Econometrica, Vol. 81, No. 4 (July, 2013), 1395–1454

MODELING EARNINGS DYNAMICS

BY JOSEPH G. ALTONJI, ANTHONY A. SMITH, JR., AND IVAN VIDANGOS[1]

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

KEYWORDS: Wage growth, job mobility, unemployment, inequality, indirect inference.

1. INTRODUCTION

IN THIS PAPER, WE BUILD AND ESTIMATE A MODEL OF EARNINGS. We have three main goals. The first is to advance the literature in labor economics on how employment, hours, wages, and earnings are determined over a career. We examine the effects of education, race, experience, employment duration, job tenure and unobserved heterogeneity, employment shocks, shocks to general skills, and draws of new job opportunities offering different hours and wages. We trace out the response of wages, hours, and earnings to the various shocks and determine the channels through which they operate. Our analysis addresses a number of long-standing questions in labor economics. For example, we provide estimates of the relative importance of general skill accumula-

[1]We are grateful to Richard Blundell, Mary Daly, Rasmus Lentz, Costas Meghir, Paul Oyer, Luigi Pistaferri, and three anonymous referees for helpful discussions and suggestions, and to Lena Yemelyanov for excellent research assistance. We also thank participants in seminars at the Bank of Spain, UC Berkeley, University of British Columbia, CEMFI, University of Chicago, the Federal Reserve Bank of San Francisco, the Federal Reserve Board, Georgetown University, Harvard University, Pennsylvania State University, Princeton University, University of Rochester, Stanford University, Vanderbilt University, and Yale University, and conference sessions at the Society of Economic Dynamics (June 2005), the World Congress of the Econometric Society (August 2005), the Cowles Foundation Macro/Labor Economics Conference (May 2006), NBER (Nov. 2006), the Econometric Society Winter Meetings (January 2007), and the Society for Computational Economics (June 2008) for valuable comments. Our research has been supported by the Cowles Foundation and the Economic Growth Center, Yale University, and by NSF Grant SES-0112533 (Altonji). The views expressed in the paper are our own and not necessarily those of the Federal Reserve Board, Yale University, NBER, or other members of their staffs. We are responsible for the remaining shortcomings of the paper.

 DOI: 10.3982/ECTA8415

tion, job shopping, and job tenure for career wage growth, and we quantify the specific channels through which an exogenous employment shock affects the path of wage rates, hours, and earnings. We study the effects of shocks on the future variance of earnings changes as well as on the average path.

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

Our third goal is to provide a richer model of earnings for use in studies of consumption and saving as well as in dynamic stochastic general-equilibrium models that are a cornerstone of modern macroeconomics and public finance. Such models have been used to study the distribution of wealth, the costs of business cycles, asset pricing, and other important questions.[2] 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 of earnings innovations.[3]

Almost all of the existing structural studies base their modeling and calibration choices for the earnings process on the large empirical literature on univariate statistical models.[4] Much has been learned from this work. With only one indicator, however, even richly specified univariate models cannot identify the various sources of earnings fluctuations, their relative importance, their dynamic behavior, or the economics underlying how labor market outcomes are determined. Without such information, it is difficult to think about the potential welfare consequences of specific sources of variation or of policies such as unemployment insurance, employment regulations, wage subsidies, or earned income tax credits that insure against particular types of shocks to income. Furthermore, the innovations in the univariate representation of a multivariate time series process may be aggregates of current and past shocks in the multivariate representation. This will lead to mistakes in characterizing what the surprises to the agent are even under the assumption that the agent's information set is the same as the econometrician's.

Only a few 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 av-

[2]Examples include Huggett (1996), Krusell and Smith (1998), and Castañeda, Díaz-Giménez, and Ríos-Rull (2003) on consumption and wealth, Krusell and Smith (1999) and Storesletten, Telmer, and Yaron (2001) on the costs of business cycles, and Krusell and Smith (1997) and Storesletten, Telmer, and Yaron (2007) on asset pricing.

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

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

erage model of the first difference in consumption, family income, earnings, hours, wages, and unemployment. The models that we consider, in contrast to those mentioned above, incorporate discrete events such as job changes, employment loss, interactions between job changes and wages, and effects of these discrete events on the variance of wage and hours shocks.[5]

There are two distinct paths that one might take in formulating a multivariate model of earnings. The first approach is the development of a statistical model of the process with little attention to an underlying theory of household decisions and constraints. This approach is in the spirit of the literature on univariate earnings processes, but the absence of theory limits what one can learn about how earnings are determined. The second approach is to develop a model that is based on lifetime utility maximization. Grounding the model of the income process in a utility maximization framework provides a foundation for using the results to analyze policies when earnings are partially endogenous. The main disadvantage is the difficulty of specifying and estimating a model that incorporates labor supply choices, job search decisions, hours constraints, voluntary separations, and involuntary job changes. Indeed, we do not know of any papers that have studied work hours and employment using a lifecycle utility maximization model that incorporates job-specific hours constraints, let alone job mobility decisions.[6] Estimation of a structural model that is as rich as the one that we work with would require solving an intertemporal model of job search, labor supply (in the presence of hours constraints), and savings as part of the estimation strategy, and is probably out of reach at the present time from a computational point of view. Low, Meghir, and Pistaferri (2010) took a major step in this direction by studying earnings risk and social insurance in the context of an intertemporal model of consumption, employment participation, wages, and mobility. They worked with a simpler model of the earnings process than we do, but were able to measure welfare costs of the risk associated with innovations in the persistent wage component, an employer-specific wage component, and job loss and unemployment. Our study is complementary to theirs.

Although our model falls short of a fully specified behavioral model, the equations can be viewed as approximations to the decision rules relating choices to state variables that would arise in a structural model based on lifetime utility maximization. The parameters of the rules depend on an underlying set of "deep" parameters that characterize labor supply preferences, job search technology, etc. The class of models that we consider is rich enough to address a number of core questions in labor economics, but tractable enough

[5] A number of recent studies provide structural models of wage rates, job mobility, and employment dynamics, including Barlevy (2008), Buchinsky, Fougère, Kramarz, and Tchernis (2010), and Bagger, Fontaine, Postel-Vinay, and Robin (2011), who provided references to a few additional studies. Wolpin (1992) is an early effort. We discuss the evidence below.

[6] Ham and Reilly (2002) tested for hours restrictions in an intertemporal labor supply framework. Blundell and MaCurdy (1999) surveyed the labor supply literature.

to be used in place of univariate income models that dominate the literature on savings, portfolio choice, etc. Furthermore, it provides a natural path to future analyses that include other important economic risks that individuals face, including changes in family structure through marriage, divorce, and the death of a spouse.

We estimate the model using data on male household heads from the Panel Study of Income Dynamics (PSID). Given the presence of interactions among discrete and continuous variables, unobserved heterogeneity and state dependence in multiple equations, measurement error, and a highly unbalanced sample, conventional maximum likelihood and method of moments approaches are not feasible. For this reason, we use indirect inference (henceforth I-I), which is one of a family of simulation-based approaches to estimation that involve comparing the distribution of artificial data generated from the structural model at a given set of parameter values to features of the actual data.[7] We use the smoothing procedure suggested by Keane and Smith (2003), which allows us to use gradient-based numerical optimization methods in the presence of both discrete and continuous endogenous variables. Estimation of our model is not straightforward, and a secondary contribution of our research is to explore the feasibility and performance of I-I in large models with a mix of discrete and continuous variables.[8]

Our main results are as follows. First, education, race, and the two forms of unobserved permanent heterogeneity play an important role in employment transitions and job changes. Second, consistent with most of the large literature on the labor supply of male household heads, wages have only a small effect on employment and on annual work hours. Third, even after accounting for unobserved individual heterogeneity and job-specific heterogeneity, we find a strong negative tenure effect on job mobility, particularly for less educated workers. Fourth, consistent with job search theory, job changes are induced by high outside offers and deterred by the job-specific wage component of the current job.

Fifth, unemployment at the survey date is associated with a large decline of .6 log points in annual earnings. About two thirds of the reduction is due to work hours, which recover almost completely after one year. The other third is due to a decline of .2 in the log hourly wage rate. Lost tenure and a drop in the job-specific wage component contribute .06 and .02, respectively, to the wage reduction. The wage recovers by about .02 in the first year and more slowly after that.

Sixth, wages do not contain a random walk component, but are highly persistent. The persistence is the combined effect of permanent observed and un-

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

[8]Other recent papers that apply I-I to panel data include Bagger et al. (2011) and Nagypal (2007).

observed heterogeneity, the job-specific wage component, which depends positively on offers in previous jobs, and strong persistence in a stochastic component representing the value of the worker's general skills.

Seventh, shocks leading to unemployment or to job changes have large effects on the variance as well as the mean of earnings changes. Eighth, job shopping, the accumulation of tenure, and the growth in general skills account for log wage increases of .13, .11, and .61, respectively, over the first thirty years in the labor market.

Finally, the variance decompositions depend somewhat on the education subgroup, model specification, and assumptions about measurement error, but in all cases we find that job mobility and unemployment play a key role in the variance of career earnings. For our main specification, the job-specific hours and wage components, unemployment shocks, and job shocks together account for 43.0%, 53.2%, and 58.9% of the variance in lifetime earnings, wages, and hours, respectively. Job-specific wage shocks are more important than job-specific hours shocks for earnings. Job-specific wage shocks dominate for wages, while job-specific hours shocks dominate for hours. Education accounts for about 30% of the variance in lifetime earnings and wages, but makes little difference for hours. Variables determined by the first year of employment, including unobserved heterogeneity, education, and the initial draws of the general skill and job-specific wage components, collectively account for 55.3% for lifetime earnings, 44.6% for lifetime wages, and 39.5% for lifetime hours in the full sample, although these values differ somewhat across model specifications.

The paper continues in Section 2, where we present the earnings model. In Section 3, we discuss the data, which are drawn from the Panel Study of Income Dynamics (PSID), and in Section 4, we discuss estimation. Section 5 contains the main results, beginning with a discussion of the parameter estimates and then turning to an analysis of the fit of the model, impulse response functions to various shocks, and variance decompositions. In Section 6, we briefly discuss results for whites by education level. In Section 7, we present an alternative model of employment transitions and job mobility. We conclude with a summary of our main findings and a research agenda.

2. A MODEL OF EARNINGS DYNAMICS

We use two models in the paper, which we refer to as the baseline model and the multinomial model. They differ only in the specification of the equations governing transitions from employment to unemployment and job changes. The main features of the models are as follows. Labor market transitions, wages, and hours depend on three exogenous variables—race, education, and potential experience—as well as on two permanent unobserved heterogeneity components. The unobserved heterogeneity components can be labelled, loosely speaking, "unobserved productivity or ability" and "propensity to move." A typical worker enters the labor market after leaving school and receives initial draws of an employment status shock that determines employment and an autoregressive wage component capturing part of "general productivity" that has the same value in all jobs. The worker also draws an initial job-specific wage component and an initial job-specific hours component. There is state dependence in the value of the current job relative to unemployment and in the value of the current job relative to an alternative job. Consequently, there is state dependence in both employment and job-to-job transitions. Each period, an unemployed worker receives an unemployment transition shock. An employed worker receives a shock to the value of the current job, a shock affecting the value of moving, and a draw of the job-specific wage component for the new job. If the worker remains employed from one period to the next, then whether the worker changes jobs depends on the draw of the job-specific wage component for the new job, the current job-specific wage component, potential experience, job seniority, the two permanent heterogeneity terms, and an independent and identically distributed (i.i.d.) shock. A typical worker's wage depends on one of the heterogeneity terms (ability), the autoregressive general-productivity component, the job-specific wage component, potential experience, and seniority. Unemployment spells reduce the autoregressive general-productivity component, and workers draw new job-specific wage and hours components when they leave unemployment. Annual hours depend on employment status, the heterogeneity terms, the wage, and a job-specific hours component that is identical across jobs. Finally, earnings are determined by wages and hours.

2.1. *Equations of the Models*

A word about notation first. The subscript i, which we sometimes suppress, refers to the individual. The variable t_i is potential years of labor market experience of i for a particular observation. We sometimes refer to it as "time" even though it is potential experience rather than calendar time, and usually suppress the i subscript.[9] The subscript $j(t)$ refers to i's job at t. The notation $j(t)$ makes explicit the fact that individuals may change jobs. In particular, $j(t) \neq j(t-1)$ if i experiences a job change without being unemployed at either t or $t-1$ or if i is employed at t but was unemployed at $t-1$. The γ parameters refer to intercepts and to slope coefficients. For each intercept and slope parameter, the superscripts identify the dependent variable. The subscripts of slope parameters identify the explanatory variable. We use δ to denote coefficients on the fixed person-specific unobserved heterogeneity components μ_i and η_i, the job-match wage component $\upsilon_{ij(t)}$, and the job-specific hours component $\xi_{ij(t)}$. The superscripts for the δ parameters denote the dependent variable, and the subscripts μ and η identify the heterogeneity component. We use

[9]Wages, hours, and earnings are net of economy-wide year effects which we remove using a regression procedure discussed in Section 4.

ρ with appropriate subscripts to denote autoregression coefficients. The ε_{it}^k are i.i.d. $N(0, \sigma_k^2)$ random variables, where k corresponds to the dependent variable affected directly by ε_{it}^k. In what follows, we focus our discussion on the baseline model.

2.1.1. *Log Wages*

The log wage rate wage_{it} is determined by the following system of equations:

$$\text{wage}_{it} = E_{it}\text{wage}_{it}^{\text{lat}}, \tag{1}$$

$$\text{wage}_{it}^{\text{lat}} = \left[X_{it}\gamma_X^w + \gamma_{t^3}^w t^3\right] + P(\text{TEN}_{it})\gamma_{\text{TEN}}^w + \delta_\mu^w \mu_i + \omega_{it} + \upsilon_{ij(t)}, \tag{2}$$

$$\omega_{it} = \rho_\omega \omega_{i,t-1} + \gamma_{1-E_t}^\omega(1 - E_{it}) + \gamma_{1-E_{t-1}}^\omega(1 - E_{i,t-1}) + \varepsilon_{it}^\omega, \tag{3}$$

$$\upsilon_{ij(t)} = (1 - S_{it})\upsilon_{ij(t-1)} + S_{it}\upsilon'_{ij'(t)}, \tag{4}$$

$$\upsilon'_{ij'(t)} = \rho_\upsilon \upsilon_{ij(t-1)} + \varepsilon_{ij(t)}^\upsilon. \tag{5}$$

Equation (1) says that, for employed individuals (i.e., $E_{it} = 1$), wage_{it} equals the "latent wage" $\text{wage}_{it}^{\text{lat}}$. For an unemployed individual, $\text{wage}_{it}^{\text{lat}}$ captures the process for wage offers. At a given point in time, the individual might not have such an offer. The formulation parsimoniously captures the idea that worker skills and worker-specific demand factors evolve during an unemployment spell. It allows us to deal with the fact that wages are only observed for jobs that are held at the survey date.

Equation (2) states that $\text{wage}_{it}^{\text{lat}}$ depends on five components. The first is $X_{it}\gamma_X^w + \gamma_{t^3}^w t^3$, where X_{it} is a vector of exogenous variables consisting of the race indicator BLACK_i, years of education EDUC_i, a quadratic in t, and a constant, and t^3 is the cube of experience. Since we control for both tenure effects and gains from job shopping, the effect of t is a general human capital effect and/or an aging effect. The second term $P(\text{TEN}_{it})$ consists of the first four powers of employer tenure, TEN_{it}. The third term is the unobserved ability component μ_i.

The fourth term is a stochastic component ω_{it}, which, according to (3), depends on $\omega_{i,t-1}$, the current value and the first lag of unemployment $(1 - E_{it})$, and the error component ε_{it}^ω. The dependence of ω_{it} on its past reflects persistence in the market value of the general skills of i and/or the fact that employers base wage offers on past wages.

The fifth wage component is the job-match-specific term $\upsilon_{ij(t)}$. Each period, individuals are assigned a potential offer from an alternate job j' with job-specific component $\upsilon'_{ij'(t)}$. Its value depends on $\upsilon_{ij(t-1)}$ and the shock $\varepsilon_{ij(t)}^\upsilon$ as specified in (5). When agents leave unemployment or move from job to job without unemployment, the employer change indicator S_{it} is 1 and $\upsilon_{ij(t)}$ becomes $\upsilon'_{ij'(t)}$. Growth in $\upsilon_{ij(t)}$ with experience and with job mobility is endogenously determined through the influence of $\upsilon_{ij(t-1)}$ and $\upsilon'_{ij'(t)}$ on mobility, as we discuss momentarily. In standard search models with exogenous offer arrivals, the job-specific component of the offer, $v'_{ij'(t)}$, does not depend on $v_{ij(t-1)}$, although accepted offers $v_{ij(t)}$ do. In such models, the correlation between accepted offers $v_{ij(t)}$ and $v_{ij(t-1)}$ arises only because the reservation wage is a positive function of $v_{ij(t-1)}$. Nevertheless, we also allow offers $v'_{ij'(t)}$ to depend directly on $v_{ij(t-1)}$ through the parameter ρ_v, for three reasons. The first is that employers may base offers to prospective new hires in part on wages in the prior firm, including the firm-specific component. Bagger et al. (2011), building on Postel-Vinay and Robin (2002) and Postel-Vinay and Turon (2010), is one of a few recent papers in which outside firms tailor offers to surplus in the current job. This surplus will be related to $v_{ij(t-1)}$ to the extent that $v_{ij(t-1)}$ is the worker's portion of a job-specific productivity component. In contrast to those papers, however, we do not allow the current employers to change $v_{ij(t-1)}$ in response to outside offers. (Wages do change with ω_{it}.) The second reason $v'_{ij'(t)}$ depends on $v_{ij(t-1)}$ is that $v_{ij(t-1)}$ is not likely to be entirely job-specific in the presence of demand shocks affecting jobs in a narrowly defined industry, occupation, and region. The third is that the network available to an individual may be related to the quality of the job that he is in. As it turns out, our estimates of ρ_v are large—about .70.[10] We were not successful in limited experimentation with estimating models in which the link between $v_{ij(t)}$ and $v_{ij(t-1)}$ when agents move from job to job without unemployment ($\mathrm{JC}_{it} = 1$) differs from the link following unemployment, although standard job search models with exogenous layoffs imply that it should.

2.1.2. *Employment Transitions* (EE_t) *and Job Changes* (JC_t)*, Baseline Model*

The dummy EE_{it} indicates whether a worker who was employed in $t-1$ remains employed. In the baseline model, it is determined by

$$\begin{aligned}(6)\qquad \mathrm{EE}_{it} = I\big[&X_{i,t-1}\gamma_X^{\mathrm{EE}} + \gamma_{\mathrm{ED}}^{\mathrm{EE}}\min(\mathrm{ED}_{i,t-1}, 9) + \gamma_{\mathrm{TEN}}^{\mathrm{EE}}\mathrm{TEN}_{i,t-1} + \gamma_{w^s}^{\mathrm{EE}}\mathrm{wage}_{it}^s \\ &+ \delta_\mu^{\mathrm{EE}}\mu_i + \delta_\eta^{\mathrm{EE}}\eta_i + \varepsilon_{it}^{\mathrm{EE}} > 0\big] \quad \text{given} \quad E_{i,t-1} = 1,\end{aligned}$$

where $I(\cdot)$ is the indicator function, $\mathrm{ED}_{i,t-1}$ is lagged employment duration and is determined endogenously by $\mathrm{ED}_{it} = E_{it}(\mathrm{ED}_{i,t-1} + 1)$, and wage_{it}^s is what the wage would be in t if the individual were to continue employment in the job

[10]Industry-specific and/or occupation-specific human capital are not accounted for in the model and are likely to influence estimates of ρ_v more than ρ_ω, given that industry and occupation changes tend to occur across employers. They would also affect the estimates of the return to seniority that we import from Altonji and Williams (2005). See Neal (1995), Parent (2000), and Kambourov and Manovskii (2009) for somewhat conflicting evidence on the importance of occupation-specific, industry-specific, and firm-specific human capital. Extending the model to distinguish occupation and/or industry is conceptually straightforward but would require modeling of occupation and industry transitions and attention to measurement error. We leave this to future work.

held at $t-1$. The variable wage^s_{it} is the value of $\text{wage}^{\text{lat}}_{it}$ determined by (2), (3), (4), with $E_{it}=1$, $E_{i,t-1}=1$, $\text{TEN}_{it}=\text{TEN}_{i,t-1}+1$, and $S_{it}=0$. The vector $X_{i,t-1}$ is the same as X_{it} except that it contains $(t-1)$ and $(t-1)^2$ rather than t and t^2. In the econometric work, we exclude $\text{TEN}_{i,t-1}$ because, in simulation experiments for this specification, we had trouble distinguishing the effects of $\text{TEN}_{i,t-1}$ and $\text{ED}_{i,t-1}$. Standard labor supply models imply that employment at t should depend on the current wage opportunity, which we proxy with wage^s_{it}. EE_{it} also depends on the permanent ability component μ_i as well as the hours preference and mobility component η_i ("propensity to move"). Both μ_i and η_i also directly affect transitions out of unemployment, job changes, and work hours, but η_i is excluded from the wage model. One may think of η_i as a factor that is related to labor supply and to job and employment mobility preferences but not productivity.

The job change equation, conditional on remaining employed, is

$$\begin{aligned}(7)\qquad \text{JC}_{it} = I\big[&X_{i,t-1}\gamma^{\text{JC}}_X + \gamma^{\text{JC}}_{\text{TEN}}\text{TEN}_{i,t-1} + \delta^{\text{JC}}_{v'j'(t)}v'_{ij'(t)} + \delta^{\text{JC}}_{vj(t-1)}v_{ij(t-1)} \\ &+ \delta^{\text{JC}}_{\mu}\mu_i + \delta^{\text{JC}}_{\eta}\eta_i + \varepsilon^{\text{JC}}_{it} > 0\big] \quad \text{given} \quad E_{it}=E_{i,t-1}=1.\end{aligned}$$

Standard job search and job matching models predict a negative coefficient on $v_{ij(t-1)}$, since higher values of the job match component of the current job should reduce search activity and raise the reservation wage. In the model, each worker is assigned a potential draw of $v'_{ij'(t)}$ based on (5), which we discuss momentarily. Search models predict a positive coefficient on $v'_{ij'(t)}$, but the magnitude should depend on the probability that the worker actually receives the offer. That is, the relative magnitudes of the two coefficients should depend on offer arrival rates and need not be equal.[11] We include $\text{TEN}_{i,t-1}$ as well as $(t-1)$ because models of firm-financed or jointly financed specific capital investment suggest that it will play a role, and the decline in separation rates with $\text{TEN}_{i,t-1}$ in cross section data is very strong. However, little is known about how much of the association between $\text{TEN}_{i,t-1}$ and JC_{it} is causal because of the difficulty of distinguishing state dependence from the individual heterogeneity (μ and η) and job match heterogeneity (v) in dynamic discrete choice models, particularly when data are missing on early employment histories for most sample members. Indeed, Buchinsky et al. (2010) is the only other study that we know that accounts for both individual and job-specific heterogeneity and deals with initial conditions problems when estimating the effects of TEN and t on job changes.[12]

[11]One could introduce parameters corresponding to fixed offer arrival rates for employed workers and for unemployed workers into the model and add the value of $v'_{ij'(t)}$ into the unemployment equation. Low, Meghir, and Pistaferri (2010) worked with such a specification. In our job change equation, $v_{ij(t-1)}$ may reduce mobility both because it raises the reservation wage and because it lowers search intensity.

[12]Buchinsky et al. (2010) also found negative effects in a simultaneous model of wages, employment, and job changes. Farber (1999) discussed models of the effect of tenure on mobility

In this model of employment transitions and job changes, the main distinction is between job changes from employment and job changes that involve unemployment. We believe that this is the most important distinction for the determination of wages and annual work hours. Equation (6) determines whether a worker who was employed in $t-1$ has an employment option that is better than unemployment, while (7) determines whether that option is a new job or the old job. Given the sequential structure, the probability that $\mathrm{JC}_{it} = 1 | E_{i,t-1} = 1$ depends on the variables that appear in (6) as well as (7), including the current wage wage^s_{it} and $\mathrm{ED}_{i,t-1}$.

In Section 7 and Appendix C of the Supplemental Material (Altonji, Smith, and Vidangos (2013)), we present a multinomial formulation of employment and job transitions consisting of equations for the value of staying with the current employer relative to unemployment and the value of moving to a new employer relative to unemployment. In some respects, we prefer the multinomial formulation, but it is less well-behaved numerically. This poses problems when we move to education subgroups, particularly with respect to the feasibility of bootstrapping. Consequently, we rely primarily on the baseline model. Most of the results about earnings dynamics are not sensitive to the choice between the two models.

2.1.3. *Unemployment to Employment Transition (*UE_t*)*

Movement from unemployment to employment is determined by

$$\begin{aligned}\mathrm{UE}_{it} = I\big[&X_{i,t-1}\gamma_X^{\mathrm{UE}} + \gamma_{\mathrm{UD}}^{\mathrm{UE}}\mathrm{UD}_{i,t-1} + \delta_\mu^{\mathrm{UE}}\mu_i \\ &+ \delta_\eta^{\mathrm{UE}}\eta_i + \varepsilon_{it}^{\mathrm{UE}} > 0\big] \quad \text{given} \quad E_{i,t-1} = 0,\end{aligned} \tag{8}$$

where $\mathrm{UD}_{i,t-1}$ is the number of years unemployed at the survey date and $\mathrm{UD}_{it} = (1 - E_{it})(\mathrm{UD}_{i,t-1} + 1)$. Because there are relatively few multiyear unemployment spells, we end up restricting $\gamma_{\mathrm{UD}}^{\mathrm{UE}}$ to 0 in the empirical work. We experimented with specifications containing the lagged latent wage rate or the expected value of the period-t wage, but had difficulty pinning down the effects of these variables, perhaps because we observe relatively few unemployment spells. We do include the heterogeneity components μ_i and η_i.

Note that E_{it} is given by

$$E_{it} = \mathrm{EE}_{it}E_{i,t-1} + \mathrm{UE}_{it}(1 - E_{i,t-1}).$$

2.1.4. *Log Annual Hours*

Log annual work hours are determined by

$$\mathrm{hours}_{it} = X_{it}\gamma_X^h + \gamma_{t^3}^h t^3 + \big(\gamma_E^h + \xi_{ij(t)}\big)E_{it} + \gamma_w^h \mathrm{wage}_{it}^{\mathrm{lat}} + \delta_\mu^h\mu_i + \delta_\eta^h\eta_i + \varepsilon_{it}^h. \tag{9}$$

---

and surveyed the empirical evidence. He presented evidence showing a negative effect of tenure when one uses prior mobility as a control for individual heterogeneity.

The hours_{it} equation includes X_{it} and t^3. It also includes η_i, μ_i, and the product of the job-specific hours component $\xi_{ij(t)}$ and E_{it}. We include $\xi_{ij(t)}$ because there is strong evidence that work hours are heavily influenced by a job-specific component. This component presumably reflects work schedules imposed by employers.[13] A new value of $\xi_{ij(t)}$ is drawn when individuals take a new job. The i.i.d. error component ε^h_{it} picks up transitory variation in straight time hours worked, overtime, multiple job holding, and unemployment conditional on employment status at the survey. It may reflect temporary shifts in worker preferences as well as hours constraints.

Hours also depend on $\text{wage}^{\text{lat}}_{it}$ and E_{it}. For most observations, $\text{wage}^{\text{lat}}_{it}$ is the actual wage. However, many individuals who are unemployed at the survey date work part of the year. For these individuals, $\text{wage}^{\text{lat}}_{it}$ is 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 regard the coefficient on the latent wage as the response to a relatively permanent wage change rather than the Frisch elasticity.

2.1.5. *Log Earnings*

$$\begin{aligned} \text{(10)} \qquad & \text{earn}_{it} = \gamma^e_0 + \gamma^e_w \text{wage}^{\text{lat}}_{it} + \gamma^e_h \text{hours}_{it} + e_{it}, \\ & e_{it} = \rho_e e_{i,t-1} + \varepsilon^e_{it}. \end{aligned}$$

Log earnings earn_{it} depends on $\text{wage}^{\text{lat}}_{it}$ and hours_{it}. The coefficients γ^e_w and γ^e_h might differ from 1 for a number of reasons, including overtime, multiple job holding, bonuses and commissions, job mobility, and the fact that, for some salaried workers, the wage reflects a set work schedule but annual hours worked may vary. Note that X_{it} is excluded but influences earnings through $\text{wage}^{\text{lat}}_{it}$ and hours_{it}. We also include a first-order autoregressive error component e_{it} to capture some of these factors. In previous drafts of the paper, we freely estimated γ^e_w and γ^e_h and obtained values very close to 1 for some more restrictive specifications of the model than the one we use here. For richer versions of the model, it is helpful to restrict the coefficients to be 1, which we do below.

2.1.6. *Error Components and Initial Conditions*

The fixed person-specific error components μ_i and η_i are $N(0,1)$, i.i.d. across i, independent of each other, and independent of all transitory shocks and measurement errors. Without loss of generality, we impose the sign normalizations $\delta^w_\mu > 0$ and $\delta^{\text{JC}}_\eta > 0$.

[13]For evidence, see Altonji and Paxson (1986) and Blundell, Brewer, and Francesconi (2008).

The job match hours component $\xi_{ij(t)}$ and the innovation $\varepsilon^{v}_{ij(t)}$ in equation (5) are $N(0, \sigma^2_{\xi})$ and $N(0, \sigma^2_{v})$, respectively. The shocks ε^{h}_{it}, $\varepsilon^{\omega}_{it}$, and ε^{e}_{it} are $N(0, \sigma^2_{k})$, where $k = h, \omega$, and e. The shocks $\varepsilon^{\mathrm{EE}}_{it}$, $\varepsilon^{\mathrm{UE}}_{it}$, $\varepsilon^{\mathrm{JC}}_{it}$ are $N(0, 1)$. They are i.i.d. across i and t and independent from one another.

The initial conditions are

$$\text{(11)} \qquad \text{Employment:} \quad E_{i1} = I\big[b_{0g} + \delta^{\mathrm{EE}}_{\mu}\mu_i + \delta^{\mathrm{EE}}_{\eta}\eta_i + \varepsilon^{\mathrm{EE}}_{i1} > 0\big],$$

$$\text{(12)} \qquad \text{Wages:} \quad \text{wage}^{\text{lat}}_{i1} = X_{i1}\gamma^{w}_{X} + \gamma^{w}_{t3} + \delta^{w}_{\mu}\mu_i + \omega_{i1} + v_{ij(1)},$$

$$\text{General productivity:} \quad \omega_{i1} \sim N\big(0, \sigma^2_{\omega 1,g}\big),$$

$$\text{Wage job match:} \quad v_{ij(1)} \sim N\big(0, \sigma^2_{v1}\big),$$

$$\text{Earnings error:} \quad e_{i1} \sim N\big(0, \sigma^2_{e}\big),$$

Other initial conditions:

$$\text{TEN}_{i1} = 0, \quad \text{ED}_{i1} = E_{i1}, \quad \text{UD}_{i1} = 1 - E_{i1}, \quad \text{JC}_{i1} = 0.$$

The random components ω_{i1}, $v_{ij(1)}$, and e_{i1} are mutually independent and independent of the shocks in the model. The intercept b_{0g} of the initial employment condition and the variance of initial wages $\sigma^2_{\omega 1,g}$ depend on the race-education group g, where the groups are defined by (BLACK & EDUC $\leq$ 12), (BLACK & EDUC $>$ 12), (not BLACK & EDUC $\leq$ 12), and (not BLACK & EDUC $>$ 12).

2.1.7. *Measurement Error and Observed Wages, Hours, and Earnings*

The observed (measured) variables are

$$\text{(13)} \qquad \text{wage}^{*}_{it} = E_{it}\big(\text{wage}^{\text{lat}}_{it} + m^{w}_{it}\big),$$

$$\text{(14)} \qquad \text{hours}^{*}_{it} = \text{hours}_{it} + m^{h}_{it},$$

$$\text{(15)} \qquad \text{earn}^{*}_{it} = \text{earn}_{it} + m^{e}_{it}.$$

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

2.2. *Additional Discussion of the Model*

When interpreting results for EE_{it} and JC_{it}, one must keep in mind that our employment indicator refers to the survey date. We undoubtedly miss short spells of unemployment that fall between surveys. Due to data limitations, we cannot tell whether a person has changed jobs between surveys only once or multiple times. Furthermore, if a person is employed at $t-1$, unemployed for part of the year, and employed in a new job at t, we would count this as a job-to-job change even if, for example, the job change is due to a layoff into unemployment. A relatively simple alternative would be to make use of information on the number of weeks that the individual was unemployed during the year. However, one would want to distinguish between short spells of unemployment that are associated with temporary layoffs with the strong expectation of recall and unemployment spells due to a permanent layoff. This is possible only at the survey date. Fortunately, earnings depend on employment through annual work hours, and the transitory error component in the hours equation should capture the effect on hours of unemployment spells of varying duration. The 25th, 50th, 75th, and 90th percentiles of hours of unemployment are 160, 688, 1080, and 1560 when $E_{it} = 0$, and 0, 0, 0, and 80 when $E_{it} = 1$.[14]

We have not considered models with an ARCH error structure. However, the model implies that the variance of wage, hours, and earnings changes are state dependent and also depend on t. This is because the odds of a job change and an unemployment spell depend on TEN, ED, t, and $v_{ij(t)}$ and because job changes and unemployment spells are associated with innovations in $v_{ij(t)}$, $\xi_{ij(t)}$, and ω_{it}.[15] The variances also depend on the permanent components of X_{it} (education and race) and on the unobserved heterogeneity components μ_i and η_i.

3. DATA

We use the 1975–1997 waves of the PSID to assemble data that refer to the calendar years 1975–1996. Programs and data are in the Data Appendix of the Supplemental Material. We use the stratified random sample (SRC), but also include nonsample members who married PSID sample members.[16] The sample is restricted to single or married male household heads. We present estimates for the full sample, and for whites by education level.

Observations for a given person-year are used if the person is between age 18 and 62, was working, temporarily laid off, or unemployed at the survey date in a given year, was not self-employed, had valid data on education (EDUC), and had no more than 40 years of potential experience. We treat persons on temporary layoff as employed. We eliminate a small number of observations in

[14] This calculation excludes years after 1993 because the edited hours of unemployment variable is not available for later years. We could have specified the model on a quarterly or monthly basis.

[15] An earlier draft of this paper, Altonji, Smith, and Vidangos (2009; henceforth ASV) presented results for an alternative model (Model B) that excludes $v_{ij(t)}$ but allows the variance of the shocks $\varepsilon_{it}^{\omega}$ to depend on whether the person is starting a new job.

[16] In ASV (2009), we reported similar results for the baseline model using the combined SRC and SEO samples. The SEO sample consists primarily of households that were low income in 1968 and substantially overrepresents blacks.

TABLE I

COMPOSITION OF PSID SRC SAMPLE BEFORE SAMPLE SELECTION BASED ON EMPLOYMENT STATUS[a]

Emp. Status	Percentage
Working	91.0
Temp. Laid off	1.5
Unemployed	3.8
Retired	1.0
Disabled	1.2
Housewife	0.2
Student	1.1
Other	0.2

[a]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.

which the individual reports being retired, disabled, a housewife, a student, or "other" (see Table I and Table II).[17]

Potential experience t_i is $\text{age}_{it} - \max(\text{EDUC}_i, 10) - 5$. BLACK_i is 1 if the individual is black and 0 otherwise. ED_{it} is the number of years in a row that a

TABLE II

PERCENTAGE OF OBSERVATIONS EXCLUDED BASED ON EMPLOYMENT STATUS, BY POTENTIAL EXPERIENCE (t)[a]

t	Percentage	t	Percentage	t	Percentage	t	Percentage
1	21.0[b]	11	1.7	21	1.9	31	5.0
2	11.1	12	2.1	22	1.7	32	5.5
3	7.5	13	1.9	23	2.8	33	7.6
4	6.0	14	2.5	24	3.3	34	7.6
5	5.2	15	1.8	25	3.4	35	8.4
6	3.0	16	2.0	26	3.2	36	11.1
7	3.0	17	2.1	27	3.6	37	12.1
8	2.3	18	2.0	28	3.6	38	12.7
9	2.2	19	1.8	29	4.0	39	16.8
10	2.0	20	2.1	30	4.8	40	20.4[c]

[a]The table presents the percentage of observations excluded, based on employment status at the survey date, for each value of potential experience (t).

[b]Of those excluded at $t = 1$, 97.3% are students.

[c]Of those excluded at $t = 40$, 70.8% are retired, 23.9% disabled.

[17]We allow persons to come out of retirement and include future observations following a retirement spell if the individual is working, temporarily laid off, or unemployed. The percentage of the PSID sample who report their employment status in a given year as disabled is 1.2.

person is employed at the survey date. In 1975 and for persons who joined the sample after 1975, we set ED_{it} to tenure with the current employer (TEN).[18] The variable $\mathrm{UD}_{i,t-1}$ is the number of consecutive years up to $t-1$ that the individual has not been employed at the survey date. We set $\mathrm{UD}_{i,t-1}$ to 0 if the first time we observe i is in year t. Few unemployment spells exceed one year, so the error is probably small. The wage measure 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.[19]

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.

We restrict the sample to individuals who are observed for at least three years because many of the key equations in the auxiliary model involve lags. The sample contains 2712 individuals who contribute a total of 31,330 person-year observations. The sample is highly unbalanced.[20] As we have already noted, an advantage of indirect inference is that, by incorporating the sample selection process into the simulation, one can handle unbalanced data. We assume that observations are missing at random, although there is reason to believe that the heterogeneity components and shocks influence attrition from the sample.[21]

In Table III, we present the mean, standard deviation, minimum, and maximum of the key variables in our baseline sample. The mean of E_{it} is .966, so we

[18]An alternative would be to apply exactly the same censoring that occurs in the PSID to the simulated data. The value of ED_{it} for the simulated case for i would be set to tenure when t in the simulated case is equal to $t_{i\min}$, the value of t for the first observation on i in the PSID.

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

[20]Each individual contributes between 3 and 22 observations. The 5th, 25th, median, 75th, and 95th percentile values of the number of observations a given individual contributes are 4, 7, 11, 16, 21. The number of observations per year varies from 702 in 1975 to 1571 in 1993 and exceeds 1100 in all years except 1975.

[21]In principle, one could augment the model with an attrition equation. Alternatively, it would be straightforward to simply use sample weights to reweight the PSID when evaluating the likelihood function of the auxiliary model if suitable weights were available. However, PSID sample weights are designed to keep the data representative of successive cross sections of the U.S. population that originate in the families present in the base year. They do not adjust for factors that alter the U.S. population, such as differences in birth rates by race or education. Furthermore, there are no sample weights for persons who move into PSID households through marriage. Consequently, we do not use weights.

TABLE III

DESCRIPTIVE STATISTICS—PSID SRC SAMPLE[a]

Variable	Obs.	Mean	St. Dev.	Min	Max
Potential Experience	31,330	17.735	9.582	1	40
Black	31,330	0.062	0.242	0	1
Education (years)	31,330	13.336	2.307	6	17
E_t	31,330	0.966	0.181	0	1
$E_t \| E_{t-1} = 1$ (EE_t)	28,170	0.975	0.157	0	1
$E_t \| E_{t-1} = 0$ (UE_t)	872	0.774	0.418	0	1
$\mathrm{JC}_t \| E_{t-1} = 1$	28,170	0.112	0.315	0	1
ED_t	30,742	10.416	7.959	0	42.25
UD_t	31,330	0.043	0.257	0	7
TEN_t	30,065	8.420	8.015	0	42.25
$\mathrm{wage}_t^* \| E_t = 1$[b]	30,265	2.737	0.494	1.14	4.96
hours_t^{*}[b]	31,330	7.760	0.284	5.30	8.42
earn_t^{*}[b]	31,330	3.546	0.658	−5.17	6.54

[a]The table presents descriptive statistics for key variables in our baseline PSID SRC sample.

[b]Economy-wide year effects have been removed from variables wage_t^*, hours_t^*, and earn_t^* by first regressing measured hours, wages, and earnings on a cubic in potential experience, BLACK, EDUC, and a set of year dummies, and subtracting the estimated year effects. The reference year is 1996. The effects of the potential experience polynomial, BLACK, or EDUC have not been removed from these variables. See discussion in Section 4.

observe relatively few unemployment spells. Note also that the mean of EE_{it} is .975. Given these magnitudes, even relatively large movements in the latent variable index determining EE_{it} have only a small effect on whether EE_{it} is 1 or 0.

4. ESTIMATION METHODOLOGY

We begin with a brief overview of our estimation procedure. We then define the auxiliary model used in the estimation procedure as well as additional moment conditions that we use.[22]

[22]We remove economy-wide year effects by first regressing measured hours, wages, and earnings on X_{it} and a set of year dummies and subtracting the estimated year effects. The variables hours_{it}^*, wage_{it}^*, and earn_{it}^* refer to the adjusted measures. We *do not* subtract the effects of X_{it}, as is done in many studies of earnings dynamics. The coefficients on X_{it} are estimated by indirect inference, simultaneously with the other parameters of the model, so that sample selection in employment can be accounted for. Simultaneously estimating the large number of year effects with the rest of the model parameters would dramatically increase computational complexity. One could account for sampling error in the year effects in the parametric bootstrap procedure that we discuss below, but we have not done so. The correction is unlikely to make much difference because estimated standard errors for the year effects are relatively small (about .016 for the wage, .01 for hours, and .02 for earnings), the year effects explain less than 1 percent of the variance in earnings, and they are very weakly correlated with BLACK_i, EDUC_i, and t_i.

4.1. *Indirect Inference (I-I)*

For clarity, we will refer to the model presented in Section 2 as the "structural" model, even though the model does not express the parameters of the decision rules for EE, UE, JC, etc., in terms of preference parameters and parameters governing job search, mobility, and exogenous layoffs. We denote the k "structural" parameters by β. Indirect inference involves the use of an "auxiliary" statistical model that captures properties of the data. This auxiliary model has $p \geq k$ parameters θ. The method involves simulating data from the structural model (given a hypothesized value of β) and choosing the estimator $\hat{\beta}$ of β to make the simulated data match the actual data as closely as possible according to some criterion that involves θ.

Let the observed data consist of a set of observations on N individuals in each of T periods: $\{y_{it}, x_{it}\}$, $i = 1, \dots, N$, $t = 1, \dots, T$, where y_{it} is endogenous to the model and x_{it} is exogenous. The auxiliary model parameters θ can be estimated using the observed data as the solution to

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

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 assumed values of β, we use the structural model to generate M statistically independent simulated data sets $\{\tilde{y}_{it}^m(\beta)\}$, $m = 1, \dots, M$. Each of the M simulated data sets has N individuals and is constructed using the same observations on the exogenous variables, x. For each of the M simulated data sets, we compute $\tilde{\theta}_m(\beta)$ as

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

where the likelihood function associated with the auxiliary model is evaluated using the mth simulated data set $\tilde{y}_m(\beta) \equiv \{\tilde{y}_{it}^m(\beta)\}$ rather than the real data. Denote the average of the estimated parameter vectors by $\tilde{\theta}(\beta) \equiv M^{-1} \sum_{m=1}^{M} \tilde{\theta}_m(\beta)$.

I-I generates an estimate $\hat{\beta}$ of the structural parameters by choosing β to minimize the distance between $\hat{\theta}$ and $\tilde{\theta}(\beta)$ according to some metric. As described in Keane and Smith (2003) and elsewhere, there are (at least) three possible ways to specify such a metric. Here we choose $\hat{\beta}$ to minimize the difference between the constrained and unconstrained values of a pseudo likelihood function of the auxiliary model evaluated using the observed data. In particular, we calculate

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

Gourieroux, Monfort, and Renault (1993) showed that $\hat{\beta}$ is a consistent and asymptotically normal estimate of the true parameter vector β_0. The reason is that as N becomes large, holding M and T fixed, $\tilde{\theta}(\hat{\beta})$ and $\hat{\theta}$ both converge to the same "pseudo" true value $\theta_0 = h(\beta_0)$, where h is a nonstochastic function.

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 some cases, it is convenient to estimate auxiliary models in which missing observations are replaced with some arbitrary value such as 0 or the sample mean. In such circumstances, the same principle applies: use the same arbitrary values in both the simulated and observed data sets. The fact that the first period that we observe people is typically after $t_i = 1$ would pose extremely serious "initial conditions" problems if we were using standard panel data methods, but is handled naturally by I-I because the missing early observations will affect the probability limits of $\tilde{\theta}(\hat{\beta})$ and $\hat{\theta}$ in the same way.[23]

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

To circumvent these difficulties, we use Keane and Smith's (2003) modification to I-I, which they called generalized indirect inference. Suppose that the simulated value of a binary variable $\tilde{y}_{it}^m$ 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 $\tilde{\theta}(\beta)$, we use a continuous function $g(\tilde{u}_{it}^m(\beta); \lambda)$ of the latent utility. The function g is chosen so that, as a 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 as the observed sample size goes to in-

[23] Heckman (1981), Wooldridge (2005), and others discussed how to deal with initial conditions by using a flexible form for the distribution of the first observation for each i and its relationship to error distributions in the outcome equations for $t_i > t_{i\min}$, where $t_{i\min}$ is first observation on i. Using their approaches, the econometrician does not impose any links between the parameters of the main model and the parameters of the initial condition. The parameters depend on $t_{i\min}$, which varies substantially in our sample. Consequently, these approaches to the initial conditions problem are not attractive in a setting such as ours, which involves a multiple equation model with a large number of endogenous state variables and substantial variation in $t_{i\min}$.

finity ensures that $\tilde{\theta}(\beta_0)$ converges to θ_0, thereby delivering consistency of the I-I estimator of β_0. Our choice of g is

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

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

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

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

4.2. *The Auxiliary Model*

Our auxiliary model is the sum of two parts. The first part provides information on the structural model parameters γ_X^w, $\gamma_{t^3}^w$, γ_X^h, and $\gamma_{t^3}^h$ that determine the effects of BLACK_i, EDUC_i, and t_i on wages, hours, and earnings. It consists of the equations

$$\text{wage}_{it}^*=\big[X_{it},t^3\big]\theta_1^w+u_{it}^w,$$

$$\text{hours}_{it}^*=\big[X_{it},t^3\big]\theta_1^h+u_{it}^h,$$

and the associated sum of squares criterion function[24]

$$\begin{aligned}&\mathcal{L}_1\big(\text{wage}^*_{it}, \text{hours}^*_{it}; X_{it}, t_i^3, \tilde{\theta}_1^w(\beta), \tilde{\theta}_1^h(\beta)\big)\\ &\quad = \sum_{i,t}\big(\text{wage}^*_{it} - \big[X_{it}, t_i^3\big]\tilde{\theta}_1^w\big)^2 + \sum_{i,t}\big(\text{hours}^*_{it} - \big[X_{it}, t_i^3\big]\tilde{\theta}_1^h\big)^2.\end{aligned}$$

The second and main part of the auxiliary model consists of a system of seemingly unrelated regressions (SUR) with seven equations and 25 covariates that are common to all seven equations. Let $\widetilde{\text{wage}}^*_{it} = \text{wage}^*_{it} - [X_{it}, t_i^3]\tilde{\theta}_1^w$, $\widetilde{\text{hours}}^*_{it} = \text{hours}^*_{it} - [X_{it}, t_i^3]\tilde{\theta}_1^h$, and $\widetilde{\text{earn}}^*_{it} = \text{earn}^*_{it} - [X_{it}, t_i^3]\tilde{\theta}_1^w - [X_{it}, t_i^3]\tilde{\theta}_1^h$. One may write the system as

$$Y_{it} = \Pi Z_{it} + u_{it}; \quad u_{it} \sim N(0, \Sigma); \quad u_{it} \text{ i.i.d. over } i \text{ and } t, \tag{16}$$

where

$$\begin{aligned}Y_{it} = \big[&E_{it}E_{i,t-1}, E_{it}(1 - E_{i,t-1}), \text{JC}_{it}E_{it}E_{i,t-1},\\ &\widetilde{\text{wage}}^*_{it}, \widetilde{\text{hours}}^*_{it}, \widetilde{\text{earn}}^*_{it}, \ln\big(1 + \widetilde{\text{wage}}^{*2}_{it}\big)\big]'\end{aligned}$$

and

$$\begin{aligned}Z_{it} = \big[&\text{Const}, (t_i - 1), (t_i - 1)^2, \text{BLACK}_i, \text{EDUC}_i, \text{ED}_{i,t-1}, \text{UD}_{i,t-1},\\ &\text{TEN}_{i,t-1}, E_{i,t-1}E_{i,t-2}, E_{i,t-2}E_{i,t-3}, E_{i,t-1}(1 - E_{i,t-2}),\\ &E_{i,t-2}(1 - E_{i,t-3}), \text{JC}_{i,t-1}E_{i,t-1}E_{i,t-2}, \text{JC}_{i,t-2}E_{i,t-2}E_{i,t-3},\\ &\widetilde{\text{wage}}^*_{i,t-1}, \widetilde{\text{wage}}^*_{i,t-2}, \widetilde{\text{hours}}^*_{i,t-1}, \widetilde{\text{hours}}^*_{i,t-2}, \widetilde{\text{earn}}^*_{i,t-1}, \widetilde{\text{earn}}^*_{i,t-2},\\ &\widetilde{\text{wage}}^*_{i,t-1}(t_i - 1), \widetilde{\text{wage}}^*_{i,t-1}(t_i - 1)^2, \widetilde{\text{wage}}^*_{i,t-1}\text{JC}_{it},\\ &\widetilde{\text{wage}}^*_{i,t-2}\text{JC}_{i,t-1}, \widetilde{\text{wage}}^*_{i,t-2}E_{i,t-1}\big]'.\end{aligned} \tag{17}$$

In estimating the model, we use the likelihood function $\mathcal{L}_2(Y; Z, \tilde{\theta}_2(\beta))$ that corresponds to (16), where $\theta_2 = (\Pi, \Sigma)$. The assumption $u_{it} \sim N(0, \Sigma)$, with u_{it}

[24]In ASV (2009) we estimated γ_X^w, $\gamma_{t^3}^w$, γ_X^h, and $\gamma_{t^3}^h$ directly from a first stage regression of wages and hours on BLACK, EDUC, and a polynomial in t. However, this procedure is subject to selection bias associated with the employment decision. This is why we now treat θ_1^w and θ_1^h as auxiliary model parameters rather than structural parameters. When we used the two-step procedure, we included a second constant in the hours equation that has no effect on earn^*_{it} because it improved the fit of the model. We kept the second constant in the hours model when we changed the estimation procedure, but the estimated value is close to 0, so dropping it would make little difference. The constant γ_h^0 reported in Table IV has a coefficient of 1 in the earnings equation.

i.i.d. over i and t, is false for several reasons, including the fact that Y contains binary variables and that both $\widetilde{\text{wage}}^*_{it}$ and $\ln(1+\widetilde{\text{wage}}^{*2}_{it})$ appear. The fact that we use a misspecified likelihood affects efficiency but not consistency.

Our choice of what to include in (16) is motivated by the following principles. First, we use a rich auxiliary model rather than focus on a few features of the data. We do this because our model is intended to explain both contemporaneous and dynamic interrelationships among key labor market variables. Given our objective, it makes more sense to use a rich auxiliary model that can capture these relationships rather than focus on a few features of the data.

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

Third, since the purpose of the structural model is to explain the behavior of its dependent variables, we use each of the dependent variables of the structural model as a dependent variable in the auxiliary model. This accounts for the first six equations. It is also natural to use the explanatory variables in the structural model as explanatory variables in the auxiliary model. This accounts for the presence of (t_i-1), $(t_i-1)^2$, EDUC_i, BLACK_i, $\text{ED}_{i,t-1}$, and $\text{TEN}_{i,t-1}$. We also include $\text{UD}_{i,t-1}$ even though we constrain $\gamma^{\text{UE}}_{\text{UD}}$ to equal 0. Since the model is dynamic and includes state dependence terms in most equations, we include two lags of each of the dependent variables for the first six equations. The lags help distinguish between state dependence and heterogeneity. We include the interaction terms $\widetilde{\text{wage}}^*_{i,t-1}(t_i-1)$, $\widetilde{\text{wage}}^*_{i,t-1}(t_i-1)^2$ to capture change with potential experience in the degree of persistence in wages. Finally, the terms $\widetilde{\text{wage}}^*_{i,t-1}\text{JC}_{it}$, $\widetilde{\text{wage}}^*_{i,t-2}\text{JC}_{i,t-1}$, and $\widetilde{\text{wage}}^*_{i,t-2}E_{i,t-1}$ capture effects of current and past job mobility and past employment on state dependence in wages. The seventh equation has $\ln(1+\widetilde{\text{wage}}^{*2}_{it})$ as the dependent variable. This helps identify parameters of the model that influence the level and change in the variance of wages and earnings over time, including the dependence on job changes and unemployment. Appendix F of the Supplementary Material reports PSID estimates of Π and Σ.

Since Π has 25×7 elements and Σ is a 7×7 covariance matrix with 28 unique elements, the auxiliary model has 203 parameters. $\mathcal{L}_1$ has another 10 slope coefficients and 2 variance terms. In contrast, the model has only 55 parameters that we estimate by I-I (not counting the measurement error parameters, tenure coefficients, and ρ_ω). As we discuss momentarily, a few extra parameters are identified in part using additional moment conditions. Consequently, the number of features of the data used to fit the structural model greatly exceeds the number of parameters.

The criterion function is Weight $\cdot \mathcal{L}_1 + \mathcal{L}_2$. We set Weight to a large value to give primacy to $\mathcal{L}_1$ for purposes of identifying γ_X^w, $\gamma_{t^3}^w$, γ_X^h, and $\gamma_{t^3}^h$, although the estimator is consistent for any positive value.

4.3. *Additional Moments and Other Information Sources*

The auxiliary model is poorly suited to identify the parameters of equations (11) and (12) for initial employment status and the initial wage because, in (17), the first three observations for each individual are lost due to lags. (In the case of $\mathcal{L}_1$, we do use the first three observations.) It also makes it difficult to identify changes with experience in the variance of shocks at the beginning of a career. To address this, we incorporated additional moment conditions.

In the case of initial employment, we estimate the intercepts b_{0g} as $\hat{b}_{0g} = \hat{b}_{0g}^* \hat{\sigma}_{E1}$, where $\hat{\sigma}_{E1} = \sqrt{(\hat{\delta}_\mu^{\mathrm{EE}})^2 + (\hat{\delta}_\eta^{\mathrm{EE}})^2 + 1}$ for four groups g defined by race and whether the person has more than a high school education. We estimate $\hat{b}_{0g}^*$ using a Probit regression of E_{it} on BLACK_i and the indicator for $(\mathrm{EDUC}_i > 12)$. We exclude the interaction to avoid computational problems in computing bootstrap standard errors that arise from the small sample sizes for blacks. We use the first five years rather than simply the first because we have relatively few observations for each group when $t = 1$.[25]

To identify $\sigma_{\omega 1}$, we use the fact that the model implies that the variance of the observed wage residuals, $\widetilde{\mathrm{wage}}_{i1}^*$, of an employed individual from race-education group g is

$$\begin{aligned}\mathrm{Var}\big(\widetilde{\mathrm{wage}}_{i1}^*; g\big) &\equiv \mathrm{Var}\big(\mathrm{wage}_{i1}^* - \big[X_{it}, t^3\big]\theta_1^w\big) \\ &\approx \big(\delta_\mu^w\big)^2 + \sigma_{\omega 1,g}^2 + \sigma_{v1}^2 + \sigma_{mw}^2.\end{aligned}$$

The relationship is approximate because the auxiliary model parameters differ slightly from the structural parameters $[\gamma_X^w, \gamma_{t^3}^w]$ because of sample selection. Because of sample size considerations, we estimate $\mathrm{Var}(\widetilde{\mathrm{wage}}_{i1}^*; g)$ as the variance of (residual) wage observations in the PSID corresponding to $t \leq 5$. $\widehat{\mathrm{Var}}(\widetilde{\mathrm{wage}}_{i1}^*; g)$ equals .066 for blacks with a high school degree or less, .099 for

[25] As it turns out, $\hat{b}_{0g}^*$ is 1.188 for blacks with a high school degree or less, 1.484 for blacks with more than high school, 1.424 for whites with high school or less, and 1.720 for whites with more than high school. The values are similar when we include $\mathrm{BLACK} * (\mathrm{EDUC} > 12)$: 1.124, 1.412, 1.429, and 1.716. We obtain similar results for other model parameters when we constrain $\hat{b}_{0g}^*$ to be the same for all groups and use $t \leq 3$ to estimate it.

blacks with more than a high school degree, .110 for whites with high school or less, and .143 for whites with more than high school. We then obtain $\hat{\sigma}^2_{\omega 1,g}$ by setting it to $\hat{\sigma}^2_{\omega 1,g} = \widehat{\mathrm{Var}}(\widetilde{\mathrm{wage}}^*_{i1}; g) - (\hat{\delta}^w_{\mu})^2 - \hat{\sigma}^2_{v1} - \hat{\sigma}^2_{mw}$ at each iteration of the I-I procedure, where in our base case $\hat{\sigma}^2_{mw}$ is preset to the outside estimate $.084^2$ on the basis of measurement error studies for the PSID, as discussed below. Estimates of other parameters are not very sensitive to constraining $\sigma^2_{\omega 1,g}$ to be the same for all groups and estimating $\sigma^2_{\omega 1,g}$ using $t \leq 3$ rather than 5.

To identify ρ_ω, we use a large number of moment conditions spanning a much longer time span than the three lags in our auxiliary model. The long span helps distinguish persistence due to ω_{it} from persistence due to μ_i and $v_{ij(t)}$. For workers who are continuously employed between $t-m$ and $t+n$ and who do not change jobs between t and $t+n$,

$$\mathrm{cov}(\widetilde{\mathrm{wage}}_{i,t+n} - \widetilde{\mathrm{wage}}_{it}, \widetilde{\mathrm{wage}}_{i,t-m}) = \left(\rho_\omega^{n+m} - \rho_\omega^m\right) \mathrm{Var}(\omega_{t-m}|t-m). \tag{18}$$

We approximate $\mathrm{Var}(\omega_{t-m}|t-m)$ with a constant plus a second-order polynomial in $t-m$. We compute $\widehat{\mathrm{cov}}(\widetilde{\mathrm{wage}}^*_{i,t+n} - \widetilde{\mathrm{wage}}^*_{it}, \widetilde{\mathrm{wage}}^*_{i,t-m})$ for each n, m combination satisfying $1 \leq n \leq n_{\max}$ and $1 < m \leq m_{\max}$. We estimate ρ_ω and the parameters of the polynomial approximation to $\mathrm{Var}(\omega_{t-m}|t-m)$ by weighted minimum distance using the size of the samples used to estimate $\widehat{\mathrm{cov}}(\widetilde{\mathrm{wage}}^*_{i,t+n} - \widetilde{\mathrm{wage}}^*_{it}, \widetilde{\mathrm{wage}}^*_{i,t-m})$ for each n, m combination as the weights, eliminating moments estimated using fewer than five observations. We use the average value of .908 as $\hat{\rho}$.[26] For a simpler version of the model reported in ASV (2009), we obtained .913 when we ignore (18) and rely on freely estimating ρ_ω simultaneously with the other parameters by I-I.

We impose Altonji and Williams' (2005) estimates of tenure-wage polynomial coefficients γ^w_{TEN} based on PSID data for the years 1975–2001 rather than attempting to estimate it by I-I, which would have required adding several variables to the auxiliary model.[27]

[26]In the full sample, the number of moments varies from 850 when $n_{\max}$ and $m_{\max}$ are 5 to 2429 when $n_{\max}$ and $m_{\max}$ are 9. The point estimates and approximate standard errors for $n_{\max}, m_{\max} = 5$, $n_{\max}, m_{\max} = 6$, $n_{\max}, m_{\max} = 7$, $n_{\max}, m_{\max} = 8$, and $n_{\max}, m_{\max} = 9$ are .89 (.018), .90 (.013), .91 (.009), .92 (.007), and .92 (.006), respectively. (The standard errors reported in this note account for heteroskedasticity but not for correlation among the moments, which use overlapping data. They are probably understated.) Equation (18) is an approximation because the model implies that employment transition probabilities depend on ω_t through the wage. This means that the evolution of ω_t depends on the number of periods of continuous employment: $n+m$. This should not matter much because employment transitions are not very sensitive to the wage level.

[27]The profile that we use corresponds to Table 6, Panel D, column 2 of their paper. It is $.0272563 \cdot \mathrm{TEN} - .0023283 \cdot \mathrm{TEN}^2 + .00815 \cdot \mathrm{TEN}^3/100 - .000914 \cdot \mathrm{TEN}^4/1000$. It is obtained using Altonji and Shakotko's (1987) instrumental variables approach, which treats t as exogenous and uses the within-job variation in TEN_{ijt}, TEN^2_{ijt}, TEN^3_{ijt}, and TEN^4_{ijt} to identify the effects of tenure. The implied estimates (standard errors) of the return to 2, 5, 10, and 20 years of tenure are .046 (.0064), .008 (.0011), .112 (.016), and .119 (.029), respectively. Our finding of a modest

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. Some studies have attempted to directly estimate the variances of measurement error in wages, hours, and earnings under a classical measurement error assumption (e.g., Altonji et al. (2002)). Here, we draw loosely upon studies of measurement error in the PSID and other panel data sets as well as patterns in the data to come up with estimates of the measurement error parameters. Our choices imply that measurement error accounts for 35% of $\mathrm{Var}(\Delta\mathrm{wage}^*_{it})$, 25% of $\mathrm{Var}(\Delta\mathrm{hours}^*_{it})$, and 25% of $\mathrm{Var}(\Delta\mathrm{earn}^*_{it})$. However, we also experiment with alternative choices and find that most of our results are robust. See Appendix D of the Supplemental Material for more details.[28] We abstract from measurement error in employment, which we believe is relatively unimportant, as well as in the job change indicator, which is probably more serious.

4.4. *Mechanics of Estimation*

Our chosen values of $\lambda = .05$ and $M = 20$ yield a smooth objective function that allows the use of fast gradient-based optimization algorithms with little evidence of bias.[29] Not surprisingly given the size and complexity of our models, the objective function displays multiple local optima with respect to some of the parameters. We experimented extensively with different starting values to make sure that we are finding the global optimum. We began the process with estimates obtained from probit or regression models relating the dependent variable in each equation of the structural model to the observed variables in that equation, with the fixed heterogeneity components ignored. We refined our search by using grid evaluations, paying particular attention to the set of parameters that appeared most problematic, and by experimenting with smaller versions of our models to help us find good initial guesses, and then building up to more complex versions of the models.

The fact that we are iterating on 55 parameters, the large size of the auxiliary model, and the number of simulations make computation very time-consuming

link between t and $v_{ij(t)}$ implies that Altonji and Williams' estimates are biased downward by a small amount.

[28]The assumption of normally distributed, classical measurement error runs counter to evidence that actual reports are a mixture of correct responses and responses with error. Furthermore, Bound et al. (2001) summarized 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.

[29]We use a standard quasi-Newton algorithm with line search, which can additionally handle simple bounds on the parameter values. 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.

even though we use a fast gradient-based optimization algorithm. To reduce estimation time, we exploit the highly parallelizable structure of our estimation methodology.[30]

4.5. *Bootstrap Standard Error Estimation*

We use a parametric bootstrap procedure to conduct inference. Given consistent estimates of the structural parameters, we repeatedly generate "artificial" observed data sets from the structural model, applying data availability rules that match the PSID sample. We treat each of the artificial data sets as if it was the PSID (with year effects removed) and apply our full estimation procedure, including computation of the values of $\hat{\rho}_{\omega}$, $\hat{b}_{0g}$, and $\hat{\sigma}^2_{\omega 1,g}$, to obtain estimates of the parameters of the structural model for each such data set. The standard deviations of the parameter estimates across the data sets serve as our standard error estimates.[31] The standard errors do not account for sampling error in $\hat{\gamma}^{w}_{\mathrm{TEN}}$ or uncertainty about the measurement error parameters. Standard errors of functions of model parameters, such as the impulse response functions and variance decompositions, are constructed as the standard deviation across parametric bootstrap replications. The bootstrap procedure is very computationally intensive, so we use 300 bootstrap replications.[32]

4.6. *Local Identification and Analysis of Estimation Bias*

Along with functional form restrictions and normality assumptions, exclusion restrictions play a key role in identification. First, we restrict the form of state dependence. For example, in the baseline case, employer tenure (TEN) is excluded from the EE equation, and employment duration (ED) is excluded from the job change equation (JC). Wages depend directly on only the current value and first lag of unemployment and depend directly on tenure but do not depend directly on employment duration. The lag of hours does not appear anywhere in the model. Second, the model places restrictions on the direct links among wages, employment, unemployment, job changes, and hours. Third, the model restricts fixed unobserved heterogeneity to have a two factor structure, and excludes one of the factors from the wage equation. These and

[30]Specifically, for a given value of the structural parameters, the $M = 20$ simulations required to evaluate the objective function are essentially independent and can be conducted simultaneously by 20 different processors. All programs are written in FORTRAN 90.

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

[32]We dropped 6 of 306 replications because the estimator failed to converge. The corresponding figures for the low and high education samples are 28 of 328 and 11 of 311. In the case of multinomial model in Section 7, we dropped 40 of 340 cases.

other restrictions help us distinguish between state dependence and heterogeneity and to identify the causal links among the variables in the model.

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 Monte Carlo experiments extensively to establish local identification and analyze the adequacy of our auxiliary model given the sample size and demographic structure of the available data and to check for bias. For a hypothesized vector of parameter values, we simulate data and then verify that the parameter values that maximize the likelihood function of the auxiliary model are close to the hypothesized values. Using a number of model specifications, including ones that differ somewhat from the ones presented in the paper, we informally experimented with varying parameter values to get a sense of how robust identification is to the particular values. We also used these experiments to investigate whether the objective function has flat regions near the solution, or multiple global optima.

In general, we have found that identification of most of the parameters is quite robust. However, our Monte Carlo studies also indicate that a few of the parameters are poorly determined given the sample size. We also found local optima involving alternative combinations of subsets of the parameters. Bringing in additional information through the moment conditions described above solved the most serious problems. However, some of the parameters remain sensitive to changes in the auxiliary model, and starting values must be chosen carefully. This is particularly true of the coefficients of the experience profiles in the EE, UE, and, to a lesser extent, the JC equations. In Table IV, there is also evidence of bias for some of the parameters in these equations. Overall, however, the relatively small values of the bootstrap standard errors in the tables indicate that, for the sample size and demographic structure of the PSID sample, our auxiliary model and the additional moment conditions are quite informative about most of the model parameters. Furthermore, in almost all cases, the means of the bootstrap replications are close to the point estimates, indicating that the degree of bias in the procedure is small for most of our parameters.

5. EMPIRICAL RESULTS

First, we discuss the parameter estimates for the baseline model. The baseline is the model presented in Section 2. 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 as a way to summarize how shocks affect the level and variance of wages, hours, and earnings. Finally, 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 the Full Sample*

Columns I, II, and III of Table IV report parameter estimates, the means of the parametric bootstrap estimates, and standard error estimates for the full sample. The row headings indicate the variable (or error component) and the parameter that the estimates correspond to. The estimates are grouped by equation, beginning with EE. In the case of binary variables EE, UE, and JC, column IV reports marginal effects on the probabilities that $\text{EE}_{it} = 1|E_{i,t-1} = 1$, $\text{UE}_{it} = 1|E_{i,t-1} = 0$, and $\text{JC}_{it} = 1|\text{EE}_{it} = 1$, respectively.[33]

5.1.1. *Employment Transitions and Job Changes*

The coefficients on $(t-1)$ and $(t-1)^2$ in the EE equation imply that the latent variable determining E_{it} conditional on $E_{i,t-1} = 1$ declines slowly with t until t is about 13 and then rises slowly. However, the implied change in the probability of a transition is small because the EE probability is high. The coefficient on $\min(\text{ED}_{t-1}, 9)$ is .028 (.025) and the marginal effect is .002, indicating a small positive duration dependence in the odds of remaining employed. The value of $\min(\text{ED}_{t-1}, 9)$ is rising over the first few years in the labor market, but the overall relationship between EE and t is weak. The fit of the experience profile of EE transitions is good, as we document below.[34] In Table V.A, we show that a regression of E_{it} on $\text{ED}_{i,t-1}$ conditional on $E_{i,t-1} = 1$ gives similar results in data simulated from the model and in PSID data. This indicates that the combined effect of duration dependence and unobserved heterogeneity in the model does a good job of matching the weak positive state dependence found in the data. As was noted earlier, we restricted $\gamma^{\text{EE}}_{\text{TEN}}$ to 0 because simulation experiments suggested difficulty in distinguishing the effects of employment duration and firm tenure.

The coefficient on wage^s_{it} is .071 (.118). The wage effect is not statistically significant and the implied marginal effect on EE is small.

[33]These are evaluated at the probability that $\text{EE}_{it} = 1|E_{i,t-1} = 1$, $\text{UE}_{it} = 1|E_{i,t-1} = 0$, $\text{JC}_{it} = 1|\text{EE}_{it} = 1, E_{i,t-1} = 1$, respectively, and are obtained by multiplying parameters of the latent indices by the standard normal density evaluated at the probabilities. (In the case of JC_{it}, we usually leave the conditioning on $E_{i,t-1} = 1$ implicit.) In the case of t, we take the quadratic term into account. In Appendix C of the Supplemental Material, we use simulated data to estimate average marginal effects on $\text{JC}_{it} = 1|E_{it-1} = 1$, taking into account the effects of variables on $\text{EE}_{it} = 1|E_{i,t-1} = 1$ and $\text{JC}_{it} = 1|\text{EE}_{it} = 1$.

[34]The specific point estimates of the coefficients on $t-1$, $(t-1)^2$, and $\min(\text{ED}_{t-1}, 9)$ should be taken with a grain of salt given the standard errors and the fact that the bootstrap replications provide evidence of some bias.

TABLE IV.A

BASELINE MODEL ESTIMATES[a]

		I	II	III	IV	V	VI	VII	VIII	IX	X
		Full SRC Sample				Whites With Low Education			Whites With High Education		
Equation/Variable	Parameter	Point Estimate	Bootstrap Mean	Standard Error	Marginal Effect	Point Estimate	Standard Error	Marginal Effect	Point Estimate	Standard Error	Marginal Effect
E-E Equation (6)											
(cons)	γ_0^{EE}	**1.389**	1.187	(0.243)		**1.280**	(0.239)		**1.214**	(0.436)	
$(t-1)$	γ_t^{EE}	**−0.025**	−0.008	(0.015)	0.0007	**−0.034**	(0.017)	0.0027	**−0.042**	(0.023)	0.0017
$(t-1)^2/100$	γ_{t2}^{EE}	**0.108**	0.065	(0.039)		**0.112**	(0.044)		**0.128**	(0.065)	
$\min(\text{ED}_{t-1}, 9)$	$\gamma_{\text{ED}}^{\text{EE}}$	**0.028**	0.030	(0.025)	0.0016	**0.078**	(0.023)	0.0056	**0.084**	(0.030)	0.0038
BLACK	$\gamma_{\text{BLACK}}^{\text{EE}}$	**−0.158**	−0.135	(0.115)	−0.0094						
EDUC	$\gamma_{\text{EDUC}}^{\text{EE}}$	**0.055**	0.050	(0.015)	0.0032	**0.036**	(0.024)	0.0026	**0.045**	(0.026)	0.0021
wage_t^s	γ_{ws}^{EE}	**0.071**	0.050	(0.118)	0.0042	**−0.125**	(0.145)	−0.0089	**−0.059**	(0.142)	−0.0027
μ	δ_μ^{EE}	**0.298**	0.305	(0.121)	0.0176	**0.190**	(0.093)	0.0135	**0.383**	(0.118)	0.0175
η	δ_η^{EE}	**−0.481**	−0.406	(0.103)	−0.0285	**−0.387**	(0.081)	−0.0274	**−0.083**	(0.139)	−0.0038
U-E Equation (8)											
(cons)	γ_0^{UE}	**1.597**	1.377	(0.487)		**1.468**	(0.718)		**1.598**	(0.727)	
$(t-1)$	γ_t^{UE}	**−0.124**	−0.102	(0.056)	0.0266	**−0.129**	(0.067)	0.0254	**−0.104**	(0.074)	0.0216
$(t-1)^2/100$	γ_{t2}^{UE}	**0.335**	0.289	(0.176)		**0.301**	(0.179)		**0.300**	(0.236)	
BLACK	$\gamma_{\text{BLACK}}^{\text{UE}}$	**−0.046**	−0.042	(0.229)	−0.0138						
EDUC	$\gamma_{\text{EDUC}}^{\text{UE}}$	**0.027**	0.023	(0.030)	0.0083	**0.053**	(0.056)	0.0166	**0.027**	(0.052)	0.0073
μ	δ_μ^{UE}	**0.308**	0.230	(0.176)	0.0926	**0.339**	(0.163)	0.1066	**0.120**	(0.236)	0.0325
η	δ_η^{UE}	**0.106**	0.076	(0.176)	0.0320	**−0.051**	(0.219)	−0.0159	**0.015**	(0.219)	0.0040

(*Continues*)

TABLE IV.A—*Continued*

		I	II	III	IV	V	VI	VII	VIII	IX	X
		Full SRC Sample				Whites With Low Education			Whites With High Education		
Equation/Variable	Parameter	Point Estimate	Bootstrap Mean	Standard Error	Marginal Effect	Point Estimate	Standard Error	Marginal Effect	Point Estimate	Standard Error	Marginal Effect
JC Equation (7)											
(cons)	γ_0^{JC}	**−0.498**	−0.462	(0.218)		**−0.360**	(0.234)		**−1.356**	(0.557)	
$(t-1)$	γ_t^{JC}	**−0.006**	−0.012	(0.019)	−0.0050	**−0.001**	(0.019)	−0.0047	**0.701**	(0.585)	−0.0060
$(t-1)^2/100$	γ_{t2}^{JC}	**−0.072**	−0.049	(0.050)		**−0.056**	(0.049)		**−0.318**	(0.189)	
TEN_{t-1}	γ_{TEN}^{JC}	**−0.066**	−0.057	(0.023)	−0.0128	**−0.117**	(0.028)	−0.0244	**−0.018**	(0.033)	−0.0033
BLACK	$\gamma_{\text{BLACK}}^{JC}$	**0.030**	0.011	(0.111)	0.0059						
EDUC	$\gamma_{\text{EDUC}}^{JC}$	**−0.022**	−0.021	(0.013)	−0.0042	**−0.017**	(0.019)	−0.0036	**−0.020**	(0.028)	−0.0037
υ_{t-1}	$\delta_{\upsilon-1}^{JC}$	**−0.833**	−0.787	(0.154)	−0.0503	**−0.801**	(0.226)	−0.0501	**−0.573**	(0.236)	−0.0324
υ_t'	δ_{υ}^{JC}	**0.496**	0.485	(0.132)	0.0336	**0.269**	(0.187)	0.0186	**0.804**	(0.201)	0.0540
μ	δ_{μ}^{JC}	**−0.067**	−0.116	(0.127)	−0.0130	**−0.004**	(0.088)	−0.0008	**−0.521**	(0.223)	−0.0954
η	δ_{η}^{JC}	**0.539**	0.526	(0.110)	0.1042	**0.437**	(0.119)	0.0913	**0.399**	(0.237)	0.0730

[a]The table presents estimation results for our base model, estimated on the full SRC sample and two subsamples: SRC Whites with Education ≤ 12, and SRC Whites with Education > 12. Estimates were obtained by Indirect Inference, unless indicated otherwise. Marginal effects are evaluated at the mean of EE, UE, and JC, respectively. The marginal effects of potential experience accounts for the quadratic term. The marginal effects of υ_{t-1} and υ_t' are the effect of a one-standard-deviation-change based on the standard deviations for the particular sample. Parametric bootstrap standard errors are in parentheses. Bootstraps are based on 300 replications. As explained in Footnote 24 in the paper, the hours equation includes a second constant that has no effect on earn_{it}^*. The point estimates of that constant are .052, .049, and −.199, for the Full, Low Education, and High Education samples, respectively.

TABLE IV.B

BASELINE MODEL ESTIMATES[a]

		I	II	III	V	VI	VIII	IX
		Full SRC Sample			Whites With Low Education		Whites With High Education	
Equation/Variable	Parameter	Point Estimate	Bootstrap Mean	Standard Error	Point Estimate	Standard Error	Point Estimate	Standard Error
Wage Equation (Eqs. (1)–(5))								
Constant		**0.001**	0.009	(0.055)	**−0.003**	(0.115)	**−0.031**	(0.131)
$(t-1)$	γ_t^w	**0.064**	0.065	(0.005)	**0.065**	(0.007)	**0.061**	(0.007)
$(t-1)^2/10$	γ_{t2}^w	**−0.021**	−0.021	(0.003)	**−0.022**	(0.003)	**−0.019**	(0.004)
$(t-1)^3/1000$	γ_{t3}^w	**0.022**	0.023	(0.004)	**0.027**	(0.005)	**0.018**	(0.006)
BLACK	$\gamma_{\mathrm{BLACK}}^w$	**−0.224**	−0.223	(0.029)				
EDUC	γ_{EDUC}^w	**0.105**	0.105	(0.004)	**0.091**	(0.010)	**0.117**	(0.008)
Tenure polynomial		**yes**			**yes**		**yes**	
μ	δ_μ^w	**0.081**	0.096	(0.035)	**0.141**	(0.035)	**0.123**	(0.050)
v_{t-1}	ρ_v	**0.691**	0.679	(0.049)	**0.611**	(0.068)	**0.719**	(0.075)
ε^v	σ_v	**0.276**	0.282	(0.009)	**0.277**	(0.012)	**0.292**	(0.017)
ε_1^v	σ_{v1}	**0.165**	0.164	(0.016)	**0.172**	(0.019)	**0.089**	(0.008)
ω_{t-1}	ρ_ω[b]	**0.908**	0.910	(0.025)	**0.865**	(0.060)	**0.921**	(0.030)
$1-E_t$	γ_{1-Et}^ω	**−0.134**	−0.137	(0.013)	**−0.170**	(0.017)	**−0.092**	(0.020)
$1-E_{t-1}$	γ_{1-Et-1}^ω	**0.049**	0.053	(0.017)	**0.072**	(0.030)	**0.015**	(0.029)
ε^ω	σ_ω	**0.089**	0.085	(0.005)	**0.066**	(0.009)	**0.097**	(0.012)
ε_1^ω (Black, Low Education)	$\sigma_{\omega 1}$[b]	**0.160**	0.245	(0.055)				
ε_1^ω (Black, High Education)	$\sigma_{\omega 1}$[b]	**0.241**	0.274	(0.050)				
ε_1^ω (White, Low Education)	$\sigma_{\omega 1}$[b]	**0.263**	0.294	(0.023)	**0.212**	(0.028)		
ε_1^ω (White, High Education)	$\sigma_{\omega 1}$[b]	**0.319**	0.318	(0.018)			**0.342**	(0.027)

(*Continues*)

TABLE IV.B—*Continued*

		I	II	III	V	VI	VIII	IX
		Full SRC Sample			Whites With Low Education		Whites With High Education	
Equation/Variable	Parameter	Point Estimate	Bootstrap Mean	Standard Error	Point Estimate	Standard Error	Point Estimate	Standard Error
Hours Equation (9)								
constant	γ_0^h	**−0.454**	−0.450	(0.015)	**−0.451**	(0.039)	**−0.220**	(0.029)
$(t-1)$	γ_t^h	**0.009**	0.009	(0.002)	**0.008**	(0.004)	**0.008**	(0.003)
$(t-1)^2/10$	γ_{t2}^h	**−0.003**	−0.003	(0.001)	**−0.002**	(0.002)	**−0.003**	(0.002)
$(t-1)^3/1000$	γ_{t3}^h	**0.002**	0.002	(0.002)	**0.000**	(0.003)	**0.001**	(0.003)
BLACK	$\gamma_{\mathrm{BLACK}}^h$	**−0.054**	−0.055	(0.015)				
EDUC	γ_{EDUC}^h	**0.011**	0.011	(0.002)	**0.017**	(0.004)	**0.012**	(0.004)
E_t	γ_E^h	**0.430**	0.431	(0.011)	**0.434**	(0.016)	**0.451**	(0.014)
ε^ξ	σ_ξ	**0.162**	0.175	(0.013)	**0.135**	(0.028)	**0.185**	(0.014)
$\mathrm{wage}_t^{\mathrm{lat}}$	γ_w^h	**−0.084**	−0.087	(0.016)	**−0.189**	(0.028)	**−0.049**	(0.019)
μ	δ_μ^h	**0.098**	0.098	(0.018)	**0.127**	(0.023)	**0.052**	(0.021)
η	δ_η^h	**−0.012**	−0.002	(0.024)	**−0.041**	(0.023)	**0.035**	(0.042)
ε^h	σ_h	**0.141**	0.140	(0.003)	**0.169**	(0.004)	**0.103**	(0.004)
Earnings Equation (10)								
constant	γ_0^e	**−0.018**	−0.018	(0.001)	**−0.010**	(0.012)	**0.010**	(0.002)
$\mathrm{wage}_t^{\mathrm{lat}}$	$\gamma_{\mathrm{w}}^{e\,\mathrm{c}}$	**1.000**			**1.000**		**1.000**	
hours_t	$\gamma_h^{e\,\mathrm{c}}$	**1.000**			**1.000**		**1.000**	
e_t	ρ_e	**0.624**	0.624	(0.009)	**0.579**	(0.015)	**0.673**	(0.013)
ε^e	σ_e	**0.169**	0.169	(0.002)	**0.179**	(0.003)	**0.152**	(0.003)

[a]The notes are the same as for Table IV.A.

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

[c]Imposed.

Moving to the UE equation, the estimated t profile implies that the exit probability declines with experience and then increases, but the standard errors are large. As we document in Section 5.2.1, the model predictions of the probability that $\mathrm{UE} = 1$ are within the 95% confidence intervals based on the PSID data for most values of t, but the PSID data are noisy. We experimented with models that included $\mathrm{UD}_{i,t-1}$ but had difficulty estimating the duration coefficient, perhaps because the overall number of unemployment spells is small and relatively few individuals were unemployed for two or more surveys in a row. (Most work on duration dependence in unemployment spells uses weekly, monthly, or quarterly data.) Simulations in Table V.A show that the equation without unemployment duration matches closely the negative link between UE and $\mathrm{UD}_{i,t-1}$ found in the PSID, presumably because of the important role played by permanent heterogeneity.

In the JC equation, the latent variable for $\mathrm{JC}_{it}|\mathrm{EE}_{it} = 1$ declines with t, and is strongly decreasing in job tenure. The coefficient on $\mathrm{TEN}_{i,t-1}$ is $-.066$ (.023), indicating that 10 years of seniority shift the index determining JC_{it} by .66 standard deviations of the job change shock $\varepsilon_{it}^{\mathrm{JC}}$. The marginal effect of an extra year of tenure on $\mathrm{JC}_{it} = 1$ is $-.013$. It is noteworthy that we obtain a large negative effect of tenure on JC even after accounting for unobserved person-specific heterogeneity (μ and η) and for job match heterogeneity.

The job match components $v_{ij(t-1)}$ and $v'_{ij'(t)}$ play an important role in job mobility without unemployment, and they have signs and relative magnitudes that are consistent with the theoretical discussion above. The coefficient on $v_{ij(t-1)}$ is $-.833$ (.154). To get a sense of the magnitude, note that the standard deviation of $v_{ij(t-1)}$ is .310. Consequently, a one-standard-deviation increase in $v_{ij(t-1)}$ lowers the JC index by $-.26$. Since the coefficient on TEN is $-.066$, this is roughly equivalent to the effect of 4 years of seniority. The average marginal effect of a one-standard-deviation shift is $-.050$. The current value $v'_{ij'(t)}$ has a coefficient of .496 (.132).[35] A one-standard-deviation shock to $v'_{ij'(t)}$ raises the job change probability by .034.

The coefficient on BLACK is $-.158$ (.115) in the equation for EE. The average marginal effect is $-.009$. BLACK has a small, negative but imprecisely estimated effect in the equation for UE. EDUC has a small positive effect in both the EE and UE equations. The effects of BLACK and EDUC on JC are small and insignificant. The coefficient on the "ability" factor μ is .298 (.121) in the EE equation, .308 (.176) in the UE equation, and $-.067$ (.127) in the JC equation. These results are sensible in light of the fact that μ has a positive sign in both the wage and hours equations. The corresponding marginal effects of a one-standard-deviation shift are .018, .093, and $-.013$. These are the partial effects of the heterogeneity components in a given period holding spell duration constant.

[35]The total effect of $v_{ij(t-1)}$ on the JC probability is -0.027, which is smaller than the partial effect because a unit shift in $v_{ij(t-1)}$ shifts the mean of $v'_{ij'(t)}$ by .691.

The mobility/hours preference component η enters the EE, UE, and JC indices with coefficients of $-.481$ (.103), .106 (.176), and .539 (.110), respectively. The results indicate that η lowers the probability of remaining employed and raises the probabilities of transiting out of unemployment (not significant) and of moving from job to job without unemployment. The marginal effects of a one-standard-deviation shift in η on the mean probabilities are $-.029$ for EE, .032 for UE, and .104 for $\mathrm{JC}|\mathrm{EE}_t = 1$. It has essentially a zero coefficient in the hours equation.

5.1.2. *The Wage Model*

We begin with the parameters of (3), the equation for the autoregressive component ω_{it}. The estimated standard deviation of the initial condition for ω_{i1} ranges from .160 for less educated blacks to .319 for highly educated whites. The autoregressive coefficient $\hat{\rho}_\omega$ is .908 (.025), which implies considerable persistence but is well below unity. The shocks ε^ω_{it} have a standard deviation of .089 (.005). This value strikes us as large given that we separately account for the effects of job-specific error components. The only other study we know that allows for a persistent general wage component, a job-specific error term, and endogenous mobility is Low, Meghir, and Pistaferri (2010). They obtained a value of .104 but set $\rho_\omega = 1$.[36] In the PSID, the standard deviation of wage changes for stayers is .128 after adjusting for measurement error.

The coefficients of $-.134$ (.013) on $(1 - E_{it})$ and .049 (.017) on $(1 - E_{i,t-1})$ imply that being unemployed at the survey date has a large effect on the mean of the wage that persists for some time, even when the value of lost tenure is held constant. As will become clear from the impulse response functions, unemployment also leads to a loss of tenure as well as to a reduction, on average, in the value of the job match component, which implies further reductions in wages.

The estimate for coefficient δ^w_μ on μ is .081 (.035). The direct contribution of unobserved permanent heterogeneity to the variance of wages is relatively small once we account for both ω_{it} and job match heterogeneity. However, the estimate of δ^w_μ is somewhat sensitive to model specification.

The parameters of the job match component $v_{ij(t)}$ are quite interesting. The initial condition $v_{ij(1)}$ has a standard deviation of .165 (.016). The autoregression parameter ρ_v is .691 (.049) and the value of $\hat{\sigma}_v$ is large: .276 (.009). As we have already noted, the substantial persistence of $v_{ij(t)}$ across jobs suggests that wage offers are based in part on salary history, that demand shocks may reflect narrow occupation, industry, and region and thus may not be entirely job-specific, and/or that the search network available to workers depends on job quality. As we shall see below, the contribution of the job-specific component to the variance of wages and earnings is substantial.

[36]They do not include the $(1 - E_{it})$ terms in their specification.

FIGURE 1.—Decomposing the experience profile of wages. Baseline model, full SRC sample. The figure displays the model's decomposition of wage growth over a career (or the experience profile of log wages) into the contributions of job shopping (the mean value of the job-specific wage component ν), the accumulation of tenure (the contribution of the mean value of tenure on the wage experience profile), and the accumulation of general human capital.

As we discuss in Appendix A, one can use the model to decompose $\mathbf{E}(\text{wage}_{it}|t)$, the experience profile of wages, into the contributions of general human capital $[t\gamma_t^w + t^2\gamma_{t2}^w + t^3\gamma_{t3}^w]$, change in the general productivity component ω_{it} due to unemployment shocks, $\mathbf{E}(\omega_{it}|t)$, gains from job mobility $\mathbf{E}(v_{ij(t)}|t)$, and accumulated job seniority $\mathbf{E}(P(\text{TEN}_{it})\gamma_{\text{TEN}}^w|t)$. Figure 1 shows the components and thus addresses the fundamental question of what accounts for wage growth over a career. We exclude $\mathbf{E}(\omega_{it}|t)$ from the graph to reduce clutter. It declines by $-.016$ over the first 10 years and $-.020$ over the first 30 years, primarily because of unemployment shocks. Most of the return to potential experience is due to general skill accumulation. Job shopping and the accumulation of tenure account for increases of .068 and .064, respectively, of the .513 increase in the mean of the log wage over the first 10 years. They account for .128 and .113 of the total increase of .830 over the first 30 years. Using Social Security records for quarterly earnings (rather than hourly wage rates), Topel and Ward (1992) found that job mobility accounts for 1/3 of wage growth during the first 10 years in the labor market, which is much larger than what we find. We suspect their estimates are overstated by the school-to-work transition and growth across jobs in hours worked early in a career, while ours are understated because we miss some job changes and do not use the first three years of wages in the component $\mathcal{L}_2$ of the auxiliary model.[37]

[37]Using the PSID, Buchinsky et al. (2010) estimated a simultaneous model of employment, job mobility, and wage rates that incorporates tenure effects, general experience, and job-specific error components. They found a large effect of human capital accumulation and returns to seniority that are more than double the values from Altonji and Williams (2005) that we impose,

5.1.3. *Hours and Earnings*

In the *hours* equation, the effect of E_{it}, $\hat{\gamma}_E^h$, is .430 (.011). The large value indicates that unemployment at the survey date is associated with relatively long completed spells of nonemployment. As we noted above, short spells will tend to be missed by E_{it} given that it is a point-in-time measure at annual frequencies. They will be captured by the hours component ε_{it}^h. The small negative wage elasticity of $-.084$ (.016) is consistent with a large literature that finds that the hours of male household heads are not very responsive to wages. The coefficients on μ and η are .098 (.018) and $-.012$ (.024), respectively, suggesting only a modest role for individual heterogeneity (net of EDUC and BLACK) in annual hours in any given year. However, permanent heterogeneity turns out to be quite important over the lifetime. The importance of μ relative to η varies across the specifications. The standard deviation of ε_{it}^h is .141, indicating substantial year-to-year variation in hours even when the job-specific component ξ does not change. The standard deviation of ξ is large—.162 (.013).

Turning to earnings, recall that the coefficients γ_w^e and γ_h^e are constrained to equal 1. The earnings component e_{it} has an autoregression coefficient of .624 (.009) and the standard deviation of the shock ε_{it}^e is .169 (.002).

Does persistence in wages, earnings, and hours stem primarily from permanent heterogeneity or from dynamics in the model and persistence in ω_{it}, $v_{ij(t)}$, and $\xi_{ij(t)}$ as well as e_{it}? To explore this question, we used simulated data from the model to regress the wage on the first lag of the wage with and without controls for EDUC_i, BLACK_i, μ_i, η_i, and ω_{i1}. We repeated the procedure using the fifth lag and the tenth lag of the wage. We also estimated similar sets of autoregressions for earnings and hours. The size of the decline in the lag coefficient when the controls are added indicates that fixed heterogeneity is a very important source of persistence in wages, earnings, and hours, but not the main source (not reported).

but did not present estimates of the gains from job mobility. Given total wage growth in the PSID, the sum of their estimates of the general human and tenure profile seems to imply zero or negative gains from mobility on average over a 30-year career. Bagger et al. (2011) did not allow for a direct effect of seniority on wages such as would arise from shared investment in firm-specific capital. They interpreted an indirect effect that arises through the response of firms to outside offers as a return to tenure. They attributed average growth of wages within the firm and growth of wages across firms to job search and to general human capital. Using Danish matched employer/employee data, they found that general human capital accumulation raises the productivity of the less educated group by about .14 over the first 10 years, but lowers productivity by about $-.09$ over the next 20 years. It is much more important for highly educated workers. The within-job and between-job wage growth that reflect job search and competition are both important contributors to wage growth for those who remain employed across periods. The between-job component is the more important of the two. However, they did not provide estimates of the overall contribution of moving to better jobs that we can compare to our estimates, because they focused on job changes without unemployment and did not take account of the fact that workers lose some of the gains from prior search when they suffer a layoff. Rubinstein and Weiss (2006) surveyed the literature on the determinants of wage growth over a career.

5.2. *Evaluating the Fit of the Model*

We use the estimated model to simulate careers for 271,200 individuals (100 times the size of our PSID sample). We construct the simulation so that, for each simulated career, education, race, and the potential experience values for which data are available match that of a corresponding PSID case. The simulated variables incorporate measurement error. To examine the fit of the model, we first compare the experience profiles of the means and standard deviations of the key variables implied by the model against corresponding values from the PSID. We then turn to a comparison of regression relationships among key variables that are implied by the model with those of the corresponding PSID estimates.

5.2.1. *Predicted and Actual Means and Standard Deviations of Key Variables, by Potential Experience*

Figure 2 compares means and standard deviations of key variables in the PSID against the corresponding model predictions. Panels (a), (b), and (c) display the 95% confidence interval estimates of the standard deviations of wage^*_{it}, hours^*_{it}, and earn^*_{it} based directly on the PSID sample to the point estimates from the model.[38] In all cases except the wage when $t = 5$ and hours when $t = 35$, the model predictions lie within the confidence intervals, although some values are close to the boundary. Across experience levels, the predicted value of $\text{SD}(\text{wage}^*_{it})$ is .50, while the actual value is .49. The model slightly overpredicts the standard deviation early in a career and understates the increase a bit.

The sample value for $\text{SD}(\text{hours}^*_t)$ is .28—close to the model value of .27. The model implies that $\text{SD}(\text{hours}^*_t)$ varies little with t and understates the increase between $t = 30$ and 35. This might reflect the effects of partial retirement not captured by the experience profiles in the model, but it also may be due to sampling error in the PSID estimates.

Panels (d) and (f) of Figure 2 compare the PSID values and the model predictions for the mean of E_t and for the mean of JC_t conditional on $E_{t-1} = 1$. The overall mean for E_t is .966 in the data and .971 based on the model. The model overstates employment when $t = 5$ by about .02, which is statistically significant. Neither the data nor the model predictions show much movement in E_t with experience. Overall, the mean of JC_t predicted by the model matches the data very closely relative to sampling error. Panel (e) reports the sample means and simulated means of EE transitions, which match reasonably well. Panel (g) shows corresponding figures for exits from unemployment (UE). The actual and simulated means of UE are .774 and .803, so the model matches

[38] For each value of t shown in the figure, the results are based on $t - 1$, t, and $t + 1$. The confidence interval estimates are based on the normal approximation using robust standard errors clustered at the individual level. We display point estimates for the model predictions rather than confidence interval estimates because the latter are very narrow in almost all cases.

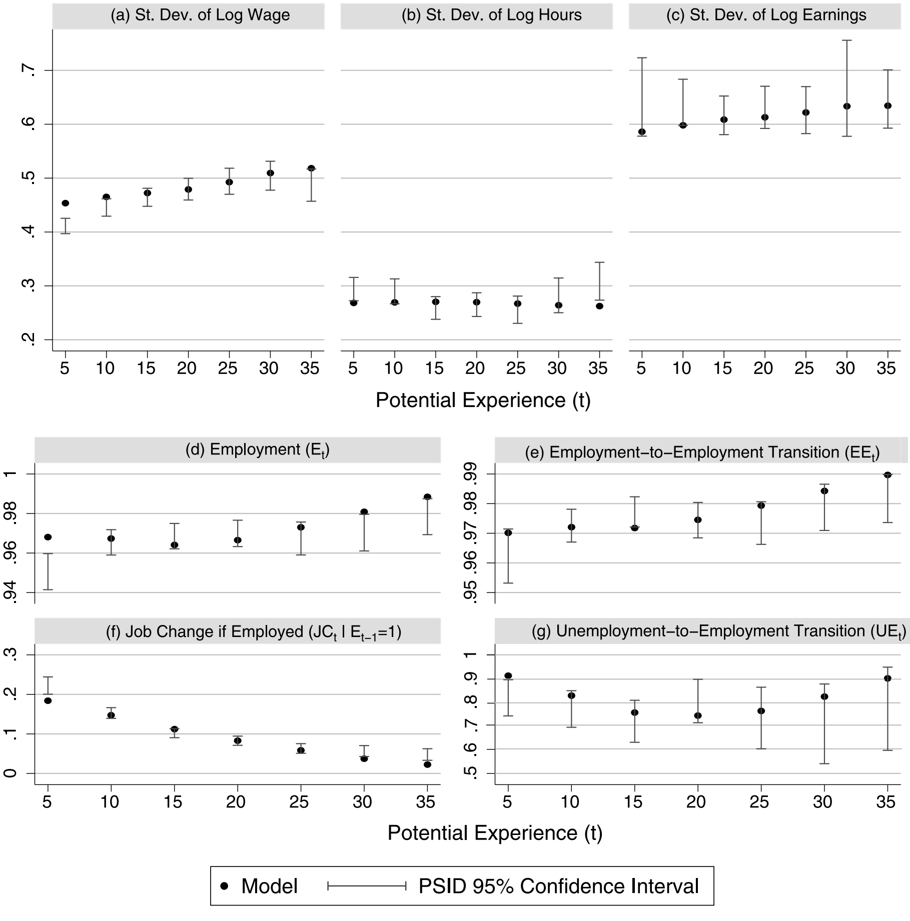


FIGURE 2.—Comparisons of PSID Data and Model Predictions by Potential Experience (t). The figure compares standard deviations and means of key variables in our baseline PSID sample against the corresponding model predictions. All panels display the 95 percent confidence interval estimates of sample statistics based on the PSID to point estimates from the model. In all panels, each statistic corresponding to potential experience t was actually calculated using potential experience $t-1$, t, and $t+1$. Panels (a), (b), and (c) display the standard deviations of wage^*_{it}, hours^*_{it}, and earn^*_{it}. The units in the vertical axis of those three panels are thus standard deviations of the log. All other panels in the figure display means of the main variable specified in the corresponding panel heading. For example, panel (e) displays the mean of employment, E_t, conditional on $E_{t-1}=1$. (*Continues*)

these transitions slightly less well. However, as the figure shows, there is a lot of noise in the sample means for particular experience values.[39]

[39]The confidence interval estimates for the UE probability in panel (g) are the exact confidence intervals for the binomial distribution. All other confidence intervals in Figure 2 are based on the normal approximation as described previously.

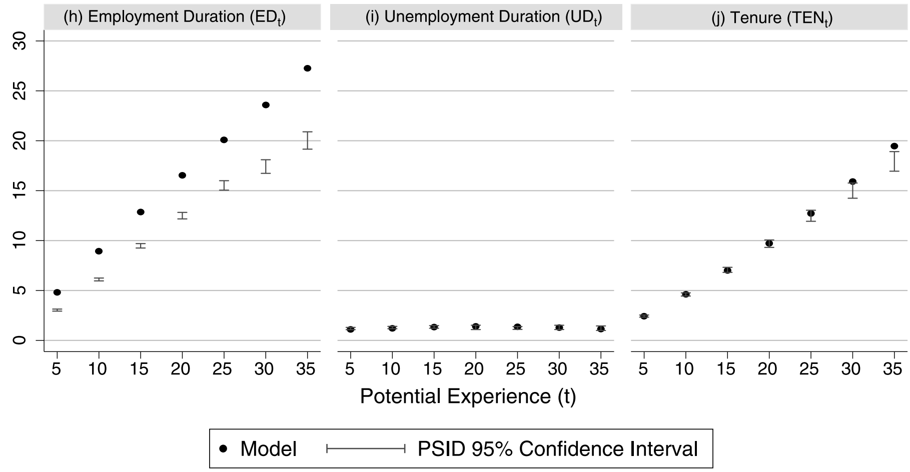


FIGURE 2.—*Continued*.

Panels (h), (i), and (j) examine the behavior of the mean of ED, UD, and TEN. The fits for UD and TEN are reasonably close, although the model slightly overpredicts tenure late in a career. The model overpredicts ED by a substantial amount. This is probably attributable to our use of TEN as the initial value for ED when an individual first enters the PSID sample.

5.2.2. *Comparison of Regression Relationships Among Key Variables*

Tables V.A and V.B report a series of descriptive regressions. Coefficients in bold are regression coefficients based on simulated data. Each of these coefficients is followed underneath by a corresponding regression coefficient based on the PSID data (in italics). Robust panel standard errors for the PSID estimates are in parentheses.[40] Column I of Table V.A reports regressions of E_t on BLACK, EDUC, $(t-1)/10$, $(t-1)^2/100$, 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 differences in the experience profiles. The coefficient on ED_{t-1} is .0027 in the simulation and .0022 (.0002) in the PSID, a fairly close correspondence.

Column II reports results for a version of the UE equation. Although the differences in the coefficients on the experience polynomial appear substantial, the model fits the experience profile reasonably well, as previously seen. The

[40]As in the previous subsection, the point estimates for the simulated data are based on a sample 100 times as large as the PSID, with the same demographic structure. The PSID standard errors provide a rough guide to whether the coefficients based on the simulated data are statistically different from the PSID regression coefficients.

TABLE V.A

REGRESSIONS COMPARING DATA SIMULATED FROM ESTIMATED BASELINE MODEL AND PSID: EMPLOYMENT AND JOB CHANGES[a]

Dependent Variable:	I E_t[b]	II E_t[c]	III JC_t[d]
Independent Variable:			
Black	**−0.0076**	**−0.0139**	**−0.0013**
	−0.0117	*−0.0190*	*−0.0131*
	(0.0039)	(0.0477)	(0.0079)
Education	**0.0018**	**0.0058**	**−0.0011**
	0.0028	*0.0197*	*−0.0034*
	(0.0004)	(0.0070)	(0.0008)
$(t-1)/10$	**−0.0241**	**−0.2760**	**−0.0621**
	−0.0017	*0.0131*	*−0.0946*
	(0.0043)	(0.0597)	(0.0083)
$(t-1)^2/100$	**0.0029**	**0.0722**	**0.0140**
	−0.0013	*−0.0088*	*0.0221*
	(0.0010)	(0.0158)	(0.0020)
ED_{t-1}	**0.0027**		
	0.0022		
	(0.0002)		
UD_{t-1}		**−0.0373**	
		−0.0364	
		(0.0232)	
$TEN_{t-1}/10$			**−0.0864**
			−0.0809
			(0.0030)
Observations	**2,620,289**	**79,711**	**2,558,223**
	25,749	*793*	*24,712*
R^2	**0.01**	**0.03**	**0.07**
	0.01	*0.02*	*0.06*
RMSE	**0.15**	**0.40**	**0.29**
	0.15	*0.40*	*0.29*

[a]The table presents least-squares regressions comparing data simulated from our estimated baseline model and PSID data. Estimates on simulated data are in bold, estimates on PSID data are in italics, standard errors (for the PSID estimates) are in parentheses. The regressions on simulated data are based on a simulated sample that is 100 times as large as the PSID sample, but has the same demographic structure (by potential experience, race, and education) as the PSID sample.

[b]Sample restricted to observations where $E_{t-1} = 1$.

[c]Sample restricted to observations where $E_{t-1} = 0$.

[d]Sample restricted to observations where $E_t = 1$ and $E_{t-1} = 1$.

model matches well the degree of persistence in unemployment spells. The coefficients of the JC equation in Column III match fairly closely.

Table V.B examines the dynamics of wage_t^*, hours_t^*, and earn_t^*. In Column I, the sums of the coefficients on the two lags of wage_t^* are very close, although

TABLE V.B

REGRESSIONS COMPARING DATA SIMULATED FROM ESTIMATED BASELINE MODEL AND PSID: WAGES, HOURS, EARNINGS[a]

Dependent Variable:	I $wage^{*b}_t$	II $wage^{*b}_t$	III $hours^*_t$	IV $earn^*_t$	V $earn^*_t$
Independent Variable:					
Black	**−0.0164**	**−0.2283**	**−0.0235**	**−0.0527**	
	−0.0244	*−0.2407*	*−0.0291*	*−0.0580*	
	(0.0053)	(0.0095)	(0.0058)	(0.0089)	
Education	**0.0078**	**0.1071**	**0.0062**	**0.0218**	
	0.0129	*0.1032*	*0.0032*	*0.0230*	
	(0.0006)	(0.0010)	(0.0007)	(0.0011)	
$t/10$	**−0.0155**	**0.6326**	**0.0466**	**0.0687**	
	−0.0352	*0.6585*	*−0.0282*	*−0.0485*	
	(0.0193)	(0.0283)	(0.0212)	(0.0325)	
$t^2/100$	**−0.0010**	**−0.1868**	**−0.0162**	**−0.0266**	
	0.0099	*−0.2376*	*0.0186*	*0.0335*	
	(0.0098)	(0.0155)	(0.0108)	(0.0166)	
$t^3/1000$	**0.0007**	**0.0193**	**0.0011**	**0.0032**	
	−0.0008	*0.0280*	*−0.0040*	*−0.0059*	
	(0.0015)	(0.0025)	(0.0017)	(0.0025)	
$TEN_t/10$		**0.3704**			
		0.4055			
		(0.0184)			
$TEN^2_t/100$		**−0.2016**			
		−0.1722			
		(0.0146)			
$TEN^3_t/1000$		**0.0300**			
		0.0280			
		(0.0031)			
$wage^*_t$			**−0.0247**		**0.9740**
			−0.0003		*0.9458*
			(0.0035)		(0.0036)
$wage^*_{t-1}$	**0.7126**				
	0.6497				
	(0.0064)				
$wage^*_{t-2}$	**0.2116**				
	0.2662				
	(0.0064)				
JC_t	**−0.0159**				
	−0.0107				
	(0.0043)				
$hours^*_t$					**0.8552**
					0.7648
					(0.0073)

(*Continues*)

TABLE V.B—*Continued*

Dependent Variable:	I wage_t^{*b}	II wage_t^{*b}	III hours_t^*	IV earn_t^*	V earn_t^*
hours_{t-1}^*			**0.3377**		
			0.3635		
			(0.0066)		
hours_{t-2}^*			**0.2736**		
			0.1633		
			(0.0063)		
earn_{t-1}^*				**0.5658**	
				0.7159	
				(0.0069)	
earn_{t-2}^*				**0.2563**	
				0.1248	
				(0.0068)	
Observations	**2,158,813**	**3,041,429**	**2,262,355**	**2,329,200**	**3,041,429**
	21,783	*29,000*	*22,687*	*23,292*	*30,265*
R^2	**0.88**	**0.36**	**0.31**	**0.70**	**0.79**
	0.85	*0.40*	*0.23*	*0.73*	*0.74*
RMSE	**0.17**	**0.40**	**0.21**	**0.34**	**0.28**
	0.18	*0.38*	*0.21*	*0.32*	*0.31*

[a]The table presents least-squares regressions comparing data simulated from our estimated baseline model and PSID data. Estimates on simulated data are in bold, estimates on PSID data are in italics, standard errors (for the PSID estimates) are in parentheses. The regressions on simulated data are based on a simulated sample that is 100 times as large as the PSID sample, but has the same demographic structure (by potential experience, race, and education) as the PSID sample.

[b]All observations in sample must satisfy $E_t = 1$.

the relative sizes of the coefficients on the first and second lags differ a little. The coefficient on JC_t is small and negative in both the simulated and PSID data.

Column III examines hours. The coefficients match reasonably closely, although the sum of the coefficients on the lags of hours is 0.61 in the simulated data but only .53 in the PSID data, and the discrepancy between the cubic polynomials is between .048 and .054 for values of t greater than 11. The wage coefficient is essentially 0 in the actual data and $-.025$ in the simulated data—a close correspondence.

Finally, Columns IV and V report earnings regressions. Note that all of the dynamics in earnings stem from dynamics in the wage, hours, and the autoregressive earnings component e_{it}. In Column IV, the sum of the coefficients on the lagged earnings $\text{earn}_{i,t-1}^*$ and $\text{earn}_{i,t-2}^*$ is .82 in simulated data and .84 in the PSID data, so that the model understates the persistence of earnings by a small amount. There are also some differences between the data and the model in the coefficients on wage_t^* and hours_t^* (Column V).

Overall, we view the match between the model and the data as good.

5.3. *Mean and Variance Impulse Response Functions*

Figure 3 reports impulse responses to shocks that occur when $t = 10$. The point estimates displayed are constructed as follows. First, using our model, we simulate a large number of individuals through $t = 9$. Then we impose the shock indicated in the figures on all individuals in period 10. After that, we continue the simulation in accordance with the model. The different panels in the figure show the mean paths of earnings, wages, and hours relative to the base case. The base case represents the mean of the simulated paths in the absence of the specified intervention in period 10.[41]

Since wages and hours are reflected in earnings with coefficients of 1, we focus the discussion on earnings, shown in panel (a), to save space. Panels (b) and (c) show the response of wages and hours. The diamond line in panel (a) reports the response of the mean of earn_{it} to a one-standard-deviation positive shock to $\varepsilon_{it}^{\omega}$, the error term in the autoregressive component of wages. Earnings rise by about .08 and the effect slowly decays, governed by the value .908 for ρ_{ω}. The pattern for earnings closely mirrors the response of wages (panel (b)) because the coefficient on the wage is 1 and the wage elasticity of hours is only $-.08$.

The line with circles shows the effect of becoming unemployed when $t = 10$. The pattern is very interesting. The log of earnings drops by about $-.6$, recovers by about two thirds after one year, and then slowly returns to the base case. The initial drop is the combination of a drop of about $-.4$ in log hours (panel (c)) and a drop of about $-.2$ in the wage (panel (b)). Hours recover almost completely after one period. The wage increases by about .02 in the first year and continues to recover slowly after that.

The drop in wages is due to three main factors. First, the distributed lag coefficients on unemployment in the wage equation and $\hat{\rho}_{\omega}$ indicate that unemployment reduces ω_{it} by $-.134$ (.013) followed by an increase a year later of .049 (.017) plus $.134 * (1 - .908)$ if the person leaves unemployment. After that, the response of ω_{it} to unemployment is governed by $\hat{\rho}_{\omega}$. Second, the loss of tenure lowers the wage by an average of .064 relative to the baseline average for persons at $t = 10$. Third, since there is no selectivity in the job change induced by the unemployment spell, on average, workers suffer a decline in $\upsilon_{ij(t)}$ equal to $(1 - \rho_{\upsilon})\mathbf{E}(\upsilon_{ij(t)} | t = 10)$, or .021. On average, endogenous mobility following the unemployment spell leads $\upsilon_{ij(t)}$ to move back up toward the base case mean for a given value of t.

The pattern of a long-lasting impact of unemployment on earnings is broadly consistent with a number of previous studies, including Jacobson, LaLonde, and Sullivan (1993), who used establishment earnings records. Using the PSID

[41] 1.984 standard error bands were obtained by computing impulse responses using each of the 300 values of the model parameters obtained by parametric bootstrap. The bands are quite narrow, so we omit them to avoid cluttering the figures. They are available upon request.

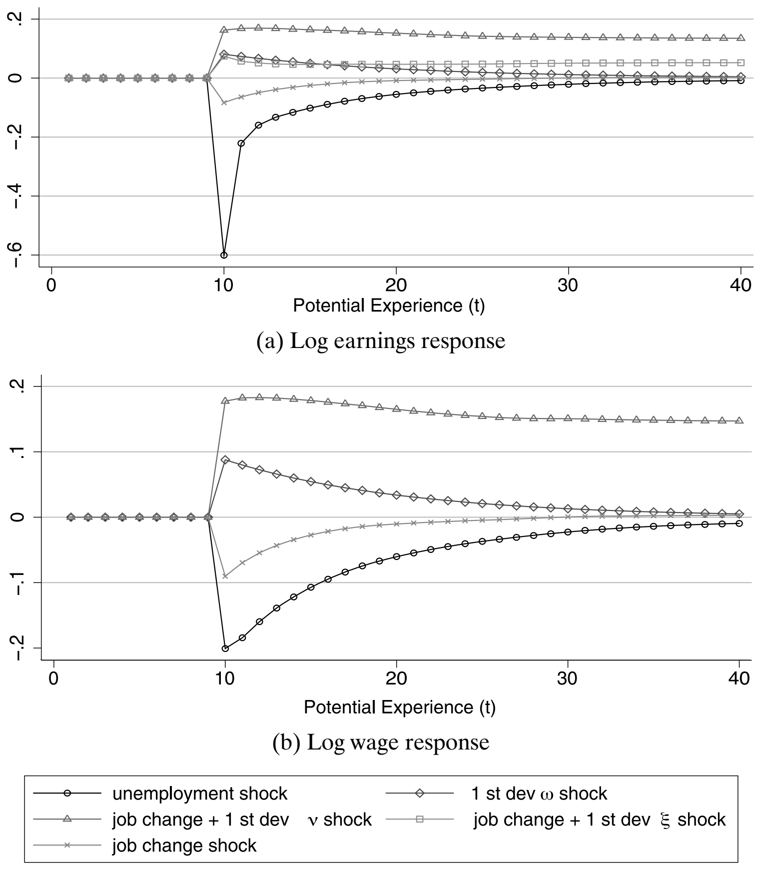
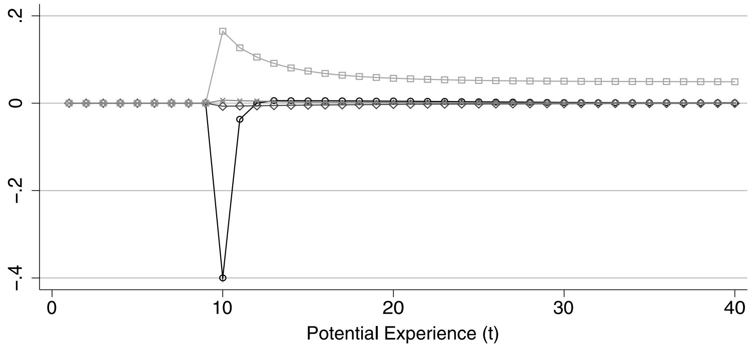


FIGURE 3.—Mean Response of Key Variables to Various Shocks at $t = 10$. The figure displays the response of the mean of log earnings, log wage, and log hours to various shocks that are imposed when potential experience $t = 10$. The shocks are an unemployment shock, a job change shock, a one- standard-deviation shock to the autoregressive component of wages, a job change shock accompanied by a one- standard-deviation shock to the job-specific wage component, and a job change shock accompanied by a one- standard-deviation shock to the job-specific hours component. To construct the point estimates, we first use the model to simulate a large number of individuals through $t = 9$. We then impose the shock indicated in the figures in period 10 on all individuals. After that, we continue the simulation in accordance with the model. The panels in the figure show the mean paths of log earnings, log wages, and log hours relative to the base case. The base case represents the mean of the simulated paths in the absence of the specified intervention in period 10. (*Continues*)

and a fixed effects strategy, Stevens (1997) found a 30% drop in earnings and a 14% drop in wages in the year of a layoff. Earnings recover substantially in the first year, but wages recover very slowly. Her estimate of the initial earnings loss is smaller than ours, perhaps because those who are laid off do not necessarily



(c) Log hours response

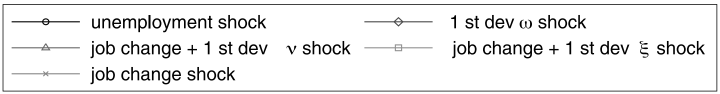


FIGURE 3.—*Continued*.

become unemployed, and those who are unemployed at the survey date tend to be in a long spell. The model permits us to examine effects that operate through wages and hours separately, as well as to identify the specific channels of influence.[42]

Finally, the figures report the response of wages, hours, and earnings to an exogenous job change. In this case, JC_{it} is set equal to 1 in period 10 for all individuals with $E_t = E_{t-1} = 1$, which one should think of as resulting from a large positive realization of the i.i.d. component $\varepsilon_{it}^{\mathrm{JC}}$ that negatively affects the relative attractiveness of the current job, rather than from a large draw of $v'_{ij'(t)}$. The line marked with "×" shows the average response. Part of the decline in earnings reflects the value of lost tenure (.064). In addition, since the job change is not selective on v_{ij}, $v_{ij(t)}$ declines by $(1-\rho_v)\mathbf{E}(v_{ij(t)}|t=10)$ or .021. The line with triangles is the effect of an exogenous job change that is accompanied by a draw of $\varepsilon_{ij(t)}^{v}$ that is one standard deviation above its mean, or .276. The net positive effect is large and highly persistent. These results are mirrored in wages (panel (b)). In addition, we show the effect of an exogenous job change that is accompanied by a one-standard-deviation increase of .162

[42]Kletzer (1998) surveyed the literature on job loss and wages. A number of studies examine how employer and industry tenure affects the size of the loss. When the problem of unobserved worker heterogeneity (but not job heterogeneity) is addressed, there appear to be modest tenure effects of the loss that are consistent with Altonji and William's (2005) estimates used here. Neal (1995) and Parent (2000) argued that industry tenure is more important than firm tenure. Kambourov and Manovskii (2009) argued that occupational tenure is more important than firm or industry tenure. As we noted earlier, one could extend the model we consider to include industry and occupation transition equations, but we leave this to future research.

in the job-specific hours component $\xi_{ij(t)}$. This is associated with a positive increase in hours worked and in earnings that decays in half in the first few years but slowly thereafter. Since $\xi_{ij(t)}$ is independent across jobs, the persistence stems from the fact that when t is greater than 10, job changes with or without unemployment are infrequent.[43]

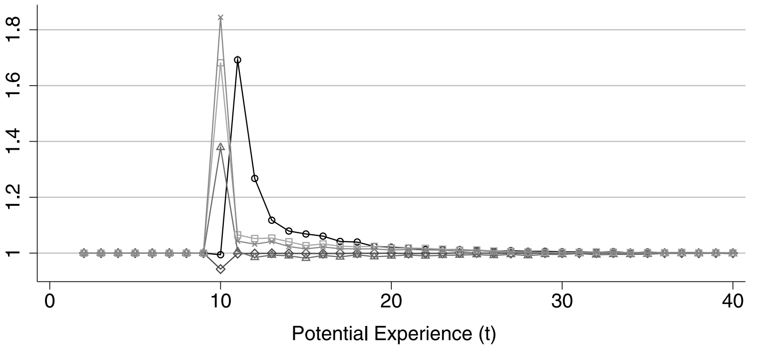
We also use the model to estimate the effects of an exogenous job loss and an exogenous job change on earnings variability using the methodology described above. The circle line in panel (a) of Figure 4 graphs the ratio of $\text{Var}(\text{earn}_{it} - \text{earn}_{i,t-1})$ following an exogenous unemployment shock when $t = 10$ to the baseline variance for the model. The variance ratio is slightly below 1 when $t = 10$, it is 1.69 when $t = 11$, declines to 1.27 when $t = 12$, and then slowly declines to 1 over the next 10 years. Panel (b) shows that the corresponding ratio for $\text{Var}(\text{earn}_{it})$ is about .85 when $t = 10$, presumably because differences in wages matter less when everyone is unemployed at the survey date. It rises to 1.10 when $t = 11$ and then slowly declines to about 1.05. An exogenous job change induces a big spike in the ratio for $\text{Var}(\text{earn}_{it} - \text{earn}_{i,t-1})$ when $t = 10$. The corresponding ratio for $\text{Var}(\text{earn}_{it})$ rises slowly following the shock, presumably because in some cases the exogenous job change induces additional ones. We have produced corresponding figures for shocks at $t = 3$ (not shown). The impact on the variance in that case is somewhat smaller and less persistent.

5.4. *Variance Decompositions*

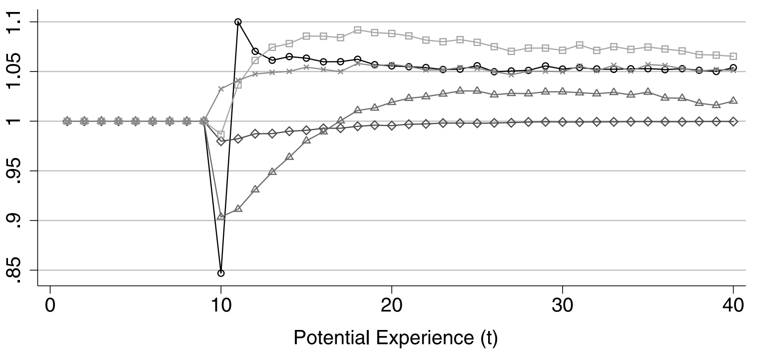
In this section, we use the model to measure the relative importance for the variance of earnings, wages, and hours of the initial condition and shocks to the autoregressive wage component, the i.i.d. hours shocks, the i.i.d. earnings shocks, job changes and employment spells and the associated shocks, the permanent heterogeneity components μ and η, and the effects of education for the white population. To do this, we first compute the variance of the sum of the annual values of the *levels* of earnings, wages, and hours over a 40-year career. We then repeat the simulation after setting the variance of the particular random component in the model to 0. We use the drop in the variance relative to the base case as the estimated contribution of the particular type of shock. Since the model is nonlinear, the contributions do not sum to 100% and may be negative.[44] We have normalized them to sum to 100. We report results

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

[44]A few of the estimated variances contributions are, in fact, negative. We have verified that variance in one shock can reduce the influence of other shocks. Variance in μ and in η increases heterogeneity in turnover behavior, which tends to reduce the variance in the sum of earnings and wages. On the other hand, the direct effects of μ on wages and hours and of η on hours increases variance.



(a) Response of cross-sectional variance of the first difference of log earnings to various shocks at $t = 10$



(b) Response of cross-sectional variance of log earnings to various shocks at $t = 10$

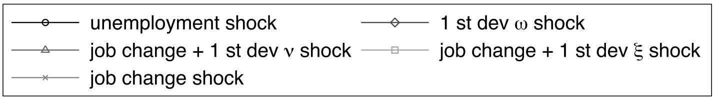


FIGURE 4.—Panel (a) in the figure displays the response of the ratio of $\text{Var}(\text{earn}_{it} - \text{earn}_{i,t-1})$ to the baseline variance for the model, to various shocks that are imposed when potential experience $t = 10$. See note in Figure 3. Panel (b) displays the response of the ratio of $\text{Var}(\text{earn}_{it})$ to the baseline variance for the model.

for the levels of variables, accounting for the experience profile in all variables. The decompositions of the sums of the annual values of logs of earnings, hours, and wages are similar (not reported). We use the parametric bootstrap distribution of $\hat{\beta}$ to estimate the standard errors of variance contributions, which are reported in parentheses.

The results are in Table VI.A. The first row refers to the sum of lifetime earnings. The earnings shocks ε_{it}^{e} account for only 5.9% of the variance in lifetime earnings even though they account for about 15% of $\text{Var}(\text{earn}_{it})$ in a given year (Table VI.B). The reason for the relatively small contribution to lifetime earn-

TABLE VI.A

DECOMPOSITION OF CROSS-SECTIONAL VARIANCE IN LIFETIME EARNINGS, WAGE, AND HOURS (IN LEVELS). BASELINE MODEL, FULL SRC SAMPLE[a]

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
	Contribution to Variance							Breakdown of 'Composite'			
Variable	ε^e	ε^h	ε^ω	Composite	η	μ	EDUC	ξ	υ	E	JC
Lifetime Earnings	**5.9** (0.3)	**1.7** (0.1)	**9.5** (1.0)	**43.0** (3.3)	**−4.7** (2.0)	**15.9** (4.2)	**28.7** (2.2)	**8.4** (1.4)	**33.9** (3.3)	**1.5** (0.4)	**−0.7** (0.3)
Lifetime Wage	**0** (0.0)	**0** (0.0)	**15.4** (1.6)	**53.2** (3.4)	**−6.0** (2.3)	**4.7** (5.0)	**32.7** (3.3)	**0** (0.0)	**52.8** (3.5)	**1.2** (0.4)	**−0.8** (0.4)
Lifetime Hours	**0** (0.0)	**3.6** (0.2)	**0.5** (0.2)	**58.9** (10.1)	**1.5** (4.0)	**32.9** (11.3)	**2.6** (0.8)	**54.2** (9.8)	**1.2** (0.6)	**3.6** (0.6)	**−0.1** (0.1)

[a]Entries in columns I to VII display the contribution of a given type of shock to the variance of lifetime earnings, wage, and hours, and are expressed as a percentage of the lifetime variance in the basecase. In the basecase, we simulate the full estimated model. To compute the contribution of a particular shock, we simulate the model again, setting the variance of a given shock to zero for all t. We then compute the variance of the appropriate variables. The difference relative to the basecase is the contribution of the given shock. Since the model is nonlinear, the contributions do not sum up to 100%. We normalize columns I to VII to sum to 100. Column III is the combined contribution of the initial draw of ω_{i1} and the subsequent shocks $\varepsilon_{it}^{\omega}$. Column IV is the combined contribution of the job match wage and hours components, employment and unemployment shocks, and job change shocks. In columns VIII through XI, we decompose column IV. Column VIII shows the marginal contribution of ξ, IX the marginal contribution of υ with $\mathrm{Var}(\xi)$ set to 0, X the marginal contribution of unemployment spells with $\mathrm{Var}(\xi)$ and $\mathrm{Var}(\upsilon)$ set to 0, and column XI displays the marginal contribution of job changes with $\mathrm{Var}(\xi)$ and $\mathrm{Var}(\upsilon)$ set to 0, and no unemployment. The variance of the levels of lifetime earnings, wages, and hours are 560,278; 90,803; and 253,504,501, respectively. Bootstrap standard errors are in parentheses.

ings is that the shocks are not very persistent. Similarly, the i.i.d. hours shocks ε_{it}^{h} contribute only 1.7% of the variance in lifetime earnings but account for between 6.2% and 7.5% of the variance in a year's earnings (Table VI.B). One can easily self-insure against these shock categories. In contrast, in column III, the initial condition $\varepsilon_{i1}^{\omega}$ and the i.i.d. shocks to ω_{it} are together responsible for 9.5% of the variance in lifetime earnings. The 9.5% figure reflects the fact that these shocks account for 15.4% of the variance in lifetime wages.[45] They contribute little to the variance in hours because the response of hours to wages is small.

The most striking result is in column IV, which shows the collective impact of job-specific hours and wage components, unemployment spells, and job changes. Altogether, mobility and unemployment-related shocks account for 43.0%, 53.2%, and 58.9% of the variance in lifetime earnings, wages, and hours, respectively. Given the interactions among the job change and employment related factors, we break down their relative contributions by first turning off the job-specific hours shocks, then turning off both job-specific hours

[45]The separate contributions of ω_{i1} and the shocks $\varepsilon_{it}^{\omega}$ are 3.0 and 6.5 percent in the case of earnings and 4.1 and 11.3 percent in the case of wages.

TABLE VI.B

DECOMPOSITION OF CROSS-SECTIONAL VARIANCE IN EARNINGS, WAGE, AND HOURS IN LEVELS AT DIFFERENT t (POTENTIAL EXPERIENCE). BASELINE MODEL, FULL SRC SAMPLE[a]

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Variable/	Contribution to Variance							Breakdown of 'Composite'				
Potential Experience	ε^e	ε^h	ε^ω	Composite	η	μ	EDUC	ξ	υ	E	JC	Variance
Earnings												
$t=1$	**11.1**	**7.5**	**28.1**	**18.4**	**0.4**	**11.9**	**22.5**	**9.4**	**7.9**	**1.0**	**0**	115.82
	(0.4)	(0.3)	(3.2)	(2.0)	(0.4)	(3.4)	(1.2)	(1.6)	(1.5)	(0.2)	(0.0)	
$t=5$	**14.6**	**6.2**	**19.3**	**29.3**	**0.5**	**10.0**	**20.1**	**8.6**	**21.5**	**0.9**	**−1.7**	256.95
	(0.6)	(0.3)	(1.7)	(2.2)	(0.8)	(3.1)	(1.2)	(1.4)	(2.2)	(0.2)	(0.2)	
$t=10$	**15.5**	**6.7**	**14.6**	**33.9**	**−1.5**	**11.0**	**19.9**	**8.2**	**26.7**	**1.0**	**−2.0**	447.63
	(0.8)	(0.4)	(1.3)	(2.6)	(1.1)	(3.2)	(1.3)	(1.3)	(2.6)	(0.3)	(0.3)	
$t=20$	**15.4**	**6.9**	**12.2**	**37.9**	**−2.1**	**10.5**	**19.1**	**7.3**	**30.4**	**1.2**	**−1.1**	745.96
	(0.9)	(0.5)	(1.2)	(2.5)	(1.5)	(3.1)	(1.4)	(1.4)	(2.6)	(0.3)	(0.2)	
$t=30$	**14.7**	**6.6**	**10.9**	**39.9**	**−1.7**	**10.8**	**18.8**	**7.7**	**31.9**	**0.7**	**−0.4**	796.75
	(0.8)	(0.5)	(1.1)	(2.6)	(2.1)	(3.2)	(1.4)	(1.3)	(2.6)	(0.2)	(0.2)	
$t=40$	**15.3**	**6.6**	**11.4**	**40.6**	**−1.8**	**9.0**	**18.9**	**7.6**	**32.9**	**0.3**	**−0.2**	771.98
	(0.9)	(0.5)	(1.2)	(2.5)	(2.1)	(3.3)	(1.4)	(1.3)	(2.6)	(0.2)	(0.1)	
Wage												
$t=1$	**0**	**0**	**52.0**	**15.9**	**0**	**4.0**	**28.1**	**0**	**15.9**	**0**	**0**	16.54
	(0.0)	(0.0)	(4.8)	(2.8)	(0.0)	(4.2)	(1.5)	(0.0)	(2.8)	(0.0)	(0.0)	
$t=5$	**0**	**0**	**33.9**	**37.1**	**0.8**	**3.2**	**24.9**	**0**	**38.5**	**0.4**	**−1.8**	34.39
	(0.0)	(0.0)	(2.9)	(2.8)	(0.7)	(3.8)	(1.7)	(0.0)	(2.8)	(0.2)	(0.3)	
$t=10$	**0**	**0**	**27.1**	**47.0**	**−2.3**	**2.9**	**25.3**	**0**	**48.1**	**0.8**	**−1.9**	56.93
	(0.0)	(0.0)	(2.4)	(3.1)	(1.1)	(4.2)	(2.1)	(0.0)	(3.0)	(0.3)	(0.4)	
$t=20$	**0**	**0**	**21.8**	**52.5**	**−2.9**	**4.2**	**24.4**	**0**	**52.2**	**1.2**	**−0.8**	94.33
	(0.0)	(0.0)	(2.1)	(3.1)	(1.8)	(4.0)	(2.3)	(0.0)	(3.1)	(0.4)	(0.4)	
$t=30$	**0**	**0**	**19.6**	**54.1**	**−1.4**	**3.8**	**23.9**	**0**	**53.5**	**0.9**	**−0.3**	107.69
	(0.0)	(0.0)	(2.0)	(3.0)	(1.9)	(3.9)	(2.3)	(0.0)	(3.0)	(0.3)	(0.2)	
$t=40$	**0**	**0**	**20.4**	**54.2**	**−0.3**	**2.8**	**22.8**	**0**	**53.9**	**0.4**	**−0.1**	116.95
	(0.0)	(0.0)	(2.1)	(3.0)	(2.0)	(4.0)	(2.4)	(0.0)	(3.0)	(0.2)	(0.1)	

(*Continues*)

and wage shocks, then turning off hours, wage, and unemployment shocks, and finally turning off hours, wage, and unemployment and the idiosyncratic job change shocks ($\varepsilon_{it}^{\mathrm{JC}}$). We choose this order because the employment transitions and job changes that are induced by $\varepsilon_{it}^{\mathrm{EE}}$, $\varepsilon_{it}^{\mathrm{UE}}$, and $\varepsilon_{it}^{\mathrm{JC}}$ matter for variance primarily because jobs pay different wages and require different hours rather than because of the direct impact of unemployment and job changes on wages and hours. The estimates are reported in columns VIII, IX, X, and XI. For earnings, job-specific wage shocks are more important than hours shocks. Job-specific wage shocks dominate for wages, while job-specific hours shocks dominate for hours but also contribute 8.4% to the earnings variance. Some of the changes in hours within and across jobs may be due to changes in preferences and may not represent risk.

TABLE VI.B—*Continued*

Variable/ Potential Experience	I ε^e	II ε^h	III ε^ω	IV Composite	V η	VI μ	VII EDUC	VIII ξ	IX υ	X E	XI JC	XII Variance
	Contribution to Variance							Breakdown of 'Composite'				
Hours												
$t=1$	**0**	**29.6**	**0.9**	**47.7**	**2.9**	**15.2**	**3.8**	**36.8**	**0.2**	**10.7**	**0**	356,086.27
	(0.0)	(1.4)	(0.3)	(5.7)	(1.8)	(5.0)	(0.3)	(5.6)	(0.1)	(0.4)	(0.0)	
$t=5$	**0**	**32.0**	**0.8**	**48.0**	**2.7**	**15.3**	**1.2**	**41.6**	**0.4**	**5.9**	**0.1**	351,037.44
	(0.0)	(1.3)	(0.2)	(6.2)	(1.6)	(5.3)	(0.4)	(6.3)	(0.3)	(1.0)	(0.0)	
$t=10$	**0**	**32.4**	**0.6**	**48.3**	**1.9**	**15.9**	**0.9**	**42.1**	**0.3**	**5.7**	**0.2**	372,140.59
	(0.0)	(1.4)	(0.2)	(6.3)	(1.4)	(5.3)	(0.4)	(6.2)	(0.3)	(0.6)	(0.1)	
$t=20$	**0**	**32.4**	**0.5**	**48.2**	**1.4**	**16.2**	**1.4**	**41.7**	**0.2**	**6.2**	**0.1**	377,299.33
	(0.0)	(1.4)	(0.2)	(6.5)	(2.0)	(5.6)	(0.5)	(6.4)	(0.5)	(0.6)	(0.0)	
$t=30$	**0**	**33.5**	**0.7**	**46.7**	**2.9**	**15.0**	**1.2**	**43.3**	**0.3**	**3.1**	**0.0**	352,501.77
	(0.0)	(1.5)	(0.2)	(6.4)	(1.6)	(5.5)	(0.5)	(6.4)	(0.5)	(0.6)	(0.0)	
$t=40$	**0**	**34.6**	**0.6**	**46.9**	**0.2**	**15.9**	**1.8**	**45.5**	**0.3**	**1.1**	**0.0**	308,443.34
	(0.0)	(1.6)	(0.2)	(6.8)	(2.0)	(5.8)	(0.5)	(6.8)	(0.4)	(0.6)	(0.0)	

[a]Entries in columns I to VII display the contribution of a given type of shock to the variance in earnings, wage, and hours for a cross section of simulated individuals with potential experience t. The contribution is expressed as a percentage of the variance in the basecase. In the basecase, we simulate the full estimated model. To compute the contribution of a particular shock, we simulate the model again, setting the variance of the given shock to zero for all t. We then compute the variance of the appropriate variables at the specified value of t. The difference relative to the basecase is the contribution of the given shock. Since the model is nonlinear, the contributions do not sum up to 100%. We have normalized columns I to VII to sum to 100. Column III is the combined contribution of the initial draw of ω_{i1} and the subsequent shocks $\varepsilon_{it}^{\omega}$. Column IV is the combined contribution of the job match wage and hours components, unemployment shocks, and job change shocks. In columns VIII through XI, we decompose column IV. Column VIII is the marginal contribution of ξ, IX is the marginal contribution of υ with Var(ξ) set to 0, X is the marginal contribution of eliminating unemployment spells with Var(ξ) and Var(υ) set to 0, and column XI is the marginal contribution of job changes with Var(ξ) and Var(υ) set to 0, and no unemployment. Column XII is the cross-sectional variance of simulated earnings, wage, and hours, across individuals with potential experience t. Bootstrap standard errors are in parentheses.

Finally, we turn to the three permanent heterogeneity components for whites: η, μ, and EDUC. Surprisingly, the estimates in column V indicate that the mobility component η does not play much of a role. The point estimate is actually negative. However, μ accounts for 15.9% (4.2%) of the variance in lifetime earnings and 32.9% (11.3%) of the variance in work hours, but only 4.7% (5.0%) of the variance in wages. The positive direct effect that μ has on the wage variance is partially offset by its role in reducing job changes and transitions into unemployment. Education is very important, contributing 28.7% of the variance in lifetime earnings and 32.7% of the variance in lifetime wages, but only 2.6% of the variance in lifetime hours. The combined variance contribution of η, μ, and EDUC and the initial draws ω_{i1} and υ_{i1} of ω_{it} and υ_{it} is 55.3% for lifetime earnings, 44.6% for lifetime wages, and 39.5% for lifetime hours.

5.5. *Sensitivity to Alternative Measurement Error Assumptions*

In Appendix D of the Supplemental Material, we present model estimates for alternative assumptions about measurement error. Relative to the standard errors, the changes in the parameters of the EE, UE, JC, earnings, and hours equations are minor. Most of the parameters of the wage equation are also insensitive to the measurement error assumptions, but there are four important exceptions. The coefficient δ^w_μ on the permanent productivity component μ falls from .081 (.035) for the parameter values we chose to .017 when we use the alternative, higher values of .130 for σ_{mw} and .121 for σ_{mh}. The decline in the importance of the fixed heterogeneity term is accompanied by an increase in ρ_v from .691 (.049) to .782, an increase in σ_{v1} from .165 to .243, a decline in σ_ω from .089 (.005) to .033, and a decline in the values of σ_{w1} for the four race-education categories. The net effect of these changes is to reduce the role of the permanent productivity component and the persistent wage component ω_{it} in the variation of wages across people and the persistence over time. The variance decompositions are qualitatively similar to the results for our assumed values of the measurement error variances, but show a large decline in the importance of μ and ω_{it} that is balanced by a large increase in the importance of shocks associated with job changes and employment transitions. The impulse response functions of earnings, wages, and hours to various shocks are virtually identical to those discussed above, with the obvious exception that the decline in σ_ω leads to a proportional decline in the effect of a one-standard-deviation shock to ω_{it}.

We strongly prefer the results based on the lower values for σ_{mw} and σ_{mh}. Given that we do not find evidence of a unit root in the wage process, a value close to 0 for δ^w_μ is implausible. For example, the substantial correlation across siblings and between parents and children in wage rates conditional on education and race points to a fixed heterogeneity component that is correlated across siblings and across generations.[46] However, it is important to emphasize that most of our results are not sensitive to the measurement error assumptions.

6. RESULTS FOR LOW AND HIGH EDUCATION SAMPLES

Columns V and VIII in Table IV report point estimates of the baseline model estimated on SRC subsamples of whites with a high school degree or less and whites with more than high school education. We focus on whites to avoid confounding the effect of education with the effect of race, given that blacks tend to be less educated than whites. Results for all whites (not shown) are basically similar to the results for the full sample.

[46]See Black and Devereux (2011) for a survey of the literature on family correlations in economic outcomes.

The point estimates are quite similar overall, given standard errors.[47] However, a few differences are worth noting. First, mobility is much less sensitive to seniority in the high education sample. Second, JC responds more positively to outside offers and more negatively to μ in the high education sample. Third, the relative importance of μ and η in the EE equation is reversed in the two subsamples. Fourth, unemployment is less common for the high education sample. The decompositions of the experience profile of wages in Figures B.1 and B.2 of the Supplemental Material show more growth in v with t in the educated sample: .148 versus .114 after 35 years. Both σ_ω and $\sigma_{\omega 1}$ are considerably larger for the higher education sample, and autoregressive parameters ρ_v and ρ_ω are also larger. These differences result in a much larger contribution of ω to the variance of wages and earnings for the more highly educated. The standard deviation of the i.i.d. component of hours is much larger for the less educated sample (.17 versus .10), which probably reflects greater variation in overtime hours and in unemployment spells. On the other hand, σ_ξ is larger for the high education group.

Figures B.3 through B.6 of the Supplemental Material report impulse response functions for the two groups. They are similar to those for the full sample.

The variance decompositions show that shocks associated with mobility and employment transitions play a key role in the variance of lifetime earnings, wages, and hours for both samples (Tables B.I–B.IV of the Supplemental Material). They account for 50.2% of variance in the sum of earnings for the high education group and 42.4% for the low education group. The corresponding values for wages are 56.2% and 66.6%. Employment shocks and i.i.d. hours shocks are more important for the low education sample. The job-specific hours component ξ is more important for the high education sample. The persistent wage component ω_{it} makes a big variance contribution in the high education sample and a small one in the low education, while the roles are reversed in the case of μ, which plays only a small role in the high education sample. (The point estimate of the contribution of μ to the variance of wages is actually a small negative.) Within-group variation in education is important in the high education sample. Although we have focused on the percentage contributions, it is important to point out that the variance of the sum of lifetime earnings is much larger for the high education sample: \$990,304 versus \$140,540 in year-2000 dollars.

[47]We modified the bootstrap standard error procedure to avoid computational problems associated with the relatively small size of the low education and the high education subsamples. Specifically, we computed each parametric bootstrap replication using double the size of the corresponding PSID subsample. The reported standard errors are $\sqrt{2}$ times the standard deviation of the parameter estimates across the 300 bootstrap replications.

7. A MULTINOMIAL MODEL OF EMPLOYMENT TRANSITIONS AND JOB CHANGES

In this section, we replace the sequential model of employment transitions and job-to-job changes consisting of equations (6) and (7) with a multinomial choice model of the decision to stay in the current job, move to another job, or leave employment. Let the latent variable EE^{S*} denote the value of remaining employed in the current job relative to the value of unemployment:

$$(19)\qquad \begin{aligned}\mathrm{EE}_{it}^{S*} &= X_{i,t-1}\gamma_X^{\mathrm{EE}^S} + \gamma_{\mathrm{TEN}}^{\mathrm{EE}^S}\mathrm{TEN}_{i,t-1} + \gamma_{\mathrm{ED}}^{\mathrm{EE}^S}\min(\mathrm{ED}_{i,t-1}, 9)\\ &\quad + \gamma_{w^s}^{\mathrm{EE}^S}\mathrm{wage}_{it}^s + \delta_{vj(t-1)}^{\mathrm{EE}^S} v_{ij(t-1)} + \delta_\mu^{\mathrm{EE}^S}\mu_i + \delta_\eta^{\mathrm{EE}^S}\eta_i\\ &\quad + \varepsilon_{it}^{\mathrm{EE}^S} \quad \text{given} \quad E_{i,t-1} = 1.\end{aligned}$$

The equation for EE_{it}^{S*} has the cleanest interpretation if continuing with a firm is entirely up to the worker. In reality, the coefficients capture both the effects of variables on the worker's valuation of the current job relative to unemployment and effects that operate through the worker's net value to the firm. The shocks $\varepsilon_{it}^{\mathrm{EE}^S}$ are also a mix of preference shocks and shocks to the productivity of the job match that are not fully reflected in the wage. A large negative shock to $\varepsilon_{it}^{\mathrm{EE}^S}$ could arise from a temporary labor supply shock or from a decline in firm productivity that leads to a layoff.[48] Employment duration and tenure both capture state dependence that arises from locational decisions, arrangements within the household, employment-based social networks, and other factors. In the case of employment duration, habit formation in work preferences could play a role. In the case of $\mathrm{TEN}_{i,t-1}$, part of the effect is the value of tenure-related increases in nonwage fringe benefits, such as pensions and paid vacation. Part is through the effects of $\mathrm{TEN}_{i,t-1}$ on the layoff probability that arise because firms share in specific human capital investments and/or follow seniority-based layoff policies. Note that all of the determinants of wage_{it}^s are allowed to have an independent influence on EE_{it}^{S*}, with the exception of ω_{it}, which affects log wages in all jobs equally.

EE_{it}^{Q*} is the value of moving to a new job relative to the utility of unemployment:

$$(20)\qquad \begin{aligned}\mathrm{EE}_{it}^{Q*} &= X_{i,t-1}\gamma_X^{\mathrm{EE}^Q} + \gamma_{\mathrm{TEN}}^{\mathrm{EE}^Q}\mathrm{TEN}_{i,t-1} + \gamma_{w'}^{\mathrm{EE}^Q}\mathrm{wage}_{it}' + \delta_{v'j'(t)}^{\mathrm{EE}^Q} v_{ij'(t)}'\\ &\quad + \delta_\mu^{\mathrm{EE}^Q}\mu_i + \delta_\eta^{\mathrm{EE}^Q}\eta_i + \varepsilon_{it}^{\mathrm{EE}^Q} \quad \text{given} \quad E_{i,t-1} = 1,\end{aligned}$$

where $v_{ij'(t)}'$ is a draw of the job-specific component for an alternative job $j'(t)$ in t and wage_{it}' is the value of $\mathrm{wage}_{it}^{\mathrm{lat}}$ evaluated at $v_{ij(t)} = v_{ij'(t)}'$ and $\mathrm{TEN}_{i,t-1} = 0$.

[48] It would be interesting in future work to expand the model to distinguish between quits and layoffs on the basis of self reports.

We include $\mathrm{TEN}_{i,t-1}$ because it may influence the costs of changing jobs. The coefficient $\delta^{\mathrm{EE}^Q}_{v'j'(t)}$ allows $v'_{ij'(t)}$ to influence job-to-job mobility independently of its effect on wage'_{it}.[49] The odds of actually getting an alternative offer are reflected in the parameters.

The relationship between $\varepsilon^{\mathrm{EE}^Q}_{it}$ and $\varepsilon^{\mathrm{EE}^S}_{it}$ depends on the relative importance of transitory availability of job opportunities versus labor supply preferences in determining the employment of male household heads at a point in time. If labor supply preferences are key, then one would expect a positive correlation between the two, since both compare the value of employment opportunities to unemployment. However, if labor market frictions and job destruction is important, then the two may be only weakly correlated. We assume that, for male household heads, frictions dominate, so that $\varepsilon^{\mathrm{EE}^Q}_{it}$ and $\varepsilon^{\mathrm{EE}^S}_{it}$ are uncorrelated. Keep in mind that both equations contain the permanent heterogeneity components μ_i and η_i.

The indicator EE^S_{it} for whether the worker remains employed in his current job is determined by

$$\mathrm{EE}^S_{it} = I\big(\mathrm{EE}^{S*}_{it} > \mathrm{EE}^{Q*}_{it} \text{ and } \mathrm{EE}^{S*}_{it} > 0\big) \quad \text{given} \quad E_{i,t-1} = 1. \tag{21}$$

The indicator EE^Q_i for whether the worker remains employed and moves to a new job is

$$\mathrm{EE}^Q_{it} = I\big(\mathrm{EE}^{S*}_{it} < \mathrm{EE}^{Q*}_{it} \text{ and } \mathrm{EE}^{Q*}_{it} > 0\big) \quad \text{given} \quad E_{i,t-1} = 1. \tag{22}$$

The indicator EE_{it} is $\mathrm{EE}^S_{it} + \mathrm{EE}^Q_{it}$. The indicator JC_{it} for a job change conditional on remaining employed is $\mathrm{EE}^Q_{it} | \mathrm{EE}_{it} = 1$. Since a job-to-job move involves a comparison of EE^{S*} and EE^{Q*} as well as a comparison of the value of changing jobs to unemployment, variables and error components influence EE_{it}, EE^Q_{it}, and JC_{it} through both (19) and (20).

7.1. *Results for the Multinomial Model*

Appendix C of the Supplemental Material presents the results for the multinomial model. Here we provide a few highlights. The predictions and fit of the model to the PSID are similar to those for the baseline model. Both employment duration and tenure raise EE, and tenure has a substantial negative effect on JC. The values of $v_{ij(t-1)}$ and $v'_{ij'(t)}$ both increase the probability that the worker remains employed. As in the baseline model, the probability that

[49] An independent effect should arise from at least two mechanisms. First, $v'_{ij'(t)}$ is job-specific and differs in persistence from the other wage components. Second, to the extent that it reflects match-specific productivity that is shared by the worker and firm, it will also be positively associated with the firm's valuation of the match conditional on the wage and thus be negatively related to the layoff probability. Workers should value the security.

an employed worker moves to a new job depends negatively on $\upsilon_{ij(t-1)}$ and positively on $\upsilon'_{ij'(t)}$, but the effect of $\upsilon_{ij(t-1)}$ is close to 0 and $\upsilon'_{ij'(t)}$ is relatively more important in the multinomial case. Overall, the estimates of the response of earnings, the wage rate, and hours to various shocks are very similar to those for the baseline model. In particular, an unemployment shock leads to a decline in log earnings of $-.58$ due to declines in both hours and wages. Hours recover quickly, but the wage loss persists. The positive effect of an unemployment shock on the variance of the first difference in earnings is more persistent in the multinomial model. General skill accumulation, job shopping (i.e., growth in $\upsilon_{ij(t)}$), and accumulation of tenure account for 81.0%, 7.3%, and 13.9% of the implied .819 increase in log wages over the first 30 years of a career. The contribution of job shopping is smaller than in the baseline model and seems implausibly low to us.

Qualitatively, the multinomial results are similar to those for the baseline model, in that shocks associated with employment and job mobility play a very large role. They account for 28.1%, 35.6%, and 53.6% of the variance of lifetime earnings, lifetime wage rates, and lifetime hours. These values are large, but are smaller than the baseline estimates. On the other hand, the permanent heterogeneity components η and especially μ play a more important role than in the baseline model. The μ component accounts for 24.5% of the variance in earnings, 14.6% of the variance in wages, and 22.4% of the variance in hours, which compare to 15.9%, 4.7%, and 32.9% in the baseline. The larger contribution of μ to earnings and wages stems from the fact that the factor loading δ_μ^w on μ in the wage equation is larger in the multinomial model than in the baseline model. The standard deviation of $\upsilon_{ij(t)}$ is also smaller in the multinomial case than in the baseline model, which helps explain the reduced contribution of employment and job mobility.

8. CONCLUSION

In this paper, we study earnings across individuals and over careers. To this end, we construct a model of earnings dynamics from equations governing employment transitions, job changes without unemployment, wages, and work hours. Since both state dependence and heterogeneity are important and one cannot determine the role of one without accounting for the other, our models incorporate state dependence in employment, job changes, and wages, while also including multiple sources of unobserved heterogeneity as well as job-specific error components in both wages and hours. These turn out to play an important role in the variance of lifetime earnings. The equations of our model can be viewed as approximations to the decision rules suggested by structural models of employment transitions, job search, and labor supply, while at the same time providing a rich statistical description of the earnings process. Our simulation-based estimation strategy permits us to handle a highly unbalanced sample in the context of a dynamic model that mixes discrete and continuous variables and allows for both state dependence and multifactor heterogeneity and for measurement error. Vidangos (2009) showed the potential for using models of the type we develop by studying the implications of a related multi-equation model of family income for precautionary behavior, welfare, and the value of insurance within the context of a lifecycle consumption model.[50]

Our results address many important questions concerning wages, hours, and earnings over a career. In accord with many other studies, we find that education, race, and unobserved permanent heterogeneity all play important roles in employment transitions and job changes, and that labor supply of male household heads is inelastic. By accounting for both unobserved individual heterogeneity and job-specific heterogeneity, we are able to show that a substantial portion of the strong negative relationship between job seniority and job mobility found in many previous studies is causal. Job changes are induced by high outside offers and deterred by the job-specific wage component of the current job. Job offers are strongly positively related to the job-specific component of the current job, in contrast to the usual assumption in the search literature that offers are drawn at random. The dependence may arise because firms base offers to prospective new hires in part on wages in the prior firm, because the job-specific component partially reflects demand shocks affecting jobs in a narrowly defined industry, occupation, and region, and/or because an individual's job search network depends on the quality of his current job.

Overall, wages are highly persistent, but do not contain a random walk component. The persistence results from permanent heterogeneity, the job-specific wage component, and strong persistence in the stochastic component that reflects the value of the worker's general skills.

We also contribute to the displaced workers literature by providing a full decomposition of earnings losses from unemployment. Short-term earnings losses from unemployment are dominated by hours and long-term losses are dominated by wages, with lost tenure, movement to a lower-paying job, and a drop in the autoregressive skill component all playing a role. We find general human capital accumulation is the dominant source of wage growth over a career, although job tenure and job mobility both play significant roles.

Finally, job mobility and unemployment play a key role in the variance of career earnings. They operate primarily by leading to large changes in job-specific components of wages and hours rather than through their direct effects on wages and hours. For whites in our full sample, job-specific hours and wage components, unemployment shocks, and job shocks together account for 43.0%, 53.2%, and 58.9% of the variance in lifetime earnings, wages, and hours, respectively. Job-specific wage shocks are more important than job-specific hours shocks for earnings. Job-specific wage shocks dominate for wages, while job-specific hours shocks dominate for hours. Education accounts

[50]He allowed for additional sources of variation in family income, such as health and disability shocks, but used a simpler model of job mobility. See also Low, Meghir, and Pistaferri (2010).

for about 30% of the variance in lifetime earnings and wages, but makes little difference for hours. The combined variance contribution of variables determined by the first year of employment (η, μ, and EDUC and the initial draws ω_{i1} and υ_{i1} of ω_{it} and υ_{it}) is 55.3% for lifetime earnings, 44.6% for lifetime wages, and 39.5% for lifetime hours.

There are a number of extensions to the model that would be worth exploring. Thus far, we simply remove year effects from wages, hours, and earnings, but it would be natural to add aggregate shocks to the model. It would also be natural to extend the model to explore changes in the stability of earnings, building on work by Gottschalk and Moffitt (1994, 2008), Haider (2001), Shin and Solon (2008), DeBacker, Heim, Panousi, Ramnath, and Vidangos (2013), and others. This would require a very different auxiliary model. With matched employer-employee data such as those used by Abowd et al. (1999) and Bagger et al. (2011), one could distinguish firm-specific risk associated with observed as well as unobserved variables from job-match-specific risk. A much more ambitious extension would be 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.

Given the large number of issues that the paper already addresses, we do not attempt the formidable task of seeking to identify how much of the stochastic variation in earnings that we analyze is anticipated by agents, how far in advance they anticipate it, how much is insured, and how much is an endogenous response to changes in opportunity sets or preferences. Adding a family income model (with private and public transfers), as in Vidangos (2009), gets partially at the question of insurance. Dealing with expectations is more difficult. One needs either data on expectations or an expanded model that incorporates decisions that depend on and/or reveal the information set of the agent, such as consumption choices. Work by Blundell and Preston (1998), Blundell, Pistaferri, and Preston (2008), Cunha, Heckman, and Navarro (2005), and Guvenen and Smith (2010) illustrate the latter approach. A fully structural model that incorporates search frictions and hours constraints is probably needed to separate the role of preferences from labor market constraints.

APPENDIX A: DECOMPOSING CAREER WAGE GROWTH INTO THE EFFECTS OF GENERAL HUMAN CAPITAL, TENURE, AND JOB SHOPPING

Let $\text{wage}^o_{it} \equiv (\text{wage}_{it} - \delta^w_\mu \mu_i - \gamma^w_{\text{BLACK}}\text{BLACK}_i - \gamma^w_{\text{EDUC}}\text{EDUC}_i)$. Wage growth over a career is the sum of the effect of general human capital accumulation, the accumulation of job tenure, the gains from job shopping, and the cumulative effect of unemployment shocks on the general wage component ω_{it}. That is,

$$\begin{aligned}\mathbf{E}(\text{wage}^o_{it}|t) = {} & \left[t\gamma^w_t + t^2\gamma^w_{t2} + t^3\gamma^w_{t3}\right] + \mathbf{E}\big(P(\text{TEN}_{it})|t\big)\gamma^w_{\text{TEN}} \\ & + \mathbf{E}(\upsilon_{ij(t)}|t) + \mathbf{E}(\omega_{it}|t),\end{aligned}$$

where the terms are implicitly conditional on employment $E_{it}=1$, $[t\gamma_t^w + t^2\gamma_{t2}^w + t^3\gamma_{t3}^w]$ is the value of the general human capital cubic polynomial in potential experience t, $\mathbf{E}(P(\text{TEN}_{it})|t)\gamma_{\text{TEN}}^w$ is the expected value of the tenure polynomial, and $\mathbf{E}(v_{ij(t)}|t)$ and $\mathbf{E}(\omega_{it}|t)$ are the expected values of the job match and general productivity components. We use simulated data from the model to compute the values of $\mathbf{E}(\omega_{it}|t)$, $\mathbf{E}(v_{ij(t)}|t)$, and $\mathbf{E}(P(\text{TEN}_{it})|t)\hat{\gamma}_{\text{TEN}}^w$, where $\hat{\gamma}_{\text{TEN}}^w$ is taken from Altonji and Williams (2005). Figure 1 graphs

$$\mathbf{E}(\text{wage}_{it}^o|t), \quad [t\gamma_t^w + t^2\gamma_{t2}^w + t^3\gamma_{t3}^w],$$
$$\mathbf{E}(P(\text{TEN}_{it})|t)\gamma_{\text{TEN}}^w, \quad \text{and} \quad \mathbf{E}(v_{ij(t)}|t),$$

with value at $t=1$ set to 0 in each case. $\mathbf{E}(\omega_{it}|t)$ takes on the values $-.008$, $-.016$, $-.024$, $-.020$, and $-.012$ when t is 5, 10, 20, 30, and 40, respectively. It is not displayed to reduce clutter.

The above calculations include both employed and unemployed individuals and thus reflect actual wages for the employed and the "latent" wage for the unemployed, for whom TEN_{it} is 0. The values of $\mathbf{E}(\omega_{it}|t, E_{it}=1)$, $\mathbf{E}(\text{wage}_{it}^o|t, E_{it}=1)$, $[t\gamma_t^w + t^2\gamma_{t2}^w + t^3\gamma_{t3}^w]$, $\mathbf{E}(P(\text{TEN}_{it})|t, E_{it}=1)\gamma_{\text{TEN}}^w$, and $\mathbf{E}(v_{ij(t)}|t, E_{it}=1)$ are very similar to unconditional values and are not reported. The small differences reflect the fact that the distribution of ω_{it}, $v_{ij(t)}$, and TEN_{it} at each value of t is related to employment status. Note that part of the relationship between t and wages in panel data restricted to employed workers is due to selection. The positive dependence of employment on the wage means that selection into employment on μ, v, and ω varies slightly with experience. For example, $\mathbf{E}(\omega|t, E_{it}=1)$ declines by $-.01$ over the first 10 years and $-.016$ over the first 30 years. Our estimates of the experience profile account for selection bias stemming from all of the error components.

REFERENCES

ABOWD, J. M., AND D. E. CARD (1987): "Intertemporal Labor Supply and Long-Term Employment Contracts," *American Economic Review*, 77 (1), 50–68. [1396]

——— (1989): "On the Covariance Structure of Hours and Earnings Changes," *Econometrica*, 57 (2), 411–445. [1396]

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67 (2), 251–333. [1450]

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. [1405]

ALTONJI, J. G., AND R. A. SHAKOTKO (1987): "Do Wages Rise With Job Seniority?" *Review of Economic Studies*, 54, 437–459. [1417]

ALTONJI, J. G., AND N. WILLIAMS (2005): "Do Wages Rise With Job Seniority? A Reassessment," *Industrial and Labor Relations Review*, 58 (3), 370–397. [1402,1409,1417,1428,1438,1451]

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. [1396,1418]

ALTONJI, J. G., A. A. SMITH, JR., AND I. VIDANGOS (2009): "Modeling Earnings Dynamics," Working Paper 14743, NBER. [1407,1414,1417]

——— (2013): "Supplement to 'Modeling Earnings Dynamics'," *Econometrica Supplemental Material*, 81, http://www.econometricsociety.org/ecta/ Supmat/8415_miscellaneous.pdf; http://www.econometricsociety.org/ecta/Supmat/8415_data_and_programs.zip. [1404]

BAGGER, J., F. FONTAINE, F. POSTEL-VINAY, AND J. M. ROBIN (2011): "Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," Unpublished Paper, University of Bristol. [1397,1398,1402,1429,1450]

BAKER, M. (1997): "Growth-Rate Heterogeneity and the Covariance Structure of Life Cycle Earnings," *Journal of Labour Economics*, 15 (2), 338–375. [1396]

BAKER, M., AND G. SOLON (2003): "Earnings Dynamics and Inequality Among Canadian Men, 1976–1992: Evidence From Longitudinal Income Tax Records," *Journal of Labor Economics*, 21, 289–321. [1396]

BARLEVY, G. (2008): "Identification of Search Models Using Record Statistics," *Review of Economic Studies*, 75 (1), 29–64. [1397]

BLACK, S. E., AND P. J. DEVEREUX (2011): "Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics*, Vol. 4b, ed. by O. Ashenfelter and D. Card. Amsterdam: Elsevier, 1487–1541. [1444]

BLUNDELL, R., AND T. MACURDY (1999): "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics*, Vol. 3A, ed. by O. Ashenfelter and D. Card. Amsterdam: North-Holland, 1559–1695. [1397]

BLUNDELL, R., AND I. PRESTON (1998): "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics*, 113 (2), 603–640. [1450]

BLUNDELL, R., M. BREWER, AND M. FRANCESCONI (2008): "Job Changes, Hours Changes and the Path of Labour Supply Adjustment," *Journal of Labor Economics*, 26 (3), 421–453. [1405]

BLUNDELL, R., L. PISTAFERRI, AND I. PRESTON (2008): "Consumption Inequality and Partial Insurance," *American Economic Review*, 98 (5), 1887–1921. [1396,1450]

BOUND, J., C. BROWN, AND N. MATHIOWETZ (2001): "Measurement Error in Survey Data," in *Handbook of Econometrics*, Vol. 5, ed. by J. J. Heckman and E. E. Leamer. Amsterdam: Elsevier, 3705–3843. [1418]

BUCHINSKY, M., D. FOUGÈRE, F. KRAMARZ, AND R. TCHERNIS (2010): "Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.," *Review of Economic Studies*, 77 (3), 972–1001. [1397,1403,1428]

CASTAÑEDA, A., J. DÍAZ-GIMÉNEZ, AND V. RÍOS-RULL (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111 (4), 818–857. [1396]

CUNHA, F., J. J. HECKMAN, AND S. NAVARRO (2005): "Separating Uncertainty From Heterogeneity in Life Cycle Earnings, the 2004 Hicks Lecture," *Oxford Economic Papers*, 57 (2), 191–261. [1450]

DEATON, A. (1991): "Saving and Liquidity Constraints," *Econometrica*, 59 (5), 1221–1248. [1396]

DEBACKER, J., B. HEIM, V. PANOUSI, S. RAMNATH, AND I. VIDANGOS (2013): "Rising Inequality: Transitory or Persistent?: New Evidence From a U.S. Panel of Tax Returns," *Brookings Papers on Economic Activity* (forthcoming). [1450]

FARBER, H. (1999): "Mobility and Stability: The Dynamics of Job Change in Labor Markets," in *Handbook of Labor Economics*, Vol. 3B, ed. by O. Ashenfelter and D. Card. Amsterdam: North-Holland, 2439–2483. [1403]

GALLANT, R., AND G. TAUCHEN (1996): "Which Moments to Match?" *Econometric Theory*, 12, 657–681. [1398]

GOTTSCHALK, P., AND R. MOFFITT (1994): "The Growth of Earnings Instability in the U.S. Labor Market," *Brookings Papers on Economic Activity*, 1994 (2), 217–272. [1450]

——— (2008): "Trends in the Transitory Variance of Male Earnings in the U.S.: 1970–2004," Unpublished Manuscript. [1450]

GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118. [1398,1412]

GUVENEN, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review*, 97 (3), 687–712. [1396]

GUVENEN, F., AND A. A. SMITH, JR. (2010): "Inferring Labor Income Risk From Economic Choices: An Indirect Inference Approach," Unpublished Manuscript, available at http://dx.doi.org/10.2139/ssrn.1649092. [1450]

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. [1396,1450]

HAM, J., AND K. REILLY (2002): "Testing Intertemporal Substitution, Implicit Contract, and Hours Restriction Models of the Labor Market Using Micro Data," *American Economic Review*, 92 (4), 905–927. [1397]

HECKMAN, J. (1981): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time–Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data With Econometric Applications*, ed. by C. F. Manski and D. McFadden. Cambridge, MA: MIT Press, 179–195. [1412]

HUGGETT, M. (1996): "Wealth Distribution in Life-Cycle Economies," *Journal of Monetary Economics*, 38 (3), 469–494. [1396]

JACOBSON, L., R. LALONDE, AND D. SULLIVAN (1993): "Earnings Losses of Displaced Workers," *American Economic Review*, 83 (4), 685–709. [1436]

KAMBOUROV, G., AND I. MANOVSKII (2009): "Occupation Specificity of Human Capital," *International Economic Review*, 50 (1), 63–115. [1402,1438]

KEANE, M., AND A. A. SMITH, JR. (2003): "Generalized Indirect Inference for Discrete Choice Models," Unpublished Manuscript, Yale University. [1398,1411,1412]

KLETZER, L. G. (1998): "Job Displacement," *The Journal of Economic Perspectives*, 12 (1), 115–136. [1438]

KRUSELL, P., AND A. A. SMITH, JR. (1997): "Income and Wealth Heterogeneity, Portfolio Selection, and Equilibrium Asset Returns," *Macroeconomic Dynamics*, 1, 387–422. [1396]

——— (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106 (5), 867–896. [1396]

——— (1999): "On the Welfare Effects of Eliminating Business Cycles," *Review of Economic Dynamics*, 2, 254–272. [1396]

LILLARD, L., AND Y. WEISS (1979): "Components of Variation in Panel Earnings Data: American Scientists 1960–1970," *Econometrica*, 47 (2), 437–454. [1396]

LILLARD, L., AND R. WILLIS (1978): "Dynamic Aspects of Earning Mobility," *Econometrica*, 46 (5), 985–1012. [1396]

LOW, H., C. MEGHIR, AND L. PISTAFERRI (2010): "Wage Risk and Employment Risk Over the Life Cycle," *American Economic Review*, 100 (4), 1432–1467. [1397,1403,1427,1449]

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. [1396]

MEGHIR, C., AND L. PISTAFERRI (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72 (1), 1–32. [1396]

NAGYPAL, E. (2007): "Learning-by-Doing versus Learning About Match Quality: Can We Tell Them Apart?" *Review of Economic Studies*, 74 (2), 537–566. [1398]

NEAL, D. (1995): "Industry-Specific Human Capital: Evidence From Displaced Workers," *Journal of Labor Economics*, 13 (4), 653–677. [1402,1438]

PARENT, D. (2000): "Industry-Specific Capital and the Wage Profile: Evidence From the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics," *Journal of Labor Economics*, 18 (2), 306–323. [1402,1438]

POSTEL-VINAY, F., AND J.-M. ROBIN (2002): "Equilibrium Wage Dispersion With Worker and Employer Heterogeneity," *Econometrica*, 70 (6), 2295–2350. [1402]

POSTEL-VINAY, F., AND H. TURON (2010): "On-the-Job Search, Productivity Shocks, and the Individual Earnings Process," *International Economic Review*, 51 (3), 599–629. [1402]

RUBINSTEIN, Y., AND Y. WEISS (2006): "Post Schooling Wage Growth: Investment, Search and Learning," in *Handbook of the Economics of Education*, Vol. 1, ed. by E. A. Hanushek and F. Welch. Amsterdam: North Holland. [1429]

SHIN, D., AND G. SOLON (2008): "Trends in Men's Earnings Volatility: What Does the Panel Study of Income Dynamics Show?" Working Paper W14075, NBER. [1450]

SMITH, A. A., JR. (1990): "Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models," Ph.D. Thesis, Duke University. [1398]

——— (1993): "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, 8, S63–S84. [1398]

STEVENS, A. H. (1997): "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses," *Journal of Labor Economics*, 15 (1), 165–188. [1437]

STORESLETTEN, K., C. TELMER, AND A. YARON (2001): "The Welfare Costs of Business Cycles Revisited: Finite Lives and Cyclical Variation in Idiosyncratic Risk," *European Economic Review*, 45, 1311–1339. [1396]

——— (2007): "Asset Pricing With Idiosyncratic Risk and Overlapping Generations," *Review of Economic Dynamics*, 10 (4), 519–548. [1396]

TOPEL, R., AND M. WARD (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107 (2), 439–479. [1428]

VIDANGOS, I. (2009): "Household Welfare, Precautionary Saving, and Social Insurance Under Multiple Sources of Risk," Finance and Economics Discussion Series 2009-09, Board of Governors of the Federal Reserve System. [1449,1450]

WOLPIN, K. (1992): "The Determinants of Black–White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth," *Journal of Political Economy*, 100 (3), 535–560. [1397]

WOOLDRIDGE, J. (2005): "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models With Unobserved Heterogeneity," *Journal of Applied Econometrics*, 20 (1), 39–54. [1412]

Dept. of Economics, Yale University, P.O. Box 208264, New Haven, CT 06520-8264, U.S.A. and NBER; joseph.altonji@yale.edu,

Dept. of Economics, Yale University, P.O. Box 208268, New Haven, CT 06520-8268, U.S.A. and NBER; tony.smith@yale.edu,

and

Federal Reserve Board, 20th and C Streets, NW, Washington, DC 20551, U.S.A.; ivan.vidangos@frb.gov.

Manuscript received February, 2009; final revision received April, 2012.