation, and mobility.

The models that we consider, in contrast to those mentioned above, incorporate discrete events such as changes in health, job changes, employment loss, interactions between job change and wages, and effects of these discrete events on the variance of wage and hours shocks. We address multiple factor unobserved heterogeneity, state dependence, and measurement error. We deal with complicated missing data problems that stem from the fact that we observe differing portions of the careers of samples members, and we handle initial conditions.

There are two distinct paths that one might take in formulating such a model. The first is to formally model the decision to seek employment or to exit from unemployment, the decision to change jobs, and the choice of work hours, drawing on the substantial literature on each of these questions. There are both advantages and disadvantages to such an approach. The main advantage is that grounding the study of the income process in a utility maximization framework provides a foundation for the use of 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, hours constraints, voluntary quits, and involuntary job changes. For example, only a handful of papers have studied work hours and employment in a model that combines life cycle labor supply considerations, job search, and hours constraints. Estimation of a structural model that is as rich as the model 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 (2003) take an important step in this direction, but work with a simpler model of the earnings process than we do.

The second approach is the development of a statistical model of the income process with little attention to links between the descriptive model and an underlying theory of household decisions and constraints. Both approaches are complicated by the fact that panel data on earnings over a career are incomplete and contain substantial measurement error and by the fact that unobserved heterogeneity in wages, employment, etc., complicates identification of state dependence parameters.

Our model lies between these two extremes but is closer to a descriptive statistical model than a fully specified behavioral one. While it is limited in important ways, it is a major step beyond the simple, typically univariate, processes assumed by almost all of the literature. 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.

The complexity of our model and data makes the estimation problem that we consider a difficult one. The interactions among discrete and continous variables, unobserved heterogeneity and state dependence in multiple equations, the presence of measurement error, a highly unbalanced sample, and the dependence of error component variances on a number of factors together makes the likelihood and the moments of the model extremely complex functions of the model parameters. Conventional maximum likelihood and method of moments approaches are unattractive. Furthermore, we wish to develop an estimation strategy that can be scaled up to even richer models in future research, such as consideration of the household income of individuals when household formation and dissolution is taken into account. For this reason, we use a simulation based approach to estimation.

There is a family of simulation based estimation methods 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. We use a variant called indirect inference (I-I), which was originally proposed by Smith (1990, 1993) and Gourieroux, Monfort, and Renault (1993). 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 methods problematic. Given our model size, derivative free methods are also unattractive. Consequently, we use a modification suggested by Keane and Smith (2003) called Generalized Indirect Inference (G-I-I). The basic idea is to replace step functions associated with binary variables with continuous functions that approximate them arbitrarily well, conditional on choice of a smoothing parameter. Estimation of our model is not straightforward, and a contribution of our research is to explore the feasibility and performance of I-I in general and G-I-I in particular in large models with a mix of discrete and continuous variables.[2]

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

[2]Other recent papers that apply I-I to panel data include Nagypal (2004) and Tartari (2005).

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 analysis of the fit of the model, impulse response functions to various shocks, and variance decompositions. In the final section we briefly summarize our main findings and provide a research agenda.

2 Models of Earnings Dynamics

We work with two classes of models, A and B. In Model A, the wage depends on a job match component that is fixed within a job as well as on an autoregressive component. The hours equation also includes a job specific component. Model B does not include a 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 work with several variants of Model A and Model B.

We begin by listing the equations for health status, employment status transitions, job to job transitions, wages, annual work hours, earnings, initial conditions, the error structure, and measurement error for the model with job match wage components. We then discuss the parts of the model. We consider alternative model specifications in subsequent sections.

2.1 Model A.

The equations of the model and variable definitions are described below. A word about notation. The subscript i refers to the individual, t refers to potential labor market experience, and $j(t)$ refers to the job that i holds at time 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 in t but was unemployed in $t-1$. 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, η_i and ς_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, η, and ς identify the heterogeneity component. We use ρ with appropriate subscripts to denote autoregression coefficients. The ε_{it}^k are iid $N(0, \sigma_k^2)$ where the superscripts k corresponding to the dependent variable.

Health (H):

$$H_{it} = I[\gamma_{\mathbf{0}}^{H} + \gamma_{PE}^{H} PE_{i,t-1} + \rho_H H_{i,t-1} + \delta_{\varsigma}^{H} \varsigma_i + \varepsilon_{it}^{H} > 0] \tag{1}$$

where $H_{it} = 1$ if the individual has a health limitation that affects work, and 0 otherwise, $I[.]$ is the indicator function, PE_{it} is potential experience, with $PE_{it} = PE_{i,t-1} + 1$, ς_i captures fixed heterogeneity in health, and ε_{it}^{H} is an iid health shock.

Employment to Employment Transition (EE)

$$\begin{aligned} E_{it} = I[\gamma_{\mathbf{0}}^{EE} + \gamma_{PE}^{EE} PE_{i,t-1} + \gamma_{PEsq}^{EE} PE_{i,t-1}^{2} + \gamma_{H}^{EE} H_{it} + \gamma_{ED}^{EE} ED_{i,t-1} \\ + \delta_{\mu}^{EE} \mu_i + \delta_{\eta}^{EE} \eta_i + \varepsilon_{it}^{EE} > 0] \text{ given } E_{i,t-1} = 1 \end{aligned} \tag{2}$$

where E_{it} is an employment dummy and $ED_{i,t-1}$ is lagged employment duration and is determined recursively by $ED_{it} = E_{it} \cdot (ED_{i,t-1} + 1)$.

Unemployment to Employment Transition (UE):

$$\begin{aligned} E_{it} = I[\gamma_{\mathbf{0}}^{UE} + \gamma_{PE}^{UE} PE_{i,t-1} + \gamma_{PEsq}^{UE} PE_{i,t\ \ 1}^{2} \\ + \gamma_{H}^{UE} H_{it} + \gamma_{UD}^{UE} UD_{i,t-1} + \delta_{\mu}^{UE} \mu_i + \delta_{\eta}^{UE} \eta_i + \varepsilon_{it}^{UE} > 0] \\ \text{given } E_{i,t-1} = 0, \end{aligned} \tag{3}$$

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):

$$\begin{aligned} JC_{it} = I[\gamma_{\mathbf{0}}^{JC} + \gamma_{PE}^{JC} PE_{i,t-1} + \gamma_{PEsq}^{JC} PE_{i,t\ \ 1}^{2} + \gamma_{TEN}^{JC} TEN_{i,t-1} \\ + \delta_{\mu}^{JC} \mu_i + \delta_{\eta}^{JC} \eta_i + \varepsilon_{it}^{JC} > 0] \cdot E_{it} \cdot E_{i,t-1} \end{aligned}$$

where $TEN_{i,t-1}$ is employer tenure at the previous survey date and

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

Note that $j(t) \neq j(t-1)$ when $JC_{it} = 1$ or $E_{it} = 1$ and $E_{i,t-1} = 0$.

Log Wages:

$$
\begin{aligned}
(4)\qquad wage_{it}^{lat} &= \gamma_0^w + \gamma_X^w X_{it} + \gamma_H^w H_{it} + \upsilon_{ij(t)} + \omega_{it} + \delta_\mu^w \mu_i \\
(5)\qquad \upsilon_{ij(t)} &= (1 - JC_{it} - E_{it}(1 - E_{i,t-1})) \cdot \upsilon_{ij(t-1)} \\
&\quad + (JC_{it} + E_{it}(1 - E_{i,t-1})) \cdot \{[(\gamma_0^\upsilon JC_{it} + \rho_\upsilon \upsilon_{i,j(t-1)})] + \varepsilon_{ij(t)}^\upsilon\} \\
(6)\qquad \omega_{it} &= \rho_\omega \omega_{i,t-1} + \gamma_{1-E}^\omega \frac{1 - E_{it}}{1 + UD_{it}} + \varepsilon_{it}^\omega \\
(7)\qquad wage_{it} &= E_{it} \cdot wage_{it}^{lat}
\end{aligned}
$$

where $wage_{it}^{lat}$ is the "latent" wage, which we define below, X_{it} is a vector of exogenous variables including PE_{it}, $\upsilon_{ij(t)}$ is the job match specific wage component, ω_{it} is an autoregressive component of the latent wage, and $wage_{it}$ is the actual wage rate, which we define as 0 for persons who are unemployed.

Log Annual Work Hours of the Head of Household

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

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

Log earnings

$$
earn_{it} = \gamma_0^e + \gamma_X^e X_{it} + \gamma_w^e (wage_{it}^{lat} - \gamma_0^w - \gamma_X^w X_{it}) + \gamma_h^e (hours_{it} - \gamma_0^h - \gamma_X^h X_{it}) + \varepsilon_{it}^e
$$

Error Components and Initial Conditions:

The fixed person specific error components μ_i, η_i and ς_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 $hours$. [3] 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.

[3] Note that in Model A we constrain the coefficient on μ_i in the wage equation to be the same in all periods. We experiment with freeing up the coefficient on μ in the first period for the wage. We also experimented with allowing the effect of μ_i to grow linearly with experience, but did not obtain sensible results.

The job match hours component $\xi_{ij(t)}$ and the innovation $\varepsilon^{\upsilon}_{it}$ in $\upsilon_{ij(t)}$ are $N(0,\sigma^2_{\xi})$ and $N(0,\sigma^2_{\upsilon})$ respectively. The shocks ε^{EE}_{it}, $\varepsilon^{UE}_{it}, \varepsilon^{JC}_{it}, \varepsilon^{\omega}_{it}, \varepsilon^{\upsilon}_{ij(t)}$, $\varepsilon^{h}_{it}, \varepsilon^{e}_{it}$ are $N(0,\sigma^2_k)$, where $k = EE, UE, JC, \omega, \upsilon, h$, 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

$$\begin{aligned}
\textit{Employment}: &\quad E_{i1} = I[b_0 + \delta^{EE}_{\mu}\mu_i + \delta^{EE}_{\eta}\eta_i + \varepsilon^{EE}_{i1} > 0] \\
\textit{Wages}: &\quad wage^{lat}_{i1} = \gamma^w_0 + \gamma^w_X X_{i1} + \gamma^w_H H_{i1} + \upsilon_{ij(1)} + \omega_{i1} + \delta^{w1}_{\mu}\mu_i \\
&\quad \omega_{i1}\tilde{\ } N(0,\sigma^2_{\omega 1}) \\
\textit{Wage Job Match}: &\quad \upsilon_{ij(1)} \sim N(0,\sigma^2_{\upsilon 1}) \\
\textit{Other Initial Conditions}: &\quad TEN_{i1} = 0, ED_{i1} = E_{i1}, UD_{i1} = 1 - E_{i1}, JC_{i1} = 0.
\end{aligned}$$

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_t for previously employed workers depends on a quadratic in potential experience, a linear function of employment duration ($ED_{i,t-1}$), and the error $\delta^{EE}_{\mu}\mu_i + \delta^{EE}_{\eta}\eta_i + \varepsilon^{EE}_{it}$. 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. The UE transition probability has the same form as EE, with unemployment duration $UD_{i,t-1}$ replacing $ED_{i,t-1}$. We discuss the role of wages momentarily.

The JC equation refers to job to job changes for workers who are employed in both t and $t-1$. We include $TEN_{i,t-1}$ as well as $PE_{i,t-1}$ 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. Less is known about how much of the association between seniority and JC_{it} is causal, because of the difficulty of distinguishing state dependence from heterogeneity in discrete choice models, particularly when data are missing on early employment histories for most sample members (references needed). 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 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.

Most models of labor supply and models of job search would imply that the probability of remaining employed, the probability of changing jobs given that one is employed, and the probability of leaving unemployment would all depend upon wages. As we shall see below, the error components in the EE, UE, and JC equations do depend upon a permanent component that determines wages. The above basic version of model A implicitly assumes that changes in labor supply preferences, exogenous layoffs, and exogenous offer arrivals dominate the movements between employment and unemployment, while intertemporal variation in offered wages around the mean given PE and the μ plays only a small role. This is in keeping with the idea that the processes governing earnings are exogenous with respect to consumer choices, as is assumed in many structural models of consumption and savings, including Gourinchas and Parker (2002), Hubbard, Skinner and Zeldes (1994), Guvenen (2004), and Meghir and Pistaferri (2004) but not Low, Meghir, and Pistaferri (2003). In any event, we experiment with adding the job specific wage component $v_{ij(t-1)}$ to the JC and EE equations. In the case of EE one might argue that the wage rather than the job specific wage component should matter for participation. In some of our specifications, we include the lagged latent wage rate.[4]

When interpreting results for EE and JC, one must keep in mind that our definition of employment refers to the survey date. We undoubtedly miss many spells of unemployment that begin and end between surveys. Due to data limitations, we cannot distinguish

[4] We use the lag rather than the current value to keep the model recursive, given that we allow the wage to depend on current employment status. Furthermore, one might view the decision to stay employed given employment in $t-1$ as influenced by $wage^{lat}_{i,t-1}$. (Alternatively, we could have specified $wage^{lat}_{it}$ to depend on $E_{i,t-1}$ and E_{it} to depend on $wage^{lat}_{it}$.) We have not had much success in adding the wage to the UE equation, perhaps because we observe relatively few unemployment spells.

between 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 and 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. This would not be easy. Note that work hours depend on employment status, and the error component in the hours equation may in part capture the effect on hours of unemployment spells of varying duration.[5]

The wage model (4) is unusual in at least two respects. The first is 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 four components in addition to H_{it}. The first is $\gamma_X^w X_{it}$, which captures the effects of experience, education, race, and secular growth. The second is the heterogeneity component μ_i. The third is a stochastic component ω_{it}, which depends on $\omega_{i,t-1}$, JC_{it}, 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 wages on past wages. We will have more to say about the second mechanism when we turn to model B. The fourth is the job match specific term $\upsilon_{ij(t)}$. When persons experience an unemployment spell 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_0^{υ} in the

[5] It would be 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 want 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 (2002). A quarterly reformulation would substantially increase computation time 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.

case of a job change, and the shock $\varepsilon^{v}_{ij(t)}$. The dependence of $v_{ij(t)}$ on previous values captures the idea that reservation wages are likely to depend positively on $v_{ij(t-1)}$. We have not been successful in estimating models in which the link between $v_{ij(t)}$ and $v_{ij(t-1)}$ differs between JC and unemployment spells, although standard job search models with exogenous layoffs imply that it should.

The equation for $hours_{it}$ includes X_{it} and H_{it}. The error term for $hours$ includes the mobility component η_i and the job specific hours component $\xi^{h}_{ij(t)}$ times E_{it}. We include $\xi^{h}_{ij(t)}$because there is strong evidence that work hours are heavily influenced by a job specific component, which presumably reflects work schedules imposed by employers.[6] A new value of the component is drawn when individuals change jobs. The random error component ε^{h}_{it} picks up temporal variation in overtime, multiple job holding, and in annual hours 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. 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) $wage^{lat}_{it}$ and $hours_{it}$. Abstracting from overtime and variation of wages across jobs for multiple job holders, the coefficients γ^{e}_{w} and γ^{e}_{h} should be close to 1. There are a number of possible explanations for why they might differ from 1, 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. In any event, we freely estimate γ^{e}_{w} and γ^{e}_{h}. We also include a random error term ε^{e}_{it}. As it turns out, $\widehat{\gamma}^{e}_{w}$ and $\widehat{\gamma}^{e}_{h}$ are usually close to 1.

Thus far we have not considered models with an ARCH error structure. However, because the odds of a job change are higher when TEN is low, and job changes are associated with

[6] 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.

innovations in $v_{ij(t)}$ and $\xi_{ij(t)}$ the model implies that the variance of wage changes and earnings changes are state dependent.

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.[7] Instead, we draw 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} account for 35% of $var(\Delta wage_{it}^{*})$, 25% of $var(\Delta hours_{it}^{*})$, and 25% of $var(\Delta earn_{it}^{*})$.[8] 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, but we do perform some analysis of sensitivity to the assumed values.

Note that our assumption of normally distributed, classical measurement error runs counter to evidence that actual reports are a mixture of reports from individuals who give the correct response in households and individuals who respond with error. Furthermore, there is 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.

[7] Add a little more detail after reading the Bound at al. survey and the discussion of measurement error in Altonji and Devereux.

[8] Footnote detailing why we chose these values

2.3 Model B

The main difference between Model A and Model B is in the wage equation. The wage equation for Model B is

$$
\begin{aligned}
(12)\quad wage_{it}^{lat} &= \gamma_0^w + \gamma_X^w X_{it} + \gamma_H^w H_{it} + \omega_{it} \\
\omega_{it} &= \rho_\omega(1+\phi_1[JC_{it} + E_{it}\cdot(1-E_{i,t-1})])\omega_{i,t-1} + \gamma_{JC}^{\omega} JC_{it} + \gamma_{1-E}^{\omega}\frac{1-E_{it}}{1+UD_{it}} \\
&\quad +\delta_\mu^\omega \mu_i + (1+\phi_2[JC_{it} + E_{it}\cdot(1-E_{it})])\varepsilon_{it}^\omega \\
\omega_{i1} &= \theta_\mu^\omega \mu_i + \varepsilon_{i1}^\omega, \quad \text{(initial condition)}
\end{aligned}
$$

Model B excludes the job specific wage component v but introduces the coefficients ϕ_1 and ϕ_2, which 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 the productivity last year. One might expect the degree of dependence to be weaker across jobs than within jobs ($\phi_1 < 0$).

We consider versions of Model B with and without $wage_{i,t-1}$ in the EE and JC equations.

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 lags, the current values of the variables in our model range from 1978 to 1996. We include members of both the SRC and SEO sample. We also include nonsample members who married PSID sample members. All variables refer 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 64, was working, temporarily laid off or unemployed in a given year, was not self employed and had valid data on education. We consider persons who are on temporary layoff at the survey date to be employed. We eliminate a small number of observations in which the individual reports being retired, permanently disabled, a housewife, a student, other, or don't know (see Tables 1c and 1d).[9] Potential experience PE_{it} is $age_{it} - max(years\ of\ education_i, 10) - 5$.

[9] We allow persons to come out of retirement and include future observations following an retirement spell

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. In a future version of the paper we hope to improve upon this.[10] 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?"

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. Very few unemployment spells exceed 1 year, so the error is probably small. The wage $wage^*_{it}$ 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.[11]

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 lack of lead values, or missing data, we use information on 4,628 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 Table 1e). Of course, persons

if the individual is working, temporarily laid off or unemployed. In a future draft we will investigate the sensitivity of our results to this assumption. We also will examine the sensitivity of our results to exclusion of the permanently disabled, which may lead to an understatement of the consequences of health shocks. As reported in Table 1c, 1.3% of the PSID sample reports an employment status of permanently disabled in a given year.

[10] One possibility is to use employment status from 1968-1974 for those employed from 1968 on, and add tenure in 1968 for persons who are employed every year between 68 and 74. Another alternative would be to assume that, conditional on potential experience, the marginal effect of ED goes to 0 as ED exceeds 2 years. A third option would be to apply exactly the same censoring to the simulated data. That is, we could augment the person specific data availability matrix that we employ to ensure that the pattern of available data for the simulated data matches the pattern of available data in the PSID so that in the simulated data ED can be set to tenure when age in the simulated case is equal to age for the first value we see in the corresponding PSID case.

[11] 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 annual hours using the sample of individuals in 1978 for whom the reported wage exceeds \$9.98.

who are present for many years contributed disproportionately to the total of 33,915 person-year observations. The number of observations per year varies from 1200 in 1979 to 2004 in 1991.

The sample is highly unbalanced. As we have already noted, an advantage of simulation based estimators such as G-I-I is that by incorporating the sample selection process into the simulation, one can handle unbalanced data. However, our assumption that observations are missing at random is questionable. There are reasons 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 PSID sample weights to reweight the PSID when evaluating the likelihood function of the auxiliary model. However, there are no sample weights for persons who move into PSID households through marriage. Thus far, we have not explored the sensitivity of our results to alternative assumptions about attrition.[12]

In Table 1a we present the mean, standard deviation, minimum and maximum of the variables used in our structural model, with the exception of the lead values. We also report the same statistics for the person specific mean values. 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. Table 1c reports that prior to imposing our sample selection criteria, 89.23 %, 1.74%, and 5.65% of the person year observations correspond to individuals who are working, on temporary layoff, or unemployed at the survey date, respectively. Table 1d reports the percentage of observations excluded on the basis of employment status by value of PE.

4 Estimation Methodology

We first provide a brief overview of our estimation procedure. We then define the auxiliary model used in the estimation procedure. Finally we discuss the estimates of the measurement error parameters that we use, which are based upon other studies. Note 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 X_{it}. X_{it} consists of a constant, years of education, race, PE_{it}, PE^2_{it}, PE^3_{it}, and a set of year dummies. We

[12]See Fitzgerald, Gottschalk and Moffit (1998) for 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.

then work with the residuals of these variables when estimating the remaining parameters by G-I-I.[13] Our reported standard errors do not account for first stage estimation of the γ_X parameters. We doubt, however, if adjustment would make much difference given that our very large sample size leads to small estimated standard errors for the γ_X.

4.1 Generalized Indirect inference

We estimate the structural model using G-I-I, a method developed by Keane and Smith (2003). This method is a generalization of indirect inference (I-I), an estimation method introduced, under a different name, in Smith (1990, 1993) and later extended by Gourieroux, Monfort, and Renault (1993) and Gallant and Tauchen (1996). Indirect inference exploits the ease and speed with which one can simulate data from complicated structural models such as ours. The basic idea of I-I is to view both the observed data and the simulated data (generated by 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.

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

[13] Note that we include the constants γ_0^w, and 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.

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:

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

where the likelihood function associated with the auxiliary model is, in this case, evaluated using the mth simulated data set $\tilde{y}_m(\beta) \equiv \{y_{it}^m(\beta)\}$. 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.[14] The key idea underlying the consistency of I-I is that, 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 use the likelihood function associated with the auxiliary model to form a metric: 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:

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

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.

[14] 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.

4.2 Generalized Indirect Inference

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 a step function. Step functions arise when applying I-I to discrete choice models because any simulated choice $\tilde{y}_{it}^{m}(\beta)$ is a step function of 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 a step function of β. The non-differentiability of the objective surface in the presence of discrete variables prevents the use of fast gradient-based hill climbing algorithms to optimize the objective surface and requires instead the use of much slower algorithms such as simulated annealing and the downhill simplex method.

To circumvent these difficulties, we use Keane and Smith's (2003) modification to I-I, G-I-I. The basic idea is that the estimation procedures applied to the observed and simulated data sets need not be identical, provided that they both provide consistent estimates of the same vector of pseudo true parameter values. They exploit this idea to smooth the function $\tilde{\theta}(\beta)$, obviating the need to optimize a step function when using I-I to estimate a discrete choice model.

To illustrate the approach, 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 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 G-I-I estimator of β_0.

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

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

Because the latent utility is a 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 unemployment duration and job tenure depend on the history of indicator variables such as employment status and job changes. Since unemployment 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, we only need to replace the indicator functions by their continuous approximations $g(\tilde{u}_{it}^{m}(\beta); \lambda)$. All other variables that depend on β through $g(\tilde{u}_{it}^{m}(\beta); \lambda)$ will as a result also 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.

Although G-I-I permits the use of fast gradient-based hill climbing algorithms, it also introduces bias for any fixed value of λ greater than 0. We use a small value of λ (but still large enough to smooth the objective surface sufficiently); for this value of λ we conduct Monte Carlo experiments to verify that the associated bias in the estimates is small.

We could compute standard errors by computing an estimate of the asymptotic covariance matrix (see Gourieroux, Monfort, and Renault (1993) for its formula). Because the required calculations are somewhat difficult, we use instead a parametric bootstrapping procedure: given consistent estimates of the structural parameters, we repeatedly generate "fake" 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 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.

4.3 The Auxiliary Model

Our auxiliary model consists of a system of seemingly unrelated regressions (SUR) with 7 equations and 29 covariates that are common to all 7 equations. (There are 8 equations and 32 covariates when the health equation is included in the model.) We implement the model under the assumption that the errors follow a multivariate normal distribution with unrestricted covariance matrix. One may write the system as

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

where

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

$$Z_{it} = [Const, PE_{i,t-1}, PE^2_{i,t-1}, ED_{i,t-1}, UD_{i,t-1}, TEN_{i,t-1},$$
$$E_{i,t-1} \cdot E_{i,t-2}, E_{i,t-2} \cdot E_{i,t-3},$$
$$E_{i,t-1} \cdot (1 - E_{i,t-2}), E_{i,t-2} \cdot (1 - E_{i,t-3}),$$
$$JC_{i,t-1} \cdot E_{i,t-1} \cdot E_{i,t-2}, JC_{i,t-2} \cdot E_{i,t-2} \cdot E_{i,t-3},$$
$$wage^*_{i,t-1}, wage^*_{i,t-2}, hours^*_{i,t-1}, hours^*_{i,t-2}, earn^*_{i,t-1}, earn^*_{i,t-2},$$
$$AvgE_i, AvgJC_i, AvgED_i, AvgUD_i, AvgTEN_i,$$
$$wage^*_{i,t-1} - \overline{wage_i}^*, hours^*_{i,t-1} - \overline{hours_i}^*, earn^*_{i,t-1} - \overline{earn_i}^*,$$
$$wage^*_{i,t-1} \cdot PE_{i,t-1}, wage^*_{i,t-1} \cdot PE^2_{i,t-1}, wage^*_{i,t-1} \cdot JC_{it}]' \tag{14}$$

Since Π is a 29×7 parameter vector and Σ is a 7×7 covariance matrix with 28 unique elements, the auxiliary model has 231 parameters. In contrast, Model A.1 has only 37 parameters that we estimate by I-I, not counting the measurement error parameters, which we take from other sources. (As we discuss momentarily, one additional structural parameter is accounted for by the fact that we require the implied variance of the initial wage of the structural model to match the variance of observed wages when $PE_{it} \leq 3$, and the intercept of the initial condition for employment is partially identified using a probit model applied to first three years of employment data.) 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.[15]

Our choice of what to include 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

[15] Again, the fact that we use a misspecified likelihood affects the efficiency of our procedure rather than consistency. It also affects the form of the asymptotic standard error estimators were we to use that approach.

equations of the auxiliary model so we do not need 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 is 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 $PE_{i,t-1}$ and other lagged variables, because the total number of variables would get completely out of hand. We plan to explore dropping this assumption in future work, but it is important to note that 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 $PE_{i,t-1}$, $PE^2_{i,t-1}$, $ED_{i,t-1}$, $UD_{i,t-1}$, and $TEN_{i,t-1}$. Since the model is dynamic and includes state dependence terms in most equations, we include two lags of each dependent variable except $\ln(1+wage^{*\,2}_{it})$. The third is to include the person specific means $AvgE_i$, $AvgJC_i$, $AvgED_i$, $AvgUD_i$, and $AvgTEN_i$. The person specific means help distinguish the role of the fixed heterogeneity components from state dependence. The deviations from means variables $wage^*_{i,t-1} - \overline{wage_i}^*$, $hours^*_{i,t-1} - \overline{hours_i}^*$, and $earn^*_{i,t-1} - \overline{earn_i}^*$, in combination with the inclusion of lags of $wage_{it}$, $hours_{it}$, and $earn_{it}$, serve the same purpose. Finally, we include the interaction terms $wage^*_{i,t-1} \cdot PE_{i,t-1}$, $wage^*_{i,t-1} \cdot PE^{2*}_{i,t-1}$, and $wage^*_{i,t-1} \cdot JC_{it}$ to help capture any nonstationarity in wages.

One disadvantage of our choice for (13) is that the first three observations are lost on the individual. This makes it difficult to identify parameters of the models for the initial wage and employment status. It also makes it difficult to identify variation in the variance of shocks early in a career. In principle, one can add additional equations with 0, 1, or 2 lags to the auxiliary model to accommodate observations with missing data. The cost is added complexity to the 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. A possible explanation for this lies in the fact that our sample from the PSID has relatively few observations on individuals with low levels of experience, and plenty of observations on individuals with PE above 4 or 5. In the next section, we discuss use of the variance of wages when $PE \leq 3$ to help identify parameters of the initial condition for wages. We also discuss the initial condition of employment.

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

In the initial employment equation we estimate the intercept b_0 as $\hat{b}_0 = \hat{b}_0^* \cdot \hat{\sigma}_{E1}$ where $\hat{\sigma}_{E1} = \sqrt{(\hat{\delta}_\mu^{EE})^2 + (\hat{\delta}_\eta^{EE})^2 + 1}$ and $\hat{b}_0^*$ is the coefficient estimate from a Probit of E_{it} on a constant estimated on PSID data for $PE \leq 3$. We use the first three years rather than simply the first year because we have relatively few observations when $PE = 1$. As it turns out, $\hat{b}_0^* = 1.40$.

Similarly, recall that the initial condition for the observed (residual) wage of an employed individual is

$$wres_{i1}^* \equiv wage_{i1}^* - \gamma_0^w - \gamma_X^w X_{i1} = \sigma_{v1} N(0,1) + \delta_\mu^{w1} \mu_i + \sigma_{\omega 1} N(0,1) + \sigma_{mw} N(0,1).$$

Parameters σ_{v1} and δ_μ^{w1} are estimated by G-I-I. The measurement error parameter σ_{mw} is set loosely in accord with the findings of validation studies for the PSID. We estimate $\sigma_{\omega 1}$ from the condition $Var(wres_{i1}^*) = \sigma_{v1}^2 + (\delta_\mu^{w1})^2 + \sigma_{\omega 1}^2 + \sigma_{mw}^2$. More specifically, we approximate the sample moment $\widehat{Var}(wres_{i1}^*)$ by the variance of (residual) wage observations in the PSID corresponding to $PE \leq 3$, and then let $var\_wl_1 \equiv \widehat{Var}(wres_{i1}^*) - \sigma_{mw}^2$, which equals 0.1237 in the sample used to estimate Model A1. We then obtain $\sigma_{\omega 1}$ from $\sigma_{\omega 1}^2 = var\_wl_1 - ((\delta_\mu^{w1})^2 + \sigma_{v1}^2)$. [16]

In the case of model B, the equation for the initial wage is $wage_{i1}^* = \gamma_0^w + \gamma_X^w X_{i1} + \gamma_H^w H_{i1} + \omega_{i1}$ where $\omega_{i1} = \theta_\mu^w \mu_i + \varepsilon_{i1}^\omega$. We choose $\hat{\sigma}_{\omega 1}^2$ (the variance of ε_{i1}^ω) as the solution to $\hat{\sigma}_{\omega 1} = \sqrt{\widehat{var}(w_1) - (\hat{\theta}_\mu^w)^2 - \hat{\sigma}_{mw}^2}$ where $\widehat{var}(w_1)$ is the variance of wage residuals (adjusted for X) when $PE \leq 3$ in the PSID sample. In addition, 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. Thus far, in performing parametric bootstraps we have not accounted for sampling error in the above sample moments of the PSID. We intend to remedy this in future work.

[16] In practice, we also need to impose the restriction that $\sigma_{u1}^2 \geq 0$ during the optimization. We accomplish this by setting

$$\sigma_{u1}^2 = var\_wl_1 - [\frac{\exp(\Gamma)}{1 + \exp(\Gamma)} \cdot (\theta^2 + \sigma_{v1}^2) + (1 - \frac{\exp(\Gamma)}{1 + \exp(\Gamma)}) \cdot (var\_wl_1 - \epsilon)]$$

where $\Gamma = \frac{var\_wl_1 - (\theta^2 + \sigma_{v1}^2) - \in}{\lambda_{w1}}$; and ϵ and λ_{w1} are small numbers ($\epsilon = 0.01$ and $\lambda_{w1} = 0.05$).

4.5 Mechanics of Estimation

Implementation of our estimation methodology requires the choice of a value for the smoothing parameter λ and the number of simulations, M. The larger these parameters are, the smoother the objective function will be. However, large values of λ introduce bias, while large values of M increase computation time. The goal is to find a combination of λ and M that generates sufficient smoothness in the objective function, while keeping bias small and computation time manageable. An appropriate choice must be based on experimentation and will depend, among other things, on the complexity of the interactions among the discrete variables in the structural model. We use $\lambda = 0.05$ and $M = 20$.

The appropriate values of λ and M yield a smooth objective function that allows the use of fast gradient-based optimization algorithms.[17] Starting values for the algorithm are obtained from estimates of a series of "reduced-form equations" that correspond to the equations in the structural model. We experimented extensively with different starting values to address the possibility that the algorithm is finding local optima that are not also global. As it turns out, the objective function displays multiple local optima with respect to the coefficients on potential experience in the EE, UE, and JC probits. The reason for this appears to be that different combinations of parameter values within the probit models can lead to very similar experience profiles. The predictions for the endogenous variables in the model are essentially unaffected, however. We choose the values among the different local maxima that correspond to the largest value of the objective function. Other parameters in the model are not sensitive to the choice of initial guess.

The fact that we have between 35 and 50 parameters, 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. 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.[18]

[17] 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.

[18] The parallelization is implemented using the Message Passing Interface (MPI). We estimate the model

4.6 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 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 parameters does not establish identification of any particular parameter. Consequently, we use simulation to establish local identification and analyze the adequacy of our auxiliary model, and 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 that differ somewhat from the ones presented in the paper, we have informally experimented with varying parameter values to get a sense of how robust identification is to the particular values. In general we have found that identification of most of the parameters is quite robust. As we have already noted, the simulations using the parameter estimates as the true values provide the basis for the parametric bootstrap standard errors that we report below. The parametric bootstrap indicates that for the sample size and demographic structure of the PSID sample, our auxiliary model is informative about all of the model parameters. Furthermore, the degree of bias in the procedure is small for most of our parameters. Appendix I provides a bit more detail about our simulation experiments. Tables A1 and A2 report results for two cases.

5 Empirical Results

First, we discuss the parameter estimates for model A.1, which is a basic version of Model A that excludes health. We also consider alternative 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,

using Fortran 90 on 6 Linux based computers with a total of 12 3.6Ghz Xeon processors, which reduces estimation time from 40-150 hours to about 4-15 hours.

we decompose the variance of wages, hours, and earnings into the contributions of the main types of shocks in our model.

5.1 Parameter Estimates for Model A.1

Columns 1a and 1b of Table 2A present the parameter and standard error estimates for Model A.1. The row headings indicate the variable or error component that parameter estimates correspond to and also list the parameter name. The estimates are grouped by equation, beginning with EE.

The EE coefficients imply an approximately linear positive relationship between PE and the latent variable determining E_{it} conditional on $E_{i,t-1} = 1$. The difference in the latent variable determining EE when $t = 30$ and $t = 0$ is 1.898, but the effect on the odds of a transition is small because the EE probability is very high. Perhaps surprisingly, we find modest negative duration dependence in employment, with a coefficient on $ED_{i,t-1}$ of -0.091. We find a large role for heterogeneity in EE and UE, and to a lesser extent JC. The similar results for simple regressions of E_{it} on $ED_{i,t-1}$ conditional on $E_{i,t-1} = 1$ using simulated and PSID data reported below indicate that in the model heterogeneity reconciles negative duration dependence in the EE equation with positive dependence in the PSID data.

In the UE equation we find a positive, weakly concave relationship between PE and the latent variable for UE transitions. The odds of leaving unemployment rise with $UD_{i,t-1}$, although we have few long spells of unemployment with which to identify this effect, and it is not statistically significant. It should be kept in mind that $UD_{i,t-1}$ measures the number of consecutive surveys that an individual has been unemployed at the survey date and that persons who are on temporary layoff at the survey date are counted as employed. Most work on duration dependence in unemployment spells uses weekly, monthly or quarterly data.[19] Simulations reported below indicate that the model is fully consistent with the strong negative link between UE and $UD_{i,t-1}$ found in the PSID, presumably because of the important role played by permanent heterogeneity.

The job change indicator JC has only a weak relationship with PE but is strongly decreasing in job tenure. The coefficient on $TEN_{i,t-1}$ is $-.1230$, indicating that 10 years of seniority shift the index determining JC by 1.23 standard deviations of the iid job change

[19] References needed.

shock ε_{it}^{JC}. It is noteworthy that we obtain a large negative effect of tenure on JC even after accounting for unobserved heterogeneity.

The "productivity factor" μ has a large positive effect on both EE and UE, with coefficients of .80 in both equations. Since the standard deviations of μ, ε_{it}^{EE}, and ε_{it}^{UE} are all 1, the coefficients imply that the fixed heterogeneity term contributes about two thirds as much variance in the latent variables determining EE and UE as the transitory components. The factor μ has a negative effect on JC, with a coefficient of -.305. 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 enters the EE, UE and JC indices with coefficients of -1.14, .37, and .16, respectively. The sign pattern suggests that η raises the probabilities of transiting between employment states and of moving from job to job without unemployment. Its coefficient is essentially zero in the hours equation.

The system of equations determining wages is the most complex in the model, and the results are very interesting. We begin with the parameters of the autoregressive component

$$\omega_{it} = \rho_\omega \omega_{i,t-1} + \gamma_{1-E}^{\omega} \frac{1-E_{it}}{1+UD_{it}} + \varepsilon_{it}^{\omega}.$$

The standard deviation of the initial condition $\varepsilon_{i1}^{\omega}$ is 0.305. The autoregressive parameter is .905, which is well below unity, but is consistent with a great deal of persistence. The iid shocks $\varepsilon_{it}^{\omega}$ have a standard deviation of .101. 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 other evidence in the literature with which to compare it. For job stayers, the standard deviation of wage changes after adjusting for measurement error is .122.

The coefficient of -.289 (with standard error .0198) on $(1-E_{it})/(1+UD_{it})$ implies that the effect of being unemployed at the survey date on the subsequent wage is -.289/2 or -.1445, a large effect. The marginal effect of the second year is -.0963 or -.289/3. Consequently, two consecutive years of unemployment lead to a reduction in wages of -.256 (.256 $= -.289/2 * .905 - .289/3$). The effect decays relatively slowly given that $\hat{\rho}_\omega = .905$.[20] Thus far, we have not been successful in freeing up the dynamic response of ω_{it} to a spell of unemployment.

[20] In early work we had difficulty pinning down parameters when we experimented with more flexible functions for the effect of UD.

The coefficient δ_{μ}^{w} on μ is .1428 (.0199), which indicates that unobserved permanent heterogeneity contributes only about .020 to the variance of wages in a given year through a direct effect. (Note that μ also has an additional effect on wages through its connection to employment transitions and job changes.) One should keep in mind that these variances are net of the contribution of X_{it}, which contains the important permanent variables education and race.

The parameters of the job match component $\upsilon_{ij(t)}$ are quite interesting. The initial condition $\upsilon_{ij(1)}$ has a standard deviation of .1005. The autoregression parameter ρ_{υ} is .65 and the value of $\hat{\sigma}_{\upsilon}$ is very large, .2372. As we shall see below, the contribution of the job specific component to the variance of wages and earnings is substantial. In addition, $\gamma_0^{\upsilon} = .0388$, which indicates that JC leads to a mean shift of .0388 in υ_{ij} relative to case in which $E_{it} \cdot (1 - E_{it-1}) = 1$.

In the $hours_{it}$ equation, $\hat{\gamma}_E^h$, the effect of E_{it} is .4265. This suggests 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.) We obtain a small, negative wage elasticity. The coefficients on μ and η are .0947 and -.0187 respectively, suggesting only a modest role for individual heterogeneity in annual hours in a given year. The standard deviation σ_h of the iid hours error is .1632. The standard deviation of the job specific error component ξ is large—.1959.

Turning to earnings, we find the coefficients γ_w^e and γ_h^e on the wage and hours to be close to 1. The iid component ε_{it}^e has a standard deviation of .2334. This estimate is sensitive to our particular assumption about measurement error. Furthermore, there is evidence below that suggests that allowing ε_{it}^e to be positively serial correlated will improve the ability of the model to fit the persistence in earnings that is in the data.

5.2 Experiments with Alternative Specifications

5.2.1 Effects of Wages on Employment and Job Changes

Columns 2a and 2b of Table 2A present estimates and standard errors for Model A.2. This model is identical to A.1 except for the fact that we allow $\upsilon_{ij(t-1)}$ to enter the JC equation. It enters with a coefficient of -1.1277. To get a sense of the magnitude, note that the standard deviation of $\upsilon_{ij(t-1)}$ is .2475. Consequently, a 1 standard deviation increase in $\upsilon_{ij(t-1)}$ lowers the JC index by .2791. This is equivalent to the effect of 2 years of seniority.

An increase from 0 to .2475 (1 SD) in the value of $\upsilon_{ij(t-1)}$ lowers the probability of a job change for an individual with 2 years of seniority, ten years of experience, $\mu = 0$ and $\eta = 0$ from .2645 to .1818.

Allowing $\upsilon_{ij(t-1)}$ to influence JC leads to changes in a number of the other parameters. The value of ρ_ω rises to .954. The main change is that the signs on μ in the EE, UE, JC, and *hours* switch. We do not fully understand why this happens. It may be difficult to distinguish links among wages and labor market transitions that reflect the role of job match heterogeneity from links that reflect the role of unobserved productivity.

The effect of PE on JC becomes positive. Note that the negative response of JC to $\upsilon_{ij(t-1)}$ implies an increase in the value of $\upsilon_{ij(t-1)}$ with experience. This may account for the rise in the effect of PE holding $\upsilon_{ij(t-1)}$ constant. However, it would be a mistake to make too much of the change in the PE effect, because we have found that estimates of experience profiles are somewhat sensitive to the specification of model.

In column 3a we add $\upsilon_{ij(t-1)}$ to the EE model. It enters with a small positive coefficient, indicating that $\upsilon_{ij(t-1)}$ lowers the odds that individuals leave employment. The other parameters change very little from those in Model A.2

5.2.2 A Model with Health

Columns 4a and 4b report the point estimates and standard errors of Model A.4. It corresponds to Model A.1, with the health equation (1) added, and health terms included in the EE, UE, *wage*, and *hours* equations, as shown in the specification in Section 2. In thinking about the sign of the health parameters note that H_{it} is 1 for individuals with a health limitation that affects work. The mean of the variable is 0.0374 for persons with PE between 1 and 5 and 0.1293 for persons with PE between 35 and 40.[21] The parameters of the health equation are reported at the bottom of the table. We obtain 1.43 for σ_ς, which implies that ς accounts for 67% of the variance of the composite error term for the health latent variable. As expected, there is also strong state dependence in health and the positive coefficient on PE indicates that health status tends to worsen over time.[22]

As it turns out, the effects of health are small for EE, UE, and very small for wages. The most important effect is on hours worked. The coefficient of -.062 implies that people

[21] In computing these values we include individuals who later are lost due to missing data on lags.

[22] It would be preferable to use age rather than PE in the health equation. However, using PE for all variables simplifies the simulations, because we do not have to keep track of differences between individuals in PE relative to age.

in poor health work about 6% fewer hours per year, everything else equal.

5.2.3 Estimates of Model B

Table 2B reports 4 versions of Model B. Columns 1a and 1b report results for Model B.1, which incorporates the wage model (12), excludes *health*, and excludes $wage_{i,t-1}$ from the JC and EE equations. 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. In Model B, μ affects the first period wage without a constraint, and also affects the drift of the wage process, which is essentially autoregressive.

The results are basically similar to those for Model A.1. There is clear evidence that job changes, whether with or without unemployment, involve substantial wage risk. For job stayers, the coefficient on lagged wages is .948. It is about 80% of this for job changes. At the same time, wage innovations have a standard deviation of .0956 for stayers but are about 3 times larger for persons who have changed jobs ((1+2.0507)*0.0956). One anomaly, however, is that μ has a coefficient of -.09 in the initial condition for wages. The estimate of δ_{μ}^{w} is .0148.

Model B.2 adds the lagged wage to the JC equation, while Model B.3 adds it to the EE and JC equations. The lagged wage enters with a positive coefficient in the EE equation and a negative coefficient in the JC model. The coefficient on μ in the initial wage becomes substantially more negative: -.33 in B.2 and -.22 in B.3. The results for *Health* in column 4a are very similar to those for Model A.

5.3 Evaluating the fit of the Model:

We simulate careers for 46,280 individuals using the parameter estimates for Model A.1 and Model B.1—10 for each individual in the PSID sample. From each simulated career we select data so that the temporal pattern matches that of a corresponding PSID case. Note that in all cases the simulated variables incorporate measurement error. We provide an informal examination of 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.

Thus far, we have not conducted formal tests of whether simulated values are statistically different from the sample values.

Predicted and Actual Mean and Standard Deviations of Key Variables, by Potential Experience The top panel of Table 3 compares the standard deviation of $wage^*_{it}$ in the PSID (row 1) to the corresponding standard deviation based on data simulated from model A.1 (row 2) and model B.1 (row 3). The column labeled *Overall* is for all experience levels. The other columns report results for the level of PE reported in the top row. For each indicated level of PE, the results combine data for PE-1, PE, and PE +1. The model matches the overall standard deviation closely. The PSID data show an increase in the SD of wages from .35 when t=5 to .40 when t=20, while the simulated data for model A.1 shows a small rise followed by a small decline. The values for model B.1 rise from .38 when t=5 to .42 when t=40.

The second panel presents results for earnings. The actual and simulated SD are 0.57 and 0.58 respectively (0.59 for model B.1). However, there is an erratic pattern in the data that is not matched by the model, which displays a smooth hump shape pattern. The values are close except when t=5 and t=40. Model B1 matches more closely at the endpoints but misses a bit between t=10 and t=30. Model A.1 overpredicts the SD of hours by a small amount, with the exception of an underprediction of .34 when t=40.

The fourth and fifth panels of the table report the mean of E_t and the mean of JC_t conditional on $E_t = 1$ and $E_{t-1} = 1$. The overall mean for E is .97 in the data and .96 based on the model. Models A.1 and B.1 tracks the values of JC reasonably closely.

The sixth panel reports on simulated and actual means of EE transitions, which match reasonably well. However, in the seventh panel, the model tends to underpredict exits from unemployment (UE). Overall, the actual and simulated means of UE are .74 and .69 (.70 for B.1), and the model does not track the experience profile well. Model B.1 does somewhat better than A.1 despite the fact that the EE and UE equations of A.1 and B.1 are the same.

The final three panels of the table examine 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 0.22 years. The models overpredict ED by a substantial amount. This probably reflects the fact that we use TEN as the initial value for ED when an individual first enters the sample. (See the Data Section.)

5.3.1 Comparison of Regression Relationships Among Key Variables

Tables 4A (i)-(iv) report a series of regressions. The "a" models are based on the PSID sample and the "B" models are based on data simulated from the estimated Model A.1 (Tables 4B (i)-(iv) display similar results for model B.1). Columns 1a and 1b of Table 4A(i) report regressions of E_t on PE_{t-1}, $PE_{t-1}^2/100$, 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 minor differences in the experience profiles. The coefficient on ED_{t-1} is .0024 in the PSID and .0018 in the simulation. Note that the positive coefficient on ED_{t-1} is opposite in sign to what we obtain for the structural model. This indicates that the effect of heterogeneity outweighs the effect of state dependence in determining the pattern of persistence in E_{it}.

Columns 2a and 2b report results for a version of the UE equation. The differences in the coefficients on the experience profile do not seem very large given the standard errors in the PSID sample. 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.

Table 4A (ii) examines the dynamics of wages. When only one lag is included, the coefficient on $wage_{t-1}^*$ is .885 in the PSID and .899 in the simulation. When two lags are included, the sums of the coefficients are very close but there is a difference in the coefficient pattern. The coefficient on JC_t is small and negative in the PSID, and small and positive in the simulated data.

Table 4A (iii) examines hours dynamics. The results in the simulated and actual data match reasonably closely, although the sum of the lagged coefficients on hours is about .6 in the simulated data and about .55 in the actual data (columns 2a and 2b). The wage coefficient is essentially 0 in the actual data, and -.0172 in the simulated data—a very close correspondence.

Finally, in Table 4A (iv) we report earnings regressions. Note that in the model all of the dynamics in earnings stem from dynamics in the wage and in hours. In the PSID data, earnings are more persistent than is implied by the model. The coefficient on $earn_{i,t-1}^*$ is .829 in the actual data and .6569 in simulated data. We suspect that this problem could be fixed by allowing for serial correlation in the component of earnings not explained by the straight time hourly wage and hours, but we have not investigated this yet. There is also some difference between the data and the model in the coefficients on $wage_t^*$ and $hours_t^*$.

Overall, we are encouraged by the match between the model and the data, although there is room for improvement. In a future draft, we hope to provide an overall goodness of fit test of the model.

5.4 Impulse Response Functions

Figures 1a-1c report impulse responses to shocks that occur when PE = 10. They are constructed as follows. First, we simulate a large number of cases up to PE=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 earnings, wages, and hours. The base case represents the mean of the simulated paths in the absence of the specified shock intervention in period 10.

The line with circles in Figure 1a reports the response of the mean of $earn_{it}$ to a one standard deviation shock to $\varepsilon_{it}^{\omega}$, the error term in the autoregressive component of wages, when PE=10. Earnings rise by about .106, and the effect slowly decays. The pattern for earnings closely mirrors the response to wages, which reflects the fact that the coefficient on the wage is essentially 1, and the response of hours to the wage is small.

The line with triangles shows the effect of becoming unemployed when $PE = 10$ and is very interesting. The log of earnings drops by -.57. It recovers by about two thirds after one year, and then slowly returns to the base case. The pattern for earnings is the combination of a drop of -.393 in log hours, which fully recovers after two periods, and an initial drop of about -.12 in the wage, followed by a second, smaller drop to about -.168, and then a slow recovery. The small second drop reflects the fact that in some instances the unemployment spell lasts for more than one survey. As we have already noted, the dynamics of the wage response to unemployment is governed by $\hat{\rho}_{\omega}$, which .905, and it would be desirable to free up the response to some extent. 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. An advantage of our model is that we examine effects that operate through wages and hours separately.

Finally, the figures report the mean effect of a job change. In this case, JC is set equal to one in period 10. The results show a small, positive, and highly persistent effect on earnings and the wage.

Figures 2a-2c look in more detail at the effect of job changes. The line with triangles measures the effect of a job change accompanied by a one standard deviation increase (.237) in the error term of the process for $v_{\iota j(t)}$. As one can see, the shock leads to an increase in log earnings of about .23, and the increase is highly persistent. Similarly, the effect of a job change accompanied by a one standard deviation increase in the job specific hours component is also large and highly persistent. In the case of hours, the persistence stems from the fact that when PE is 10, the job changes without unemployment and the unemployment spells that trigger job changes are infrequent.

In a future draft, we hope to report on the degree to which the impulse responses depend upon PE. For example, one might expect the persistence of shocks to v to be lower early in the career, given the job change rates are much higher.

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, and the permanent heterogeneity components μ and η. To do this, we first compute the variance in the means 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 and take the drop in the variance relative to the base case as the contribution of the particular type of shock. Since the model is nonlinear, the contributions do not sum to 100%. We have normalized them to sum to 100. We report results for the log variables, but decompositions of the level of earnings, the wage, and hours are similar. We focus on Model A.1.

The results are in Table 5A(i). The first row refers to $earn_{it}$. The earnings shocks ε^{e}_{it} account for 5% of the variance in the mean of lifetime earnings even though they account for about 20% of $var(earn_{it})$ in a given year (Table 5A(ii)). The reason for the relatively small contribution is that the shocks are iid rather than persistent. Similarly, the value in column II indicates that iid hours shocks ε^{h}_{it} contribute only 2.7% of the variance in mean lifetime earnings while accounting for between 9.5 and 11.0% in a given year (Table 5A(ii)). One can easily self insure against these shock categories. In contrast, in column III, the initial condition $\varepsilon^{\omega}_{i1}$ and the iid shocks to ω_{it} are together responsible for 16.1% of the variance in lifetime earnings. The earnings results reflect the fact that these shocks are responsible for 24.5% of the variance in lifetime wages. They contribute little to the variance in hours because the response of hours to wages is small. Column IV shows the collective impact of job specific hours and wage components, unemployment spells, and job changes. All together, the shocks account for 42.3%, 48.3%, and 68.3% of the variance in lifetime earnings, wages, and hours. 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 job changes. The estimates, which sum to 100%, are reported in columns VII, VIII, IX, and X. Job specific hours and wage shocks are about equally important for earnings. Wage shocks dominate for wages, with employment shocks also playing a substantial role. For hours, job specific hours shocks dominate.[23]

The estimates in column V indicate that the mobility preference component η does not play much of a role. However, μ accounts for 35.4% of the variance in lifetime earnings, 26.6% of the variance in lifetime wages, and 26.2% of the variance in work hours. It accounts for much less of the variances in a given year.

Thus far we have not analyzed the contribution of health shocks. However, it is clear from the relatively small impact of health, at least as we have measured it, on employment, wages, and hours that health shocks account for very little of the variance in lifetime earnings.

In a future draft, we will report on similar decompositions for other versions of our model. Results for Model B.1 are in Table 5B. We will get at uncertainty a different points in a career by calculating the contributions of the shocks between t and t+5 to the variance in earnings over the same period, for various values of t.

6 Conclusions and a Research Agenda (very preliminary)

In this paper we use generalized indirect inference to estimate models of earnings dynamics. Our models incorporate state dependence in employment, unemployment, job changes, and wages. They also allow for multiple sources of unobserved heterogeneity, measurement error, and in one version of our model we allow for job-specific error components in both wages and hours. These turn out to play an important role in the variance of earnings over the lifetime. A big advantage of our approach compared to previous studies that have used maximum

[23] A few of the estimated variances contributions are negative. We believe this is possible given nonlinearity in the model but are still investigating.

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.

Our results are still very preliminary. We are still experimenting with alternative model specifications that might better capture the economics underlying the sources of earnings dynamics. We will not attempt to summarize every result here. Particularly noteworthy is our finding that the short-term costs of unemployment are dominated by hours and the long-term costs by wages, that job changes and unemployment spells contribute in a major way to earnings variance by leading to large changes in job-specific components of wages and hours as well as through their direct effects on wages and hours.

Our short-term research agenda includes estimating the model separately for high and low education groups, allowing the variances of iid shocks in the model to change over time and with experience. A separate paper, Vidangos (2006), studies the implications of a related multi-equation model of household income for precautionary behavior and welfare within the context of a lifecycle consumption model. That paper considers additional sources of variation in net income such as disability and medical expenditures. 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.

Further down the road, 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. A much more ambitious extension is to construct a model of the household income of an individual that incorporates marriage, divorce, and death of a spouse. This will be pursued in separate work.

Finally, we wish to emphasize that our analysis does not permit one 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 transfers) as in Vidangos (2006) gets partially at the question of insurance. However, 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 (2004), Cunha, Heckman, and Navarro (2005), and Cunha and Heckman (2006) illustrates the latter approach.

7 References (Incomplete)

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.

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.

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

Baker, M. and Solon, G. (1999). "Earnings dynamics and inequality among Canadian men, 1976-1992: evidence from longitudinal income tax records", NBER working paper no. 7370.

Blundell, R., L. Pistaferri, and I. Preston (2004). Consumption Inequality and Partial Insurance. Technical Report WP04/28, Institute for Fiscal Studies.

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.

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, April). "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.

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.

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. (2005). "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?", Working Paper, University of Rochester.

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.

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.

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

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. (2003). "Wage risk, employment risk, and precautionary saving", 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. (2004)."Learning-by-Doing versus Learning About Match Quality: Can We Tell Them Apart?" unpublished paper, Northwestern University.

"The Cost of and Political Support for Employment

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.

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.

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. (2001c). "Asset Pricing with Idiosyncratic Risk and Overlapping Generations", C.E.P.R. Discussion Papers n. 3065.

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.

Tartari, M. (2005). "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. (2006). "Household Welfare, Precautionary Saving, and Social Insurance under Multiple Sources of Risk", unpublished manuscript, Yale University.

8 Appendix 1 - Monte Carlo Evidence on Local Identification and Estimation Bias

We use Monte Carlo experiments extensively to verify that our auxiliary model is informative about all 'structural' parameters, and to assess estimation bias. Specifically, we simulate data using our structural model and a vector of structural parameters values, apply our estimation procedure to the simulated data, and compare the estimation results to the specified parameter values. We typically use 50 replications of this procedure, generating each time a different set of simulated data using the same set of hypothesized structural parameters. We then average our estimated parameters across the 50 replications, and compare the average values to the vector of 'true' parameters used to generate the simulated data. This procedure allows us to check for the presence of bias in the estimation of any particular parameter. Generally, we do not find evidence of significant bias in any of the model parameters.

In order to check for local identification, we experiment extensively with varying (locally) the initial guess of the parameter vector in the optimization. We use these experiments to investigate whether the objective function has flat regions near the solution, or multiple global optima. We find that most parameters in the model are strongly identified. The only exception, as mentioned in the body of the paper, relates to parameters in the probit models that determine the experience profiles of employment and job changes. In this case, slightly different combinations of parameter values can sometimes lead to similar experience profiles, and hence to similar properties of the simulated data and similar values of the objective function. In this case, our procedure is to choose the values among the different local optima that correspond to the largest value of the objective function. The differences in the parameter values across local optima, however, are small, and do not affect any implications for the behavior of the endogenous variables in the model.

Tables A1 and A2 present results from Monte Carlo experiments for Model A.1 and B.1, respectively. The experiments use the same sample size and demographic structure as that of our PSID sample. The initial guess used for the results presented were the original vector of parameters as the initial guess for the optimization. Estimation of wage innovation variance parameters when rho is allowed to depend on indicator $JC_t + E_t(1 - E_{t-1})$.

9 Appendix 2. A Moment Condition for ϕ_1 and ϕ_2 in Model B.

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

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

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

(15)
$$\begin{aligned}&\omega^{obs}_{it}-\rho_\omega(1+\phi_1I_{it})\omega^{obs}_{i,t-1}-\gamma^\omega_{JC}JC_{it}-\gamma^\omega_{1-E}\frac{1-E_{it}}{1+UD_{it}}\\ &=\delta^\omega_\mu\mu_i+(1+\phi_2I_{it})\varepsilon^\omega_{it}+m^w_{it}-\rho_\omega(1+\phi_1I_{it})m^w_{i,t-1}\end{aligned}$$

Denote the left-hand side of 15 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 $R(\delta^\omega_\mu,\rho_\omega,\phi_1,\phi_2,\sigma_\omega,\sigma_{mw})$:

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

and

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

Assuming that $(\delta^\omega_\mu)^2[var(\mu_i|I_{it}=1)-var(\mu_i|I_{it}=0)]$ 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:

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

and

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

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

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

and

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

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

Table 1a
Descriptive Statistics - PSID Sample

Variable	Obs.	Mean	StDev	Min	Max
E_t	33,915	0.97	0.18	0	1
JC_t	33,915	0.08	0.28	0	1
ED_t	33,915	11.58	7.45	0	42.25
UD_t	33,915	0.05	0.31	0	8
TEN_t	33,915	9.34	7.80	0	42.25
log $wage_t$	32,816	2.73	0.49	1.25	4.98
log $hours_t$	33,915	7.73	0.29	5.30	8.34
log $earn_t$	33,915	5.78	0.65	-0.81	8.45
w_t [(a)]	32,816	0.03	0.39	-2.00	2.22
h_t [(a)]	33,915	0.04	0.28	-2.51	0.87
e_t [(a)]	33,915	0.05	0.57	-8.91	2.43
AVG_i log $wage_t$	4,620	2.60	0.46	1.35	4.76
AVG_i log $hours_t$	4,628	7.70	0.22	6.14	8.30
AVG_i log $earn_t$	4,628	5.57	0.64	1.96	7.77
AVG_i w_t [(a)]	4,628	-0.01	0.32	-1.88	2.01
AVG_i h_t [(a)]	4,628	0.02	0.21	-1.53	0.62
AVG_i e_t [(a)]	4,628	-0.03	0.52	-3.33	1.58

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	33,915	19.33	8.80	4	40
Education (years)	33,915	12.94	2.38	6	17
Black	33,915	0.29	0.45	0	1
Other Nonwhite	33,915	0.01	0.11	0	1
Calendar Year	33,915	1987.5	5.25	1978	1996

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

Table 1c
Composition of PSID sample before sample selection based on employment status.

Emp. Status	Fraction
Working	89.23
Temp. Laidoff	1.74
Unemployed	5.65
Retired	0.25
Perm. Disabled	1.31
Housewife	0.17
Student	1.13
Other	0.52

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 1d
Percentage of observations excluded based on employment status.

PE	Fraction	PE	Fraction	PE	Fraction	PE	Fraction
1	18.2 [(a)]	11	2.4	21	2.3	31	2.9
2	9.8	12	2.7	22	2.4	32	4.1
3	6.6	13	2.8	23	3.2	33	4.2
4	4.8	14	3.3	24	3.3	34	4.2
5	4.4	15	2.3	25	3.2	35	3.7
6	3.3	16	2.6	26	3.7	36	5.5
7	2.8	17	2.8	27	3.7	37	5.7
8	2.8	18	2.4	28	3.6	38	5.7
9	2.3	19	2.7	29	3.8	39	8.1
10	2.2	20	2.9	30	4.2	40	9.5

The table presents the percentage of observations excluded, based on employment status at the survey date, for each value of potential experience (PE).
(a) 14.6 are students.

Table 1e
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.

Table 2A
Point Estimates and Standard Errors - Models A.1 through A.4

		Model A.1		Model A.2		Model A.3		Model A.4	
		1a	1b	2a	2b	3a	3b	4a	4b
Equation / Variable	Parameter	Point Est.	S.E.	Point Est.	S.E.	Point Est.	S.E.	Point Est.	S.E.
E-E									
(cons)	γ^{EE}_{0}	2.7738	(0.1581)	3.4200		3.7013		3.1944	
(PE_{t-1})	γ^{EE}_{PE}	0.0636	(0.0162)	0.0706		0.0433		0.0986	
(PE^2_{t-1})/100	γ^{EE}_{PEsq}	-0.0011	(0.0348)	-0.0006		-0.0053		-0.0901	
(ED_{t-1})	γ^{EE}_{ED}	-0.0910	(0.0111)	-0.1041		-0.1019		-0.1150	
(H_t)	γ^{EE}_{H}							-0.0251	
(υ_{t-1})	δ^{EE}_{υ}					0.1391			
(w_{t-1})	γ^{EE}_{w}								
(μ)	δ^{EE}_{μ}	0.8036	(0.0789)	-1.1888		-1.3368		0.8563	
(η)	δ^{EE}_{η}	-1.1365	(0.1040)	-1.3103		-1.1822		-1.5741	
U-E									
(cons)	γ^{UE}_{0}	0.9273	(0.2688)	1.8195		1.9510		1.1260	
(PE_{t-1})	γ^{UE}_{PE}	0.0103	(0.0271)	0.0520		0.0461		0.0086	
(PE^2_{t-1})/100	γ^{UE}_{PEsq}	-0.0180	(0.0762)	-0.0016		-0.1911		-0.0098	
(UD_{t-1})	γ^{UE}_{UD}	0.0733	(0.0544)	0.1502		0.1380		-0.0635	
(H_t)	γ^{UE}_{H}							-0.0122	
(μ)	δ^{UE}_{μ}	0.8004	(0.1373)	-1.3054		-1.1736		0.7078	
(η)	δ^{UE}_{η}	0.3691	(0.1207)	-0.0866		-0.0183		0.0949	
JC									
(cons)	γ^{JC}_{0}	-0.6098	(0.0707)	-0.7497		-0.7963		-0.6542	
(PE_{t-1})	γ^{JC}_{PE}	0.0029	(0.0096)	0.0388		0.0357		-0.0122	
(PE^2_{t-1})/100	γ^{JC}_{PEsq}	-0.0798	(0.0263)	-0.0018		-0.1721		-0.0453	
(TEN_{t-1})	γ^{JC}_{TEN}	-0.1230	(0.0079)	-0.1330		-0.1125		-0.0959	
(υ_{t-1})	δ^{JC}_{υ}			-1.1277		-1.2235			
(ξ_{t-1})	δ^{JC}_{ξ}								
(μ)	δ^{JC}_{μ}	-0.3051	(0.0411)	0.2331		0.2816		-0.4765	
(η)	δ^{JC}_{η}	0.1558	(0.0270)	0.1201		0.1375		0.2002	
Wage									
(H_t)	γ^{w}_{H}								
(μ)	δ^{w}_{μ}	0.1428	(0.0199)	0.0955		0.1252		0.1341	
(JC_t)	γ^{υ}_{0}	0.0388	(0.0060)	0.0166		0.0087		0.0447	
(υ_{t-1})	ρ_{υ}	0.6522	(0.0377)	0.4661		0.4748		0.6695	
(ε^{υ})	σ_{υ}	0.2372	(0.0079)	0.2475		0.2609		0.2412	
($\varepsilon^{\upsilon}_{1}$)	$\sigma_{\upsilon 1}$	0.1005	(0.0005)	0.1003		0.1035		0.1000	
(ω_{t-1})	ρ_{ω}	0.9051	(0.0160)	0.9544		0.9598		0.9153	
($1-E_t$)	γ^{ω}_{1-E}	-0.2894	(0.0198)	-0.3880		-0.4100		-0.2965	
(ε^{ω})	σ_{ω}	0.1014	(0.0026)	0.0829		0.0727		0.0980	
(ε^{ω}_{1})	$\sigma_{\omega 1}$ (i)	0.3053	(0.0114)	0.3233		0.3120		0.3094	
Hours									
(cons)	γ^{h}_{0}	-0.3945	(0.0078)	-0.3706		-0.3713		-0.3777	
(E_t)	γ^{h}_{E}	0.4265	(0.0073)	0.4140		0.4136		0.4179	
	σ_{ξ}	0.1959	(0.0052)	0.2007		0.2025		0.2005	
(w_t)	γ^{h}_{w}	-0.0852	(0.0113)	-0.0080		0.0042		-0.0534	
(H_t)	γ^{h}_{H}							-0.0617	
(μ)	δ^{h}_{μ}	0.0947	(0.0091)	-0.0640		-0.0626		0.0651	
(η)	δ^{h}_{η}	-0.0187	(0.0079)	-0.0347		-0.0279		-0.0398	
(ε^{h})	σ_{h}	0.1632	(0.0015)	0.1626		0.1623		0.1618	
Earnings									
(cons)	γ^{e}_{0}	-0.0002	(0.0019)	0.0004		0.0007		-0.0005	
(w_t)	γ^{e}_{w}	1.0288	(0.0040)	1.0370		1.0401		1.0304	
(h_t)	γ^{e}_{h}	1.0122	(0.0071)	0.9954		0.9840		1.0071	
(ε^{e})	σ_{e}	0.2334	(0.0013)	0.2262		0.2227		0.2309	
Health									
(cons)	γ^{H}_{0}							-3.3971	
(PE_{t-1})	γ^{H}_{PE}							0.0382	
(H_{t-1})	ρ_{H}							0.4659	
(ς)	δ^{H}_{ς}							1.4339	

The table presents point estimates and standard errors for model specifications A.1 through A.4. Point estimates were obtained by Generalized Indirect Inference, unless indicated otherwise. Parametric bootstrap standard errors are in parentheses. Bootstraps are based on 50 replications.

(i) Point estimate obtained using additional moment conditions. See discussion in Section 4.

Table 2B
Point Estimates and Standard Errors - Models B.1 through B.4

		Model B.1		Model B.2		Model B.3		Model B.4	
		1a	1b	2a	2b	3a	3b	4a	4b
Equation / Variable	Parameter	Point Est.	S.E.	Point Est.	S.E.	Point Est.	S.E.	Point Est.	S.E.
E-E									
(cons)	γ^{EE}_{0}	2.5000		3.5000		2.9274		2.9372	
(PE_{t-1})	γ^{EE}_{PE}	0.0677		0.0728		0.0129		0.0464	
(PE^2_{t-1})/100	γ^{EE}_{PEsq}	-0.0526		-0.0660		0.0320		-0.0133	
(ED_{t-1})	γ^{EE}_{ED}	-0.0679		-0.1110		-0.0537		-0.0795	
(H_t)	γ^{EE}_{H}							-0.0161	
(w_{t-1})	γ^{EE}_{w}					0.0607			
(μ)	δ^{EE}_{μ}	0.8681		1.1882		0.7688		1.1774	
(η)	δ^{EE}_{η}	-0.9234		-1.3643		-0.8848		-0.6597	
U-E									
(cons)	γ^{UE}_{0}	1.0755		2.0688		1.6869		1.7931	
(PE_{t-1})	γ^{UE}_{PE}	0.0323		-0.0116		-0.0571		0.0059	
(PE^2_{t-1})/100	γ^{UE}_{PEsq}	-0.0751		-0.0509		0.1239		-0.0326	
(UD_{t-1})	γ^{UE}_{UD}	0.0932		0.2062		0.3520		0.2197	
(H_t)	γ^{UE}_{H}							-0.0131	
(μ)	δ^{UE}_{μ}	0.8844		1.1541		1.1394		1.1777	
(η)	δ^{UE}_{η}	0.2576		0.1337		0.5495		0.4825	
JC									
(cons)	γ^{JC}_{0}	-0.6275		-0.5363		-0.7028		-0.6732	
(PE_{t-1})	γ^{JC}_{PE}	0.0047		0.0108		0.0221		0.0087	
(PE^2_{t-1})/100	γ^{JC}_{PEsq}	-0.0938		-0.1024		-0.1549		-0.0791	
(TEN_{t-1})	γ^{JC}_{TEN}	-0.1162		-0.1485		-0.1150		-0.1279	
(w_{t-1})	δ^{JC}_{w}			-0.4171		-0.3968			
(μ)	δ^{JC}_{μ}	-0.2766		-0.1660		-0.1780		-0.2887	
(η)	δ^{JC}_{η}	0.2508		0.1721		0.2588		0.1542	
Wage									
(H_t)	γ^{w}_{H}							-0.0007	
(ω_{t-1})	ρ_{ω}	0.9480		0.9785		0.9671		0.9552	
	φ_1	-0.1943		-0.2221		-0.1992		-0.2063	
(JC_t)	γ^{ω}_{JC}	0.0222		0.0309		0.0298		0.0216	
($1-E_t$)	γ^{ω}_{1-E}	-0.2732		-0.3999		-0.3620		-0.2835	
(μ)	δ^{ω}_{μ}	0.0148		0.0080		0.0093		0.0130	
(ε^{ω})	σ_{ω}	0.0956		0.0797		0.0871		0.0931	
(ε^{ω}_{1})	$\sigma_{\omega 1}$ (i)	0.3389		0.1119		0.2736		0.3250	
	φ_2 (i)	2.0507		2.5921		2.3220		2.1202	
(μ)	θ^{ω}_{μ}	-0.0939		-0.3334		-0.2210		-0.1345	
Hours									
(cons)	γ^{h}_{0}	-0.4047		-0.3608		-0.3704		-0.3790	
(E_t)	γ^{h}_{E}	0.4325		0.4080		0.4135		0.4202	
(w_t)	γ^{h}_{w}	-0.1475		-0.0022		-0.0619		-0.1062	
(H_t)	γ^{h}_{H}							-0.0661	
(μ)	δ^{h}_{μ}	0.1795		0.1712		0.1754		0.1534	
(η)	δ^{h}_{η}	0.0584		0.0528		0.0452		0.1007	
(ε^{h})	σ_{h}	0.1805		0.1815		0.1809		0.1808	
Earnings									
(cons)	γ^{e}_{0}	0.0001		0.0031		0.0011		-0.0001	
(w_t)	γ^{e}_{w}	1.0353		1.0422		1.0359		1.0357	
(h_t)	γ^{e}_{h}	1.0121		0.9721		0.9978		1.0002	
(ε^{e})	σ_{e}	0.2316		0.2200		0.2271		0.2291	
Health									
(cons)	γ^{H}_{0}							-3.3836	
(PE_{t-1})	γ^{H}_{PE}							0.0370	
(H_{t-1})	ρ_{H}							0.4674	
(ς)	δ^{H}_{ς}							1.4295	

The table presents point estimates and standard errors for model specifications B.1 through B.4. Point estimates were obtained by Generalized Indirect Inference, unless indicated otherwise. Parametric bootstrap standard errors are in parentheses. Bootstraps are based on 50 replications.

(i) Point estimate obtained using additional moment conditions. See discussion in Section 4.

Table 3
Evaluation of fit of models A.1 and B.1.
Descriptive statistics of PSID sample and data simulated from the estimated models.

Sample Statistic	Model	Overall	t=5	t=10	t=20	t=30	t=40
St. Dev. log Wage	**PSID**	**0.39**	**0.35**	**0.37**	**0.40**	**0.40**	**0.40**
	A1	0.39	0.38	0.39	0.40	0.38	0.37
	B1	0.40	0.38	0.39	0.40	0.41	0.42
St. Dev. log Earnings	**PSID**	**0.57**	**0.50**	**0.60**	**0.55**	**0.55**	**0.65**
	A1	0.58	0.57	0.59	0.59	0.57	0.56
	B1	0.59	0.53	0.56	0.59	0.61	0.62
St. Dev. log Hours	**PSID**	**0.28**	**0.28**	**0.30**	**0.27**	**0.27**	**0.34**
	A1	0.30	0.30	0.30	0.30	0.30	0.30
	B1	0.29	0.31	0.30	0.29	0.28	0.28
Mean Employment	**PSID**	**0.97**	**0.95**	**0.96**	**0.97**	**0.99**	**0.97**
	A1	0.96	0.95	0.95	0.96	0.96	0.97
	B1	0.96	0.94	0.95	0.96	0.97	0.96
Mean Job Change (if employed)	**PSID**	**0.08**	**0.20**	**0.14**	**0.07**	**0.05**	**0.03**
	A1	0.09	0.19	0.15	0.09	0.04	0.01
	B1	0.09	0.19	0.15	0.08	0.03	0.01
Emp. to Emp.	**PSID**	**0.98**	**0.99**	**0.97**	**0.98**	**0.99**	**0.98**
	A1	0.97	0.96	0.96	0.97	0.98	0.98
	B1	0.97	0.96	0.96	0.97	0.98	0.97
Unemp. to Emp.	**PSID**	**0.74**	**0.56**	**0.73**	**0.86**	**0.83**	**0.62**
	A1	0.69	0.68	0.70	0.72	0.66	0.76
	B1	0.70	0.65	0.69	0.72	0.70	0.65
Mean Tenure	**PSID**	**9.66**	**3.03**	**4.80**	**9.94**	**15.06**	**19.38**
	A1	9.79	2.44	4.50	9.94	16.05	21.71
	B1	9.85	2.43	4.46	9.93	16.36	21.72
Mean Emp. Duration	**PSID**	**11.96**	**4.23**	**6.64**	**12.62**	**17.42**	**21.40**
	A1	15.24	5.07	8.89	16.02	22.20	27.23
	B1	15.35	5.02	8.81	16.20	22.60	26.94
Mean Unemp. Duration	**PSID**	**1.47**	**1.60**	**1.54**	**1.40**	**1.17**	**1.36**
	A1	1.71	1.71	1.69	1.72	1.70	1.44
	B1	1.69	1.82	1.66	1.67	1.74	1.73

The table presents descriptive statistics of the PSID sample, and of data simulated from estimated models A.1 and B.1. Lead values are excluded. The descriptive statistics of simulated data are based on a simulated sample which is 10 times as large as the PSID sample, but has the same demographic structure (by potential experience) as the PSID sample. All statistics are computed using 3-year windows around the indicated value of t. For instance, t=10 corresponds to sample moments computed over all observations where t=9,10,11. The only exception is t=40, which uses only two years: t=39,40.

Table 4A (i)
Evaluation of fit of model A.1. Regressions comparing PSID sample and data simulated from estimated model - Employment and Job Change Regressions.

	PSID			SIMULATED DATA		
	1a [1]	2a [2]	3a [3]	1b [1]	2b [2]	3b [3]
Variable	E_t	E_t	JC_t	E_t	E_t	JC_t
PE_{t-1}	-0.0007	0.0151	-0.0070	-0.0004	0.0010	-0.0051
	(0.0005)	(0.0090)	(0.0009)	(0.0002)	(0.0023)	(0.0003)
PE^2_{t-1} /100	-0.0002	-0.0332	0.0168	-0.0004	-0.0005	0.0117
	(0.0011)	(0.0217)	(0.0020)	(0.0004)	(0.0055)	(0.0007)
ED_{t-1}	0.0024			0.0018		
	(0.0002)			(0.0000)		
UD_{t-1}		-0.1063			-0.0724	
		(0.0181)			(0.0028)	
TEN_{t-1}			-0.0081			-0.0092
			(0.0003)			(0.0001)
Constant	0.9642	0.7394	0.2169	0.9542	0.7963	0.2286
	(0.0047)	(0.0866)	(0.0086)	(0.0017)	(0.0219)	(0.0029)
Observations	27640	707	27046	271782	11688	264047
R-squared	0.01	0.05	0.05	0.01	0.06	0.08
RMSE	0.14	0.43	0.26	0.17	0.45	

The table presents least-squares regression results comparing PSID data and data simulated from estimated model A.1. Regressions on simulated data are based on a simulated sample which is 10 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 4A (ii)
Evaluation of fit of model A.1. Regressions comparing PSID sample and data simulated from estimated model - Wage Regressions.

	PSID				SIMULATED DATA			
	1a	2a	3a	4a	1b	2b	3b	4b
Variable	w_t	w_t	w_t	w_t	w_t	w_t	w_t	w_t
w_{t-1}	0.8850	0.6172	0.8828	0.6142	0.8999	0.7186	0.9009	0.7188
	(0.0029)	(0.0063)	(0.0029)	(0.0063)	(0.0008)	(0.0021)	(0.0008)	(0.0021)
w_{t-2}		0.3166		0.3173		0.2025		0.2034
		(0.0063)		(0.0063)		(0.0021)		(0.0021)
JC_t			-0.0281	-0.0374			0.0119	0.0139
			(0.0041)	(0.0044)			(0.0011)	(0.0013)
Constant	0.0125	0.0139	0.0148	0.0166	0.0077	0.0081	0.0066	0.0069
		(0.0063)		(0.0063)		(0.0021)		(0.0021)
Observations	27046	22578	27046	22578	264047	216819	264047	216819
R-squared	0.77	0.8	0.77	0.8	0.81	0.82	0.81	0.82
RMSE	0.18	0.17	0.18	0.17	0.17	0.16	0.17	0.16

The table presents least-squares regression results comparing PSID data and data simulated from estimated model A.1. Regressions on simulated data are based on a simulated sample which is 10 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 (iii)
Evaluation of fit of model A.1. Regressions comparing PSID sample and data simulated from estimated model - Hours Regressions.

	PSID		SIMULATED DATA	
	1a	2a	1b	2b
Variable	h_t	h_t	h_t	h_t
PE_{t-1}	0.0001	-0.0006	0.0001	-0.0006
	(0.0009)	(0.0009)	(0.0003)	(0.0003)
PE^2_{t-1}	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
h_{t-1}	0.4007	0.3684	0.3591	0.3471
	(0.0069)	(0.0067)	(0.0020)	(0.0019)
h_{t-2}	0.1899	0.1843	0.2773	0.2740
	(0.0069)	(0.0065)	(0.0020)	(0.0019)
w_t		-0.0006		-0.0172
		(0.0036)		(0.0013)
Constant	0.0161	0.0347	0.0037	0.0305
	(0.0100)	(0.0090)	(0.0034)	(0.0033)
Observations	23906	23313	239060	229631
R-squared	0.24	0.23	0.31	0.31
RMSE	0.23	0.21	0.25	0.24

The table presents least-squares regression results comparing PSID data and data simulated from estimated model A.1. Regressions on simulated data are based on a simulated sample which is 10 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 (iv)
Evaluation of fit of model A.1. Regressions comparing PSID sample and data simulated from estimated model - Earnings Regressions.

	PSID			SIMULATED DATA		
	1a	2a	3a	1b	2b	3b
Variable	e_t	e_t	e_t	e_t	e_t	e_t
PE_{t-1}	0.0014	0.0033		0.0002	0.0007	
	(0.0011)	(0.0013)		(0.0005)	(0.0005)	
PE^2_{t-1}	-0.0001	-0.0001		0.0000	0.0000	
	0.0000	0.0000		0.0000	0.0000	
e_{t-1}	0.8294	0.6861		0.6569	0.4385	
	(0.0037)	(0.0069)		(0.0014)	(0.0019)	
e_{t-2}		0.1871			0.3344	
		(0.0070)			(0.0019)	
w_t			0.9231			0.9889
			(0.0043)			(0.0013)
h_t			0.7705			0.8978
			(0.0068)			(0.0018)
Constant	-0.0002	-0.0238	0.0206	0.0035	-0.0034	0.0041
	(0.0117)	(0.0137)	(0.0017)	(0.0049)	(0.0056)	(0.0005)
Observations	28347	23906	32816	283470	239060	325242
R-squared	0.63	0.65	0.65	0.43	0.5	0.72
RMSE	0.33	0.32	0.3	0.44	0.41	0.29

The table presents least-squares regression results comparing PSID data and data simulated from estimated model A.1. Regressions on simulated data are based on a simulated sample which is 10 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 4B (i)

Evaluation of fit of model B.1. Regressions comparing PSID sample and data simulated from estimated model - Employment and Job Change Regressions.

	PSID			SIMULATED DATA		
	1a [1]	2a [2]	3a [3]	1b [1]	2b [2]	3b [3]
Variable	E_t	E_t	JC_t	E_t	E_t	JC_t
PE_{t-1}	-0.0007	0.0151	-0.0070	-0.0004	0.0062	-0.0064
	(0.0005)	(0.0090)	(0.0009)	(0.0002)	(0.0022)	(0.0003)
PE^2_{t-1} /100	-0.0002	-0.0332	0.0168	-0.0016	-0.0154	0.0141
	(0.0011)	(0.0217)	(0.0020)	(0.0004)	(0.0052)	(0.0007)
ED_{t-1}	0.0024			0.0023		
	(0.0002)			(0.0000)		
UD_{t-1}		-0.1063			-0.0750	
		(0.0181)			(0.0029)	
TEN_{t-1}			-0.0081			-0.0090
			(0.0003)			(0.0001)
Constant	0.9642	0.7394	0.2169	0.9502	0.7702	0.2419
	(0.0047)	(0.0866)	(0.0086)	(0.0017)	(0.0212)	(0.0029)
Observations	27640	707	27046	271808	11662	264067
R-squared	0.01	0.05	0.05	0.01	0.06	0.08
RMSE	0.14	0.43	0.26	0.17	0.45	0.28

The table presents least-squares regression results comparing PSID data and data simulated from estimated model B.1. Regressions on simulated data are based on a simulated sample which is 10 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 4B (ii)
Evaluation of fit of model B.1. Regressions comparing PSID sample and data simulated from estimated model - Wage Regressions.

	PSID				SIMULATED DATA			
	1a	2a	3a	4a	1b	2b	3b	4b
Variable	w_t	w_t	w_t	w_t	w_t	w_t	w_t	w_t
w_{t-1}	0.8850	0.6172	0.8828	0.6142	0.9011	0.7256	0.9023	0.7260
	(0.0029)	(0.0063)	(0.0029)	(0.0063)	(0.0008)	(0.0021)	(0.0008)	(0.0021)
w_{t-2}		0.3166		0.3173		0.1953		0.1959
		(0.0063)		(0.0063)		(0.0021)		(0.0021)
JC_t			-0.0281	-0.0374			0.0215	0.0211
			(0.0041)	(0.0044)			(0.0011)	(0.0013)
Constant	0.0125	0.0139	0.0148	0.0166	0.0047	0.0055	0.0027	0.0037
		(0.0063)		(0.0063)	(0.0003)	(0.0004)	(0.0003)	(0.0004)
Observations	27046	22578	27046	22578	264067	216859	264067	216859
R-squared	0.77	0.8	0.77	0.8	0.82	0.83	0.82	0.83
RMSE	0.18	0.17	0.18	0.17	0.17	0.16	0.17	0.16

The table presents least-squares regression results comparing PSID data and data simulated from estimated model B.1. Regressions on simulated data are based on a simulated sample which is 10 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 4B (iii)
Evaluation of fit of model B.1. Regressions comparing PSID sample and data simulated from estimated model - Hours Regressions.

	PSID		SIMULATED DATA	
	1a	2a	1b	2b
Variable	h_t	h_t	h_t	h_t
PE_{t-1}	0.0001	-0.0006	0.0002	-0.0008
	(0.0009)	(0.0009)	(0.0003)	(0.0003)
PE^2_{t-1}	0.0000	0.0000	-0.0001	0.0014
	0.0000	0.0000	(0.0007)	(0.0007)
h_{t-1}	0.4007	0.3684	0.3071	0.2778
	(0.0069)	(0.0067)	(0.0019)	(0.0019)
h_{t-2}	0.1899	0.1843	0.2914	0.2735
	(0.0069)	(0.0065)	(0.0019)	(0.0019)
w_t		-0.0006		-0.0112
		(0.0036)		(0.0012)
Constant	0.0161	0.0347	0.0027	0.0312
	(0.0100)	(0.0090)	(0.0034)	(0.0033)
Observations	23906	23313	239060	229785
R-squared	0.24	0.23	0.26	0.24
RMSE	0.23	0.21	0.25	0.23

The table presents least-squares regression results comparing PSID data and data simulated from estimated model B.1. Regressions on simulated data are based on a simulated sample which is 10 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 4B (iv)
Evaluation of fit of model B.1. Regressions comparing PSID sample and data simulated from estimated model - Earnings Regressions.

	PSID			SIMULATED DATA		
	1a	2a	3a	1b	2b	3b
Variable	e_t	e_t	e_t	e_t	e_t	e_t
PE_{t-1}	0.0014	0.0033		0.0015	0.0013	
	(0.0011)	(0.0013)		(0.0005)	(0.0005)	
PE^2_{t-1}	-0.0001	-0.0001		-0.0031	-0.0027	
	0.0000	0.0000		(0.0011)	(0.0011)	
e_{t-1}	0.8294	0.6861		0.6650	0.4313	
	(0.0037)	(0.0069)		(0.0014)	(0.0019)	
e_{t-2}		0.1871			0.3586	
		(0.0070)			(0.0019)	
w_t			0.9231			0.9973
			(0.0043)			(0.0013)
h_t			0.7705			0.8799
			(0.0068)			(0.0019)
Constant	-0.0002	-0.0238	0.0206	-0.0177	-0.0143	0.0048
	(0.0117)	(0.0137)	(0.0017)	(0.0049)	(0.0056)	(0.0005)
Observations	28347	23906	32816	283470	239060	325265
R-squared	0.63	0.65	0.65	0.44	0.51	0.72
RMSE	0.33	0.32	0.3	0.44	0.41	0.29

The table presents least-squares regression results comparing PSID data and data simulated from estimated model B.1. Regressions on simulated data are based on a simulated sample which is 10 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 5A (i)

Decomposition of Cross-Sectional Variance in Lifetime Earnings, Wage, and Hours (as percent of total variation), according to estimated Model A.1. Shocks turned off one at a time (for all t).

	Column									
	I	II	III	IV	V	VI	VII	VIII	IX	X
	Shock						Breakdown of Composite 'Shock'			
Variable	ε^e	ε^h	ε^w	Composite	η	μ	ξ	υ	E	JC
Lifetime Earnings	5.0	2.7	16.1	42.3	-1.4	35.4	39.9	45.4	14.8	0.0
Lifetime Wage	0.0	0.0	24.5	48.3	0.5	26.6	0.4	69.9	29.7	0.0
Lifetime Hours	0.0	4.2	0.6	68.3	0.8	26.2	90.9	1.1	8.1	0.0

Entries in columns I to VI represent the contribution of a given type of shock to lifetime variance in earnings, wage, and hours, and are expressed as a percentage of the variance in the basecase. The basecase corresponds to the simulation

Table 5A (ii)
Decomposition of Cross-Sectional Variance in Earnings, Wage, and Hours (all in levels) at different t (as percent of total variation), according to estimated model A.1. Shocks turned off one at a time (for all t).

	Column									
	I	II	III	IV	V	VI	VII	VIII	IX	X
	Shock						Breakdown of Composite 'Shock'			
Variable/Horizon	ε^e	ε^h	ε^w	Composite	η	μ	ξ	υ	E	JC
Earnings										
t = 1	19.0	10.0	29.5	20.6	0.6	20.3	70.9	19.9	9.2	0.0
t = 5	19.1	9.9	23.6	30.8	-0.6	17.3	39.4	52.2	8.4	0.0
t = 10	19.1	9.5	20.6	33.9	-1.3	18.2	38.5	52.0	9.4	0.0
t = 20	19.4	9.7	19.1	36.0	-1.9	17.7	37.8	54.5	7.7	0.0
t = 30	19.4	11.0	18.3	35.2	-1.6	17.7	41.4	53.0	5.6	0.0
t = 40	20.3	10.8	18.4	33.5	-2.1	19.1	44.7	51.8	3.5	0.0
Wage										
t = 1	0.0	0.0	71.0	9.8	0.9	18.3	7.2	83.4	9.4	0.0
t = 5	0.0	0.0	53.4	35.7	-1.7	12.6	1.0	95.8	3.2	0.0
t = 10	0.0	0.0	45.6	45.9	-2.7	11.2	-0.8	94.8	6.0	0.0
t = 20	0.0	0.0	41.5	49.6	-3.5	12.4	3.2	89.6	7.2	0.0
t = 30	0.0	0.0	42.0	48.1	-3.6	13.5	3.4	90.2	6.4	0.0
t = 40	0.0	0.0	43.2	45.3	-3.4	15.0	3.2	92.0	4.8	0.0
Hours										
t = 1	0.0	29.8	0.6	52.4	5.3	11.8	78.3	1.5	20.2	0.0
t = 5	0.0	31.7	0.6	52.6	3.2	12.0	83.0	2.0	15.0	0.0
t = 10	0.0	32.0	0.6	52.9	2.6	11.8	86.2	0.3	13.5	0.0
t = 20	0.0	33.8	0.6	52.8	2.0	10.8	89.8	-0.6	10.8	0.0
t = 30	0.0	34.8	0.6	52.9	1.5	10.3	92.2	-0.1	8.0	0.0
t = 40	0.0	35.6	0.6	54.0	0.8	8.9	92.3	1.8	5.9	0.0

Entries in columns I to VI represent 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. The basecase corresponds to the simulation of 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 VI 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 VII, VIII, IX and X we decompose Column IV. Column VII is the marginal contribution of ξ, VII is the marginal contribution of υ with var(ξ) set to 0, IX is the marginal contribution of eliminating unemployment spells with Var(ξ) and Var(υ) set to 0, and X is
the marginal contribution of eliminating job changes with Var(ξ) and Var(υ) set to 0 and no unemployment.

Table 5B (i)

Decomposition of Cross-Sectional Variance in Lifetime Earnings, Wage, and Hours (as percent of total variation), according to estimated Model B.1. Shocks turned off one at a time (for all t).

Column:	I	II	III	IV	V	IV	VII	VIII
Shock:	ε^e	ε^h	ε^w	ε^w_1	JC	E	η	μ
Variable								
Lifetime Earnings	4.9	3.1	24.0	6.1	0.7	5.1	2.4	53.7
Lifetime Wage	0.0	0.0	43.7	11.5	1.4	13.4	2.7	27.4
Lifetime Hours	0.0	5.1	2.5	0.8	2.1	4.7	8.1	76.6

Entries represent the contribution of a given shock to lifetime variance in earnings, wage, and hours, and are expressed as a percentage of the variance in the basecase. The basecase corresponds to the simula of the full estimated model. To compute the contribution of a particular shock, we simulate the model agai setting the variance of the 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 i nonlinear, the contributions don't sum up to 100%. We have normalized them to sum to 100.

Table 5B (ii)
Decomposition of Cross-Sectional Variance in Earnings, Wage, and Hours (all in levels) at different levels of potential experience, according to estimated Model B.1. Shocks turned off one at a time (for all t).

Column:	I	II	III	IV	V	VI	VII	VIII
Shock:	ε^e	ε^h	ε^w	ε^w_1	JC	E	η	μ
Variable/Horizon								
Earnings								
t = 1	27.6	17.9	0.0	45.3	0.0	0.8	1.4	7.0
t = 5	20.5	13.2	23.4	17.0	10.7	0.4	1.1	13.6
t = 10	18.7	12.3	27.6	7.7	11.2	1.1	1.5	19.9
t = 20	17.3	11.9	27.8	1.8	6.7	1.5	1.8	31.3
t = 30	17.5	11.6	25.5	0.8	2.3	1.2	1.4	39.5
t = 40	17.5	11.5	23.8	0.1	0.3	0.8	0.4	45.7
Wage								
t = 1	0.0	0.0	0.0	95.8	0.0	-2.3	-0.1	6.5
t = 5	0.0	0.0	52.0	36.7	20.9	-5.4	-2.0	-2.2
t = 10	0.0	0.0	62.3	16.6	23.9	-2.9	-1.2	1.4
t = 20	0.0	0.0	65.0	5.0	16.2	0.3	0.7	12.8
t = 30	0.0	0.0	62.3	1.7	6.0	2.9	0.9	26.3
t = 40	0.0	0.0	58.0	0.6	0.1	3.7	-0.2	37.7
Hours								
t = 1	0.0	33.3	0.0	2.9	0.0	10.2	6.4	47.3
t = 5	0.0	38.0	1.7	1.6	-0.1	8.3	5.0	45.5
t = 10	0.0	41.2	2.6	0.8	0.2	6.9	5.7	42.5
t = 20	0.0	45.7	3.2	0.3	1.2	5.2	5.8	38.6
t = 30	0.0	48.1	3.4	0.1	1.4	5.0	5.9	36.0
t = 40	0.0	49.2	3.2	0.1	1.1	5.5	6.0	35.0

Entries represent the contribution of a given shock to the variance in earnings, wage, and hours, for a cross-section of simulated individuals with a particular level of potential experience, t. The contribution is expressed as a percentage of the variance in the basecase. The basecase corresponds to the simulation of 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 a given 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 them to sum to 100.

Table 5B (v)
Decomposition of Cross-Sectional Variance in Lifetime Earnings, Wage, and Hours (as percent of total variation), according to estimated Model B.1. Shocks turned off sequentially (for all t) from left to right.

Column:	I	II	III	IV	V	IV	VII	VIII
Shock:	ε^e	ε^h	ε^w	ε^w_1	JC	E	η	μ
Variable								
Lifetime Earnings	6.3	3.7	28.1	6.3	-3.1	9.1	2.3	47.3
Lifetime Wage	0.0	0.0	53.7	12.0	-6.8	26.4	0.0	14.8
Lifetime Hours	0.0	5.7	2.8	0.9	2.1	4.2	10.4	73.9

Entries represent the sequential marginal contribution of a given shock to lifetime variance in earnings, wage, and hou and are expressed as a percentage of the variance in the basecase. The basecase corresponds to the simulation of the full estimated model. To compute the (sequential) contribution of a particular shock, we simulate the model again, setting (each time) the variance of one additional shock to zero for all t (the shocks are shut off from left to right). We then compute the variance of the appropriate variables. The difference relative to the basecase is the contribution of the given shock. The sensitivity of the results to the order of the shocks will be discussed in a future dra

Table A1
Monte Carlo Experiments - Model A.1

Column:		I	II	III			I	II	III
Eq. / Variable	Param	"True"	Mean	St. Dev.	**Eq.** / Variable	Param	"True"	Mean	St. Dev.
E-E					**Wage**				
(cons)	γ^{EE}_{0}	2.7738	2.6095	(0.1581)	(H_t)	γ^{w}_{H}			
(PE_{t-1})	γ^{EE}_{PE}	0.0636	0.0779	(0.0162)	(μ)	δ^{w}_{μ}	0.1428	0.1450	(0.0199)
(PE^2_{t-1})/100	γ^{EE}_{PEsq}	-0.0011	-0.0133	(0.0348)	(JC_t)	γ^{υ}_{0}	0.0388	0.0414	(0.0060)
(ED_{t-1})	γ^{EE}_{ED}	-0.0910	-0.1087	(0.0111)	(υ_{t-1})	ρ_{υ}	0.6522	0.6701	(0.0377)
(H_t)	γ^{EE}_{H}				(ε^{υ})	σ_{υ}	0.2372	0.2480	(0.0079)
(υ_{t-1})	δ^{EE}_{υ}				($\varepsilon^{\upsilon}_{1}$)	$\sigma_{\upsilon 1}$	0.1005	0.1006	(0.0005)
(μ)	δ^{EE}_{μ}	0.8036	0.8629	(0.0789)	(ω_{t-1})	ρ_{ω}	0.9051	0.8984	(0.0160)
(η)	δ^{EE}_{η}	-1.1365	-1.1575	(0.1040)	($1-E_t$)	γ^{ω}_{1-E}	-0.2894	-0.2849	(0.0198)
U-E					(ε^{ω})	σ_{ω}	0.1014	0.0988	(0.0026)
(cons)	γ^{UE}_{0}	0.9273	0.6703	(0.2688)	(ε^{ω}_{1})	$\sigma_{\omega 1}$	0.3053	0.3197	(0.0114)
(PE_{t-1})	γ^{UE}_{PE}	0.0103	0.0255	(0.0271)	**Hours**				
(PE^2_{t-1})/100	γ^{UE}_{PEsq}	-0.0180	-0.0368	(0.0762)	(cons)	γ^{h}_{0}	-0.3945	-0.4026	(0.0078)
(UD_{t-1})	γ^{UE}_{UD}	0.0733	0.0443	(0.0544)	(E_t)	γ^{h}_{E}	0.4265	0.4254	(0.0073)
(H_t)	γ^{UE}_{H}					σ_{ξ}	0.1959	0.2148	(0.0052)
(μ)	δ^{UE}_{μ}	0.8004	0.7264	(0.1373)	(w_t)	γ^{h}_{w}	-0.0852	-0.0837	(0.0113)
(η)	δ^{UE}_{η}	0.3691	0.2809	(0.1207)	(H_t)	γ^{h}_{H}			
JC					(μ)	δ^{h}_{μ}	0.0947	0.0975	(0.0091)
(cons)	γ^{JC}_{0}	-0.6098	-0.5225	(0.0707)	(η)	δ^{h}_{η}	-0.0187	-0.0125	(0.0079)
(PE_{t-1})	γ^{JC}_{PE}	0.0029	-0.0063	(0.0096)	(ε^{h}_{t})	σ_{h}	0.1632	0.1602	(0.0015)
(PE^2_{t-1})/100	γ^{JC}_{PEsq}	-0.0798	-0.0436	(0.0263)	**Earnings**				
(TEN_{t-1})	γ^{JC}_{TEN}	-0.1230	-0.1207	(0.0079)	(cons)	γ^{e}_{0}	-0.0002	-0.0006	(0.0019)
(υ_{t-1})	δ^{JC}_{υ}				(w_t)	γ^{e}_{w}	1.0288	1.0214	(0.0040)
(ξ_{t-1})	δ^{JC}_{ξ}				(h_t)	γ^{e}_{h}	1.0122	1.0134	(0.0071)
(μ)	δ^{JC}_{μ}	-0.3051	-0.3001	(0.0411)	(ε^{e}_{t})	σ_{e}	0.2334	0.2334	(0.0013)
(η)	δ^{JC}_{η}	0.1558	0.1283	(0.0270)					

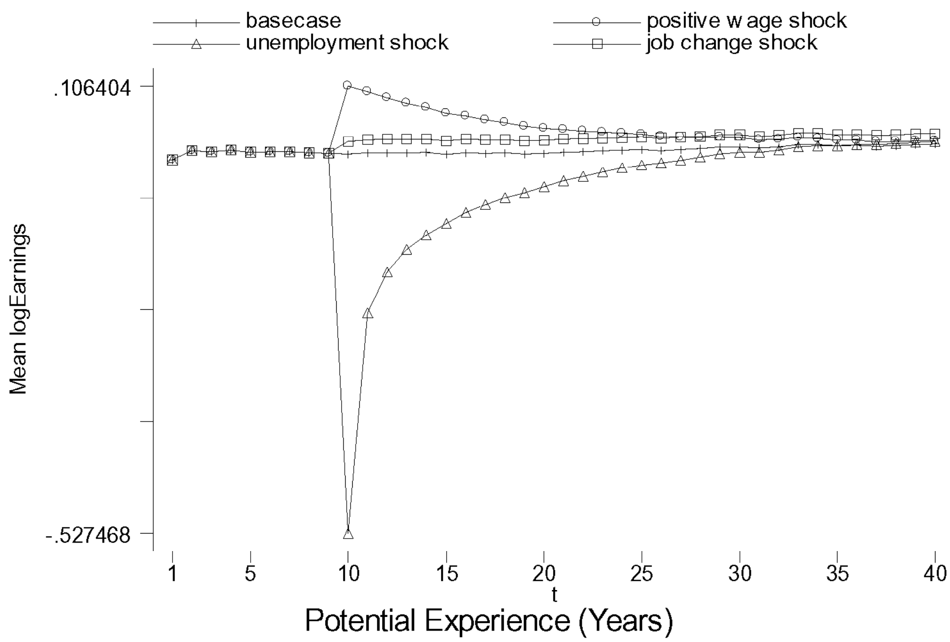
The table presents results from Monte Carlo experiments. Column I presents a vector of hypothesized structural parameters, which we call the "true" parameter values. We use this vector of hypothesized parameters to randomly generate a sample of "true" data from our structural model. We then use these artificial data and our estimation methodology to obtain a vector of parameter estimates. We repeat this procedure 50 times. Column II presents the mean vector of parameter estimates obtained across the 50 replications. Column III presents the empirical standard deviation of the estimates across the 50 replications.

Table A2
Monte Carlo Experiments - Model B.1

Column:		I	II	III			I	II	III
Eq. / Variable	Param	"True"	Mean	St. Dev.	**Eq.** / Variable	Param	"True"	Mean	St. Dev.
E-E					**Wage**				
(cons)	γ^{EE}_{0}	2.5000	2.3108	(0.1487)	(H_t)	γ^{w}_{H}			
(PE_{t-1})	γ^{EE}_{PE}	0.0667	0.0768	(0.0117)	(ω_{t-1})	ρ_ω	0.9480	0.9490	(0.0024)
(PE^2_{t-1})	γ^{EE}_{PEsq}	-0.0005	-0.0007	(0.0002)		φ_1	-0.1943	-0.1905	(0.1013)
(ED_{t-1})	γ^{EE}_{ED}	-0.0679	-0.0766	(0.0122)	(JC_t)	γ^{ω}_{JC}	0.0222	0.0255	(0.0050)
(H_t)	γ^{EE}_{H}				(1-E_t)	γ^{ω}_{1-E}	-0.2732	-0.2682	(0.0156)
(w_{t-1})	γ^{EE}_{w}				(μ)	δ^{ω}_{μ}	0.0148	0.0147	(0.0014)
(μ)	δ^{EE}_{μ}	0.8681	0.8526	(0.0873)	(ε^{ω})	σ_ω	0.0965	0.0941	(0.0018)
(η)	δ^{EE}_{η}	-0.9234	-0.9074	(0.1085)	(ε^{ω}_1)	$\sigma_{\omega 1}$ (i)			
U-E						φ_2 (i)			
(cons)	γ^{UE}_{0}	1.0755	0.7854	(0.2756)	(μ)	θ^{ω}_{μ}	-0.0939	-0.0912	(0.0264)
(PE_{t-1})	γ^{UE}_{PE}	0.0324	0.0515	(0.0322)					
(PE^2_{t-1})	γ^{UE}_{PEsq}	-0.0008	-0.0011	(0.0008)	**Hours**				
(UD_{t-1})	γ^{UE}_{UD}	0.0932	0.0885	(0.0713)	(cons)	γ^{h}_{0}	-0.4047	-0.4097	(0.0057)
(H_t)	γ^{UE}_{H}				(E_t)	γ^{h}_{E}	0.4325	0.4314	(0.0054)
(μ)	δ^{UE}_{μ}	0.8844	0.8407	(0.1421)	(w_t)	γ^{h}_{w}	-0.1475	-0.1446	(0.0091)
(η)	δ^{UE}_{η}	0.2576	0.1912	(0.0840)	(H_t)	γ^{h}_{H}			
JC					(μ)	δ^{h}_{μ}	0.1795	0.1789	(0.0051)
(cons)	γ^{JC}_{0}	-0.6275	-0.5561	(0.0692)	(η)	δ^{h}_{η}	0.0584	0.0610	(0.0092)
(PE_{t-1})	γ^{JC}_{PE}	0.0047	0.0008	(0.0112)	(ε^{h})	σ_h	0.1805	0.1808	(0.0009)
(PE^2_{t-1})	γ^{JC}_{PEsq}	-0.0009	-0.0007	(0.0003)					
(TEN_{t-1})	γ^{JC}_{TEN}	-0.1162	-0.1171	(0.0074)	**Earnings**				
(w_{t-1})	δ^{JC}_{w}				(cons)	γ^{e}_{0}	0.0001	-0.0008	(0.0021)
(μ)	δ^{JC}_{μ}	-0.2766	-0.2499	(0.0303)	(w_t)	γ^{e}_{w}	1.0353	1.0288	(0.0041)
(η)	δ^{JC}_{η}	0.2508	0.2170	(0.0277)	(h_t)	γ^{e}_{h}	1.0121	1.0141	(0.0063)
					(ε^{e})	σ_e	0.2316	0.2309	(0.0010)

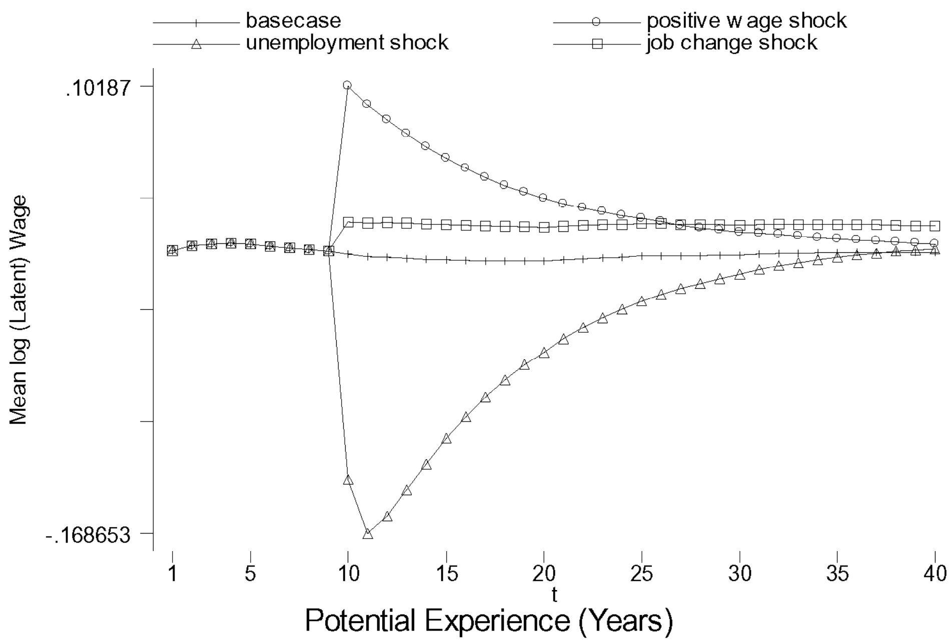
The table presents results from Monte Carlo experiments. Column I presents a vector of hypothesized structural parameters, which we call the "true" parameter values. We use this vector of hypothesized parameters to randomly generate a sample of "true" data from our structural model. We then use these artificial data and our estimation methodology to obtain a vector of parameter estimates. We repeat this procedure **20** times. Column II presents the mean vector of parameter estimates obtained across the **20** replications. Column III presents the empirical standard deviation of the estimates across the **20** replications.

Figure 1a
Mean Response of log Earnings to Various Shocks at t=10: Unemployment shock, Job Change, Wage Shock (positive, one-st.dev. shock). Model A.1.



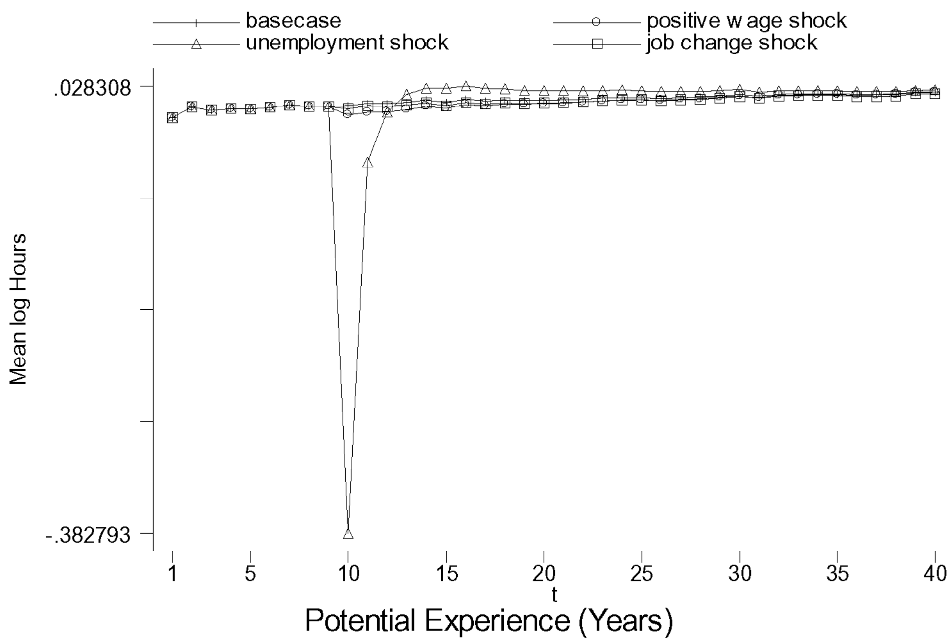
All shocks are constructed as follows. From t=1 until t=9 we simulate the careers of 50,000 individuals according to Model A.1. At t=10, we impose (exogenously) a particular 'shock' to all simulated individuals. The various shocks are defined as follows. Unemployment shock: we make all individuals unemployed at t=10. Job change shock: we make all individuals who are not unemployed at t=9 or t=10 change employers at t=10. Wage shock: we set the innovation to the autoregressive wage component equal to our estimate of σ_ω (>0) for all individuals. After imposing one of these shocks, we continue the simulation of the rest of the individuals' careers from t=11 to t=40 according to model A.1. Figure 1a displays the evolution of the cross-sectional mean of log earnings across all 50,000 individuals.

Figure 1b
Mean Response of log (Latent) Wage to Various Shocks at t=10: Unemployment shock, Job Change, Wage Shock (positive, one-st.dev. shock). Model A.1.



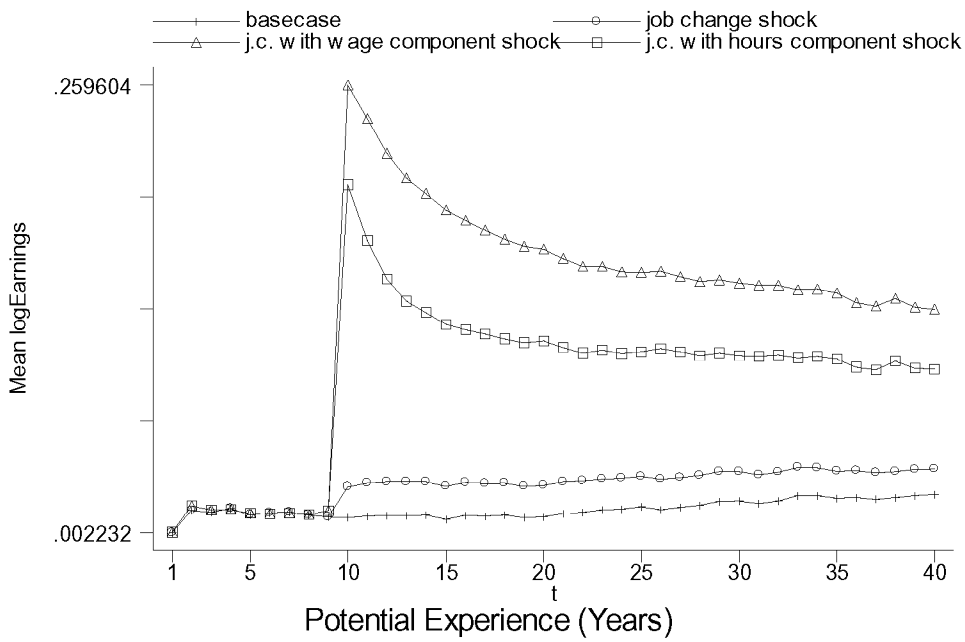
All shocks are constructed as follows. From t=1 until t=9 we simulate the careers of 50,000 individuals according to Model A.1. At t=10, we impose (exogenously) a particular 'shock' to all simulated individuals. The various shocks are defined as follows. Unemployment shock: we make all individuals unemployed at t=10. Job change shock: we make all individuals who are not unemployed at t=9 or t=10 change employers at t=10. Wage shock: we set the innovation to the autoregressive wage component equal to our estimate of σ_ω (>0) for all individuals. After imposing one of these shocks, we continue the simulation of the rest of the individuals' careers from t=11 to t=40 according to model A.1. Figure 1b displays the evolution of the cross-sectional mean of the log latent wage across all 50,000 individuals.

Figure 1c
Mean Response of log Hours to Various Shocks at t=10: Unemployment shock, Job Change, Wage Shock (positive, one-st.dev. shock). Model A.1.



All shocks are constructed as follows. From t=1 until t=9 we simulate the careers of 50,000 individuals according to Model A.1. At t=10, we impose (exogenously) a particular 'shock' to all simulated individuals. The various shocks are defined as follows. Unemployment shock: we make all individuals unemployed at t=10. Job change shock: we make all individuals who are not unemployed at t=9 or t=10 change employers at t=10. Wage shock: we set the innovation to the autoregressive wage component equal to our estimate of σ_ω (>0) for all individuals. After imposing one of these shocks, we continue the simulation of the rest of the individuals' careers from t=11 to t=40 according to model A.1. Figure 1c displays the evolution of the cross-sectional mean of log hours across all 50,000 individuals.

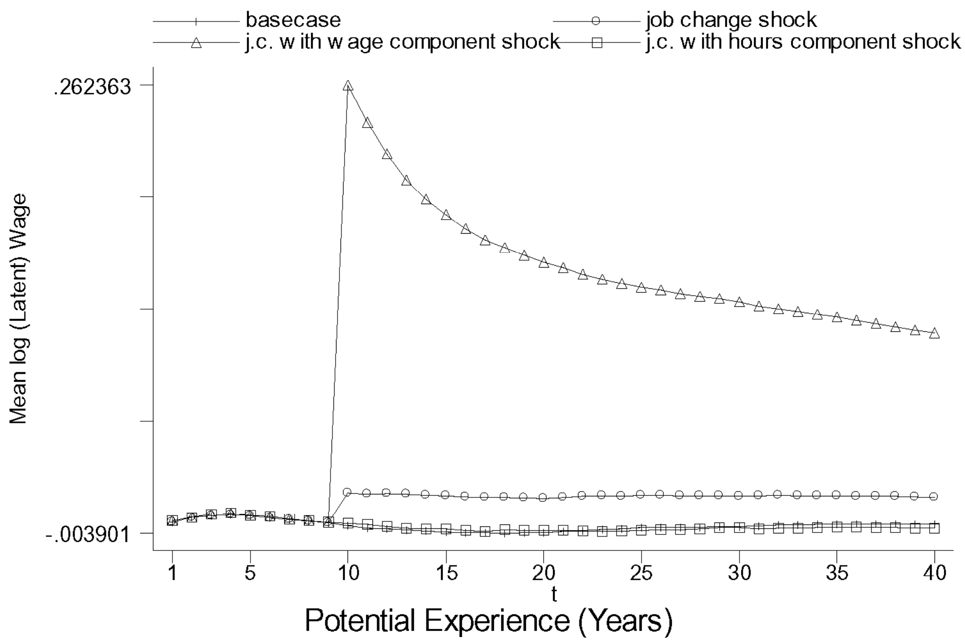
Figure 2a
Mean Response of log Earnings to Various Shocks at t=10: job change shock, job-change shock with one-standard-deviation shock to υ, job-change shock with one-standard-deviation shock to ξ. Model A.1.



All shocks are constructed as follows. From t=1 until t=9 we simulate the careers of 50,000 individuals according to Model A.1. At t=10, we impose (exogenously) a particular 'shock' to all simulated individuals. The various shocks are defined as follows. Job change shock: we make all individuals who are not unemployed at t=9 or t=10 change employers at t=10. Job-specific wage component (υ) shock: The job change shock is now accompanied by a one-standard-deviation increase (0.2372) in the error term of the process for the job-specific wage component υ. Job-specific hours component (ξ) shock: The job change shock is now accompanied by a one-standard-deviation increase (0.1959) in the error term of the process for the job-specific hours component ξ. After imposing one of these shocks, we continue the simulation of the rest of the individuals' careers from t=11 to t=40 according to model A.1. Figure 2a displays the evolution of the cross-sectional mean of log earnings across all 50,000 individuals.

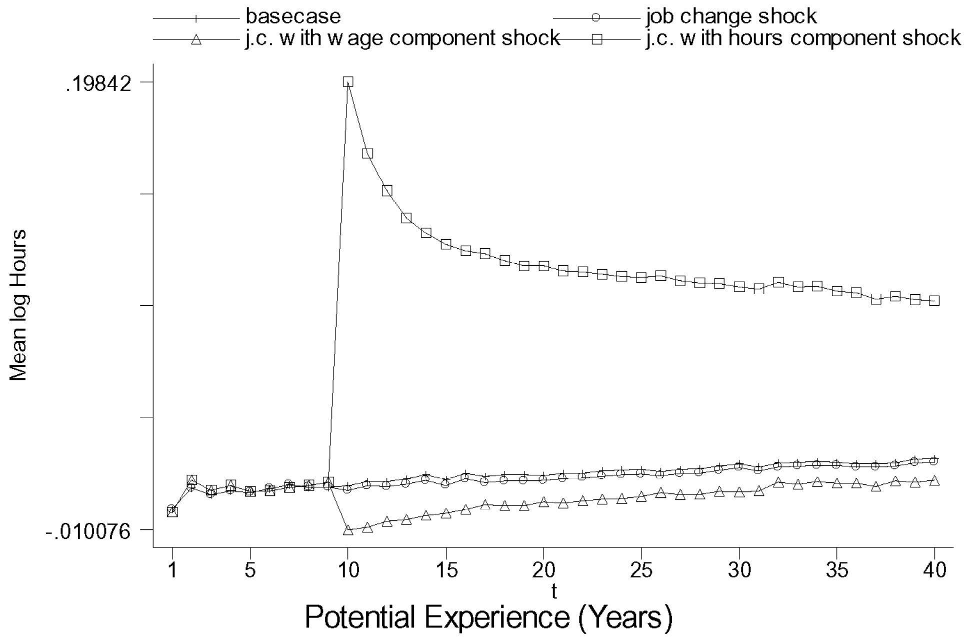
Figure 2b

Mean Response of log (Latent) Wage to Various Shocks at t=10: job change shock, job-change shock with one-standard-deviation shock to υ, job-change shock with one-standard-deviation shock to ξ. Model A.1.



All shocks are constructed as follows. From t=1 until t=9 we simulate the careers of 50,000 individuals according to Model A.1. At t=10, we impose (exogenously) a particular 'shock' to all simulated individuals. The various shocks are defined as follows. Job change shock: we make all individuals who are not unemployed at t=9 or t=10 change employers at t=10. Job-specific wage component (υ) shock: The job change shock is now accompanied by a one-standard-deviation increase (0.2372) in the error term of the process for the job-specific wage component υ. Job-specific hours component (ξ) shock: The job change shock is now accompanied by a one-standard-deviation increase (0.1959) in the error term of the process for the job-specific hours component ξ. After imposing one of these shocks, we continue the simulation of the rest of the individuals' careers from t=11 to t=40 according to model A.1. Figure 2b displays the evolution of the cross-sectional mean of log (latent) wage across all 50,000 individuals.

Figure 2c
Mean Response of log Hours to Various Shocks at t=10: job change shock, job-change shock with one-standard-deviation shock to υ, job-change shock with one-standard-deviation shock to ξ. Model A.1.



All shocks are constructed as follows. From t=1 until t=9 we simulate the careers of 50,000 individuals according to Model A.1. At t=10, we impose (exogenously) a particular 'shock' to all simulated individuals. The various shocks are defined as follows. Job change shock: we make all individuals who are not unemployed at t=9 or t=10 change employers at t=10. Job-specific wage component (υ) shock: The job change shock is now accompanied by a one-standard-deviation increase (0.2372) in the error term of the process for the job-specific wage component υ. Job-specific hours component (ξ) shock: The job change shock is now accompanied by a one-standard-deviation increase (0.1959) in the error term of the process for the job-specific hours component ξ. After imposing one of these shocks, we continue the simulation of the rest of the individuals' careers from t=11 to t=40 according to model A.1. Figure 2c displays the evolution of the cross-sectional mean of log hours across all 50,000 individuals.