Journal of Econometrics 113 (2003) 289–335

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Bayesian analysis of a dynamic stochastic model of labor supply and saving

Daniel Houser*

University of Arizona, Department of Economics, Eller College of Business and Public Administration, McClelland Hall #401, Tucson, AZ 25721-0108, USA

Abstract

This paper empirically implements a dynamic, stochastic model of life-cycle labor supply and human capital investment. The model allows agents to be forward looking. But, in contrast to prior literature in this area, it does not require that expectations be formed "rationally". By avoiding strong assumptions about expectations, I avoid sources of bias stemming from misspecification of the expectation process. A Bayesian econometric method based on Geweke and Keane (in: R.S. Mariano, T. Schuermann, M. Weeks (Eds.), Simulation Based Inference and Econometrics: Methods and Applications, Cambridge University Press, Cambridge, 1999) is used to relax assumptions over expectations. The results of this study are consistent with findings from previous research in the labor supply literature that makes the rational expectations assumption.
© 2002 Elsevier Science B.V. All rights reserved.

JEL classification: J22; J24; D84

Keywords: Life-cycle; Labor supply; Human capital; Expectations

1. Introduction

A substantial amount of recent empirical research has analyzed models of life-cycle labor supply that include human capital. Such research has typically taken dogmatic stands on the way people form expectations. In principle, however, inferences about human capital's effect on labor supply can be sensitive to the form of the expectations process. In this paper I empirically implement a dynamic, stochastic model of life-cycle labor supply that incorporates human capital but does not make strong assumptions about the way people form expectations. The estimated model provides wage and

* Tel.: +1-520-621-6223; fax: +1-520-621-8450.
E-mail address: dhouser@econlab.arizona.edu (D. Houser).

0304-4076/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0304-4076(02)00205-1

wealth elasticities of labor supply, and also allows inferences with respect to the way the incentive to accumulate human capital affects labor supply over the life cycle.

It is easily seen that the scope of the human capital effect can be restricted by the expectations process. For instance, if it is assumed that agents are not at all forward looking, then, since human capital's value lies in its future dividend stream, there can be no effect beyond any immediate impact on wages. More generally, joining particular expectations processes to many common modeling assumptions (such as expected utility maximization) can leave certain types of behaviors difficult to explain.[1] For instance, it is well known that young men in their late teens, who have completed their schooling, often choose to supply relatively little market labor (see, e.g., Card, 1994). At the same time, at least since Mincer (1974), market experience has consistently been found to be a significant and economically important determinant of wages. Rational agents should accurately forecast the return to market experience and, consequently, have the incentive to provide high labor supply early in the life-cycle.[2]

To reconcile the empirical facts with rational expectations, which is a common assumption in human capital labor supply studies (see, e.g., Eckstein and Wolpin, 1989; Shaw, 1989, Keane and Wolpin, 1997, or Altug and Miller, 1998), it is often posited that transition costs or complicated preference effects play an important role in labor supply outcomes (see, e.g., Keane and Wolpin, 1997). An alternative explanation is that people come to understand the value of human capital slowly over time. If so, then this would suggest that older men, rather than younger men, might be more influenced by the human capital accumulation incentive. A goal of this paper is to characterize the effect of this incentive on life cycle labor supply in a way that is less dependent on assumptions about the way people form expectations.

Strong assumptions about the expectations process might also lead to biased estimates of a model's structural parameters through model misspecification. This might be of particular concern in the labor supply literature, where much effort has been directed towards generating point estimates of, for instance, the intertemporal substitution elasticity (see, e.g., MaCurdy, 1981; Browning et al., 1985 or Altonji, 1986), the return to education (see, e.g., Willis and Rosen, 1979 or Willis, 1986) and the return to different types of market experience (see, e.g., Heckman and Sedlacek, 1985 or Keane and Wolpin, 1997). Unfortunately, it is usually very difficult to determine whether the expectations process is misspecified, and, if so, whether the misspecification has economically meaningful consequences. The estimates of the wage and wealth elasticities that I present in this paper should not be biased by this particular specification error.

I relax assumptions over expectations by employing an econometric method suggested by Geweke and Keane (1999). This method can be used to perform inference in a wide class of dynamic models, and, in addition to being robust to the expectations mechanism, has several advantages. First, with it one can estimate structural

[1] Epstein and Zin (1989) point out that in many models that assume rational expectations and expected utility maximization, it is impossible that agents are both highly risk averse and highly willing to substitute consumption intertemporally.

[2] In general of course, the particular labor supply profile will be sensitive to the model's parameterization (see, e.g., Blinder and Weiss, 1976).

specifications that might otherwise be intractable. The model developed below, for instance, includes serially correlated log-wage equation productivity shocks. While there is substantial empirical evidence supporting this specification (see, e.g., MaCurdy, 1982), it has not appeared in previous structural estimations due to the computational difficulties it entails (see, e.g., Keane and Wolpin, 1994). Second, the data requirements are not heavy. Inference is possible using only a single cohort of agents observed over only part of their life-cycle. In addition, with this econometric methodology one can draw inferences about agents' decision rules. With decision rules in hand the model can be simulated without computational burden, allowing one to examine wage, wealth and other labor supply effects easily.

Geweke and Keane (1999) provide Monte Carlo evidence that their procedure recovers structural parameters of dynamic models well when period return functions are correctly specified (see also Geweke et al., 2001). However, as detailed below, not all of the parameters that enter their procedure have immediate structural interpretations. Moreover, in the usual case where period return functions are not known, I point out below that the distinction between structural and other parameters can be arbitrary. In this research, my approach is to adopt a variant of the commonly used trans-log specification for the period return function, and interpret the associated parameters as "structural". Still, policy analysis based on my structural estimates should be approached with caution.

I apply the Geweke–Keane method to a life-cycle labor supply model that includes learning-by-doing human capital accumulation and savings of physical capital. Each period, agents choose either not to work, or to work part-time, full-time or overtime, and they choose levels for consumption and savings. I estimate this model using a panel drawn from the National Longitudinal Survey of Youth. Although classical estimation of the model might be possible, I perform Bayesian inference under diffuse priors. To implement the Bayesian estimation I use Markov Chain–Monte Carlo methods.

The results are largely consistent with standard results in the labor supply literature. The marginal posterior distributions for preference and log-wage equation parameters are in all cases economically reasonable. Although inference about consumption effects is quite imprecise, Bayesian point estimates suggest that consumption and leisure are complements for young men, and that consumption plays a smaller role than leisure in determining the utility of different hours alternatives. Education and market experience are each found to have a plausible positive influence on wages, and I find substantial serial correlation in the log-wage equation's error process. In addition, as reported by Altug and Miller (1998) in a study of female labor supply, my findings provide evidence that recent work experience has a stronger effect on wages than other work experience.

Simulations of the estimated model shed light on the contributions of human capital, wages and wealth to labor supply decisions. Consistent with usual findings, wages and wealth seem to contribute very little to labor supply variation. The incentive to invest in human capital is found to have a relatively more important effect on hours decisions, particularly for men in their late 20's. My results suggest that reducing the expected benefit of human capital accumulation might both shift down and flatten the age-hours profile.

This paper is in seven sections. Section 2 describes the model, Section 3 details the method of Bayesian inference and Section 4 describes the data. Section 5 reports marginal posterior distributions and the model's fit. Section 6 investigates the determinants of the labor supply of young men, and Section 7 concludes.

2. The model

2.1. Endowments, technology, preferences and timing

Each of N agents lives for $T < \infty$ periods. There is an initial period wealth endowment A_{n1}. At each age t each agent n receives a wage offer and nonlabor income. I assume that the wage offer, denoted w_{nt}, depends on his age t, market experience x_{nt}, number of hours worked in the previous year $h_{nj,t-1}$, education E_{nt} and an idiosyncratic productivity shock ε_{nt}, through the log-wage equation:

$$\log w_{nt} = \beta_0 + \beta_1 x_{nt} + \beta_2 h_{nj,t-1} + \beta_3 t + \beta_4 t^2 + \beta_5 E_{nt} + \varepsilon_{nt}, \tag{1}$$

where lagged hours in the initial period are zero. I include lagged hours because of the evidence that recent work experience has a greater effect on wages than other work experience (see, e.g., Altug and Miller, 1998; Miller and Sanders, 1997). In addition, Altug and Miller (1990, 1998), among others, provide evidence that aggregate changes in the supply and demand for labor can have significant effects on wages. Not controlling for this could confound inferences about, in particular, human capital effects. Hence, I also include year dummies in the wage equation.

I denote the log-wage regressor matrix by X_{nt}, so $\log w_{nt} = X_{nt}'\beta + \varepsilon_{nt}$. Skills may vary across individuals, and to capture this heterogeneity I allow the idiosyncratic productivity shock to be serially correlated.[3] I assume that

$$\varepsilon_{nt} = \rho\varepsilon_{n,t-1} + u_{nt} \quad (t > 1), \tag{2}$$

$$\varepsilon_{n1} \sim F_{\varepsilon}, \tag{3}$$

where the disturbance u_{nt} is independently and identically distributed over time and individuals, and the distribution of the initial shock F_{ε} will be specified below.

I assume that wealth, A_{nt}, evolves as follows:

$$A_{nt} = R_{y(n,t-1)} a_{n,t-1} + \alpha_{nt} \quad (t > 1). \tag{4}$$

Here, $a_{n,t-1}$ is the end-of-period wealth remaining after consumption expenditure and labor earnings in period $t-1$ and α_{nt} is nonlabor income that is independently and identically distributed across agents and time. The return on savings $R_{y(n,t-1)}$ depends on the year, $y(n,t-1)$, that he is age $t-1$. I assume these returns are both known and exogenous to every agent.[4] A_{n1} is taken as given.

[3] There are many ways to account for unobserved skill heterogeneity. One possibility is to assume there are a finite number of agent types, and that types differ according to their log-wage equation intercept (see, e.g., Keane and Wolpin, 1997). My specification has the advantage that a person's position in the cross-section heterogeneity distribution can evolve with age.

[4] Note that α_{nt} can be viewed as incorporating a stochastic return to savings.

After receiving the wage offer, return on savings and nonlabor income, individuals form their labor supply and consumption decisions. They choose either to work part-time, full-time or overtime, or not to work. Denote by h_{njt} the fraction of full-time hours agent n works in period t if alternative j is chosen, and assign:

$$h_{n1t} = 0.5, \quad h_{n2t} = 1.0, \quad h_{n3t} = 1.5, \quad h_{n4t} = 0.0 \tag{5}$$

for any n and t, so that alternative one corresponds to part-time hours, alternative two to full-time hours, alternative three to overtime hours and alternative four to zero market hours. Because h_{njt} depends only on j, and does not vary with n or t, when no confusion can arise I will drop the n and t subscripts. With this notation, the consumption and labor supply decisions are related to wealth and savings through the budget constraint

$$c_{njt} + a_{njt} = A_{nt} + h_j w_{nt}, \tag{6}$$

where c_{njt} is n's date t level of consumption if they choose hours alternative j.

Preferences over period consumption and leisure allocations are ordered according to the possibly stochastic function $U_t(c_{njt}, j, j^*_{n,t-1}, \zeta_{njt})$, where $j^*_{n,t-1}$ denotes the actual labor supply decision of agent n at period $t-1$, and ζ_{njt} is a preference shock.[5] The value of $j^*_{n,0}$ will be taken as given.

Finally, I assume $x_{n1} = 0$, and individual n with experience x_{nt} at age t who chooses alternative j will of course have age $t+1$ experience given by $x_{n,t+1} = x_{nt} + h_j$.

2.2. The state space

I denote the state space by S, with typical element $s \in S$. If $t > 1$, then an individual's state at the time of his labor supply and saving decisions is his current age t, current wealth A_{nt}, current interest rate $R_{y(n,t)}$, cumulative labor market experience x_{nt}, the period t realization of each random variable, last period's productivity shock realization $\varepsilon_{n,t-1}$, his level of educational attainment E_n, and his previous period's work decision, $j^*_{n,t-1}$. Consequently, the state vector for young man n at age $t > 1$ is

$$s_{nt} = \{A_{nt}, t, x_{nt}, R_{y(n,t)}, E_n, j^*_{n,t-1}, \varepsilon_{n,t-1}, \varepsilon_{nt}, \alpha_{nt}, \{\zeta_{njt}\}_{j=1,\dots,4}\}. \tag{7}$$

At $t = 1$ the state vector is

$$s_{n1} = \{A_{n1}, t = 1, x_{n1} = 0, R_{y(n,1)}, E_n, j^*_{n0}, \varepsilon_{n1}, \{\zeta_{nj1}\}_{j=1,\dots,4}\}.$$

2.3. Valuation of alternatives

Denote by $V_{njt}(s_{nt})$ the value that person n places on hours alternative j in state s_{nt}. Since the environment is dynamic, the way expectations are modeled plays a fundamental role in the way values are assigned to alternatives. For instance, if one assumes

[5] Lagged hours decisions are included to allow for transition costs associated with hours changes across consecutive years. Past research has shown that such costs seem to play an important role in generating observed hours variation (see, e.g., Keane and Wolpin, 1997).

the model's agents have rational expectations, and that they discount future rewards at rate δ, then these values can be expressed with the usual Bellman formulation

$$V_{njt}(s_{nt})=\max_{a_{njt}}\left\{U_t(c_{njt},j,j^*_{n,t-1},\zeta_{njt})+\delta E\left[\max_{k\in\{1,2,3,4\}}V_{nk,t+1}(s_{n,t+1})|s_{nt},j,a_{njt}\right]\right\} \tag{8}$$

given the laws of motion and boundary conditions

$$c_{njt}=A_{nt}+h_j w_{nt}-a_{njt} \quad (t=1,\ldots,T), \tag{9}$$

$$A_{nt}=R_{y(n,t-1)}a_{nj^*_{n,t-1},t-1}+\alpha_{nt} \quad (t=2,\ldots,T), \tag{10}$$

$$x_{n,t+1}=x_{nt}+h_j \quad (t=1,\ldots,T-1), \tag{11}$$

$$\log w_{nt}=X'_{nt}\beta+\varepsilon_{nt} \quad (t=1,\ldots,T), \tag{12}$$

$$V_{n,T+1}(s_{n,T+1})=B(s_{n,T+1}), \tag{13}$$

$$A_{n1}\ \text{given}, \quad x_{n1}=0,\ j^*_{n,0}\ \text{given} \tag{14}$$

and given the distributions for all of the stochastic variables. In general, the terminal values $B(s_{n,T+1})$ may be nonzero, to reflect, for instance, bequest motives. The conditional expectation is with taken respect to the true distribution of the state vector $s_{n,t+1}$.

Since little is known about the way people actually form expectations, I noted earlier that several difficulties could stem from positing rational expectations. In fact, even if rational expectations is the correct specification, explicitly imposing it in this environment could cause difficulties. In particular, empirically implementing (8)–(14) with the "nested maximum likelihood" algorithm (see, e.g., Wolpin, 1984, or Rust, 1987), would be extremely computationally burdensome. The reason is that this technique requires solving dynamic programming problem (8) at many trial values of the model's parameter vector. Since this model's state-space is large, the solution procedure would be tremendously time consuming (except in very special cases, such as when the payoff functions are assumed to be linear quadratic.)[6]

There are several alternatives to the nested maximum likelihood approach to implementing dynamic selection models empirically. These include the well known suggestions of Hotz and Miller (1993) and Manski (1993). The idea behind each of these methods is to use data to learn about the values of the expectations on the right-hand side of (8), and then use this information as input to a procedure that generates estimates of the structural parameters. A limitation of these approaches is that, in order to

[6] Assuming that wealth is discretized, solving the dynamic programming problem requires a high order integration over the latent state variables at every combination of (t,x,E,d_{-1},A) in the state-space. The dimension of the integration at every age-t state-point is $t+5$, corresponding to the preference, wealth and productivity contemporaneous shocks, as well as all lagged productivity shocks due to the serial correlation of that disturbance. See Keane (1994) for additional discussion of the burdens serially correlated latent variables can cause in solving and estimating selection models.

learn about expectations, each requires the data to satisfy a strict form of stationarity in order to rule out cohort effects.

In this paper I empirically implement an alternative specification that is based on a suggestion by Geweke and Keane (1999). The specification I use is tractable in the present environment, does not have heavy data requirements and does not require strong assumptions about individuals' expectation mechanisms. It requires only that individuals follow a state and choice contingent savings decision rule, $a_{njt} \equiv a(s_{nt}, j)$, and that, in any period, an alternative's total value is additively separable between a contemporaneous payoff and a "future" component, denoted respectively by $U_t(c_{njt}, j, j^*_{n,t-1}, \zeta_{njt})$ and $F(s_{nt}, j, a_{njt})$. Under these conditions, individuals assign values to alternatives as follows.

$$V_{njt}(s_{nt}) = U_t(c_{njt}, j, j^*_{n,t-1}, \zeta_{njt}) + F(s_{nt}, j, a_{njt}) \quad \text{s.t.} \quad (9)\text{–}(14) \tag{15}$$

The selection rule is that agents choose alternative j^* if and only if

$$V_{nj^*t}(s_{nt}) > V_{njt}(s_{nt}) \quad (\forall j \neq j^*).$$

Note that (15) is equivalent to (8) if $F(\cdot) = \delta E[\max_{k \in \{1,2,3,4\}} V_{nk,t+1}(s_{n,t+1}) | s_{nt}, j, a_{njt}]$ and a_{njt} is the associated optimal savings rule (assuming such a rule exists). Of course, (15) is also consistent with a variety of other expectation mechanisms. For instance, if agents are not at all forward looking when making decisions, then $F(\cdot)$ and $a(s_{nt}, j)$ may each be identically zero. In general, the shapes of $F(\cdot)$ and $a(s_{nt}, j)$ will vary with the way people form expectations. Since the way people form expectations is not generally known, neither are the forms of $F(\cdot)$ and $a(s_{nt}, j)$. Hence, I replace them with parametric flexible functional forms (polynomials) in the model's state variables. Then, I estimate the parameters that characterize the flexible functional forms jointly with the model's structural parameters.

After substituting the flexible functional form for the future component, the model bears some resemblance to the model (Hotz and Miller, 1988) used to study female life cycle labor supply and fertility. In that paper, their approach is to specify and estimate parametric index functions that have as arguments the model's state variables, and that can consequently be interpreted as approximations to some underlying structural model's decision rules. Their approach is robust with respect to both the nature of expectations and period return functions, in the sense that the underlying individual decision problem is not, and need not be, specified in their procedure.

The present model is also related to a static (Roy, 1951) selection model augmented to include nonpecuniary effects on decisions (as in Heckman and Sedlacek, 1986). However, the fact that $F(\cdot)$ in (15) is common across alternatives implies that the parameters associated with the nonpecuniary effects are held fixed across alternatives, and this restriction is not typically invoked in the estimation of static selection models. In addition, it turns out that the links between choices and states implied by the laws of motion of the state variables suggest restrictions on the future component's arguments.[7]

[7] See Geweke et al. (2001) for additional discussion on this point.

It is important to emphasize that, as a practical matter, the distinction between the parameters that characterize U_t and those that characterize F can be arbitrary. This difficulty has at its root the identification problem pointed out by Rust (1994), which applies to virtually all dynamic models that economists study and, in the present context, implies that it is not generally possible to nonparametrically separately identify U_t and F. Nevertheless, Geweke and Keane (1999) present results from a Monte Carlo study that demonstrates the parameters of U_t can be recovered well with their approach when U_t is correctly specified (see also the Monte Carlo results reported in Geweke et al., 2001). Because I cannot be sure that U_t has been correctly specified in the present case, the structural estimates that I report should be viewed with caution. In particular, one should be aware that policy analysis based on my structural estimates might be misleading. [8]

3. Bayesian inference

3.1. Specification of functional forms

This section specifies parametric forms for the functions that appear in (15) and (9)–(14). The wage equation is specified in (1), and will not be repeated here. Also, since the method of inference does not require me to solve the dynamic programming problem, it is not necessary to specify the state-contingent terminal values $B(\cdot)$ that appear in (13). If desired, inference about these values could be derived from the estimated future component. This leaves to be specified the utility function, future component and savings rule.

3.1.1. Utility function

The total amount of time available to the agent is set to unity. Accordingly, assuming there are 112 nonsleep hours available per week, an agent who works full-time spends 36% of their nonsleep hours in market work. I use a flexible variation of translog preferences that excludes a linear leisure term. This is an identifying restriction, as it would be perfectly collinear with alternative specific intercepts that appear in the future component. The utility function is:

$$U_t^*(c_j, j, j_{-1}, \zeta_j) = \theta_0^*(j, j_{-1}, t) + \theta_1^* \log(c_j) + \theta_2^* \log(c_j)^2 + \theta_3^* t \log(1 - 0.36h_j) + \theta_4^* \log(c_j) \log(1 - 0.36h_j) + \theta_5^* t \log(1 - 0.36h_j)^2 + \theta_6^* t^2 \log(1 - 0.36h_j) + \zeta_j \quad (16)$$

[8] Note that one could abandon the structural interpretation entirely and, as in Hotz et al. (1988), view the procedure as estimating reduced form decision rules.

Here, $\theta_0^*(j, j_{-1}, t)$ accounts for nonstochastic preference effects not easily expressed in terms of log leisure and log consumption, as follows.

$$\theta_0^*(j, j_{-1}, t) = C_1^* \chi(t > 1 \text{ and } j \neq j_{-1} \text{ and } j \neq 4) + C_2^* \chi(t = 1 \text{ and } j \neq 4) + \phi_1^* h_j(t+1) + \phi_2^* \chi_{j=1\,\&\,t=1} + \phi_3^* \chi_{j=2\,\&\,t=1}. \quad (17)$$

Here, χ denotes an indicator function that takes the value one if the stated condition is true, and zero otherwise, and note that $t = \text{age} - 16$. C_1^* and C_2^* are transition costs to choosing different hours alternatives across consecutive periods, although I assume there is no cost to choosing not to work. Note that C_2^* is an age 17 transition cost, and I also include age 17 preference effects in the terms ϕ_2^*, ϕ_3^* (the analogous effect on the preference for overtime is excluded as for identification purposes). The reason for these effects is that at age 17 young men are likely living at home, and this may influence their ranking of alternatives. Finally, regressors in h_j and th_j allow for additional hours effects. Since linear hours terms are not identified, I impose the restriction that the coefficients of these terms are identical, and include the single regressor $h_j(t+1)$.

3.1.2. The future component

While people are not restricted to a particular expectation mechanism, it is useful to specify the arguments of the future component in a way that is consistent with the laws of motion of the underlying behavioral model. For example, contemporaneous realizations of serially independent stochastic variables, such as wealth and preference shocks, contain no information relevant for forecasting future outcomes, and so should not enter the arguments of the future component's flexible functional form. Also, since the preference shocks are serially uncorrelated, their contemporaneous realizations hold no information useful for predicting future outcomes and so they should not enter the future component's regressors.

The model implies $F(s_{nt}, j, a_{njt})$ has arguments in period $t's$ choice (since that is period $(t+1)'s$ lagged choice, which is payoff relevant) and the terms $(t+1), x_{nt}+h_j, E_{nt}$, $R_{y(n,t)}a_j$, and $\rho\varepsilon_{nt}$. The interpretation is that peoples' predictions about the effect their current choice will have on their future outcomes depend on the state they expect to realize the next period, conditional on their current state and choice.[9]

In order to accommodate a wide class of expectation mechanisms I choose to model the future component as a high order polynomial, as follows (the n and t subscripts are suppressed for clarity).

$$F^*(s, j, a_j; \pi^*) = \pi_0^* \chi_{j=1} + \pi_1^* \chi_{j=2} + \pi_2^* \chi_{j=3} + \pi_3^* \chi_{j=4} + \pi_4^*(t+1) + \pi_5^* R_{y(n,t)} a_j + \pi_6^*(x + h_j) + \pi_7^* \rho\varepsilon + \pi_8^* E + \pi_9^*(t+1)^2 + \pi_{10}^*(R_{y(n,t)} a_j)^2 + \pi_{11}^*(x + h_j)^2 + \pi_{12}^*(\rho\varepsilon)^2 + \pi_{13}^* E^2 + \pi_{14}^*(t+1) R_{y(n,t)} a_j + \pi_{15}^*(x + h_j)(t+1)$$

[9] One way to think about this is that I impose the restriction that the information relevant to agents is consistent with the model, but then do not impose strong restrictions on the way they use this information to form forecasts.

$$
\begin{aligned}
&+\pi^*_{16}(t+1)\rho\varepsilon+\pi^*_{17}(t+1)E+\pi^*_{18}Ra_j(x+h_j)\\
&+\pi^*_{19}R_{y(n,t)}\rho\varepsilon a_j+\pi^*_{20}R_{y(n,t)}a_jE+\pi^*_{21}(x+h_j)\rho\varepsilon+\pi^*_{22}(x+h_j)E\\
&+\pi^*_{23}\rho\varepsilon E+\pi^*_{24}(t+1)^3+\pi^*_{25}(x+h_j)^3+\pi^*_{26}E^3\\
&+\pi^*_{27}(t+1)^2(x+h_j)+\pi_{28}(t+1)^2E+\pi^*_{29}(t+1)(x+h_j)^2\\
&+\pi^*_{30}(t+1)E^2+\pi^*_{31}(x+h_j)^2E+\pi^*_{32}(x+h_j)E^2\\
&+\pi^*_{33}(x+h_j)(t+1)E. \qquad (18)
\end{aligned}
$$

Here, χ is an indicator function that takes value one if the stated condition is true, and zero otherwise. I assume that agents' expectations are such that their future component lies along this flexible functional form.

3.1.3. The savings rule

I assume the savings decision rule is a flexible functional form in current wealth A, current age t, market experience x, level of education E and offered wage w,

$$a(s,j|\gamma)=\gamma_1+\gamma_2h_jw_t+\gamma_3E+\gamma_4(x_t+h_j)+\gamma_5t+\gamma_6A_t \qquad (19)$$

so long as this level of savings is consistent with positive consumption, and does not exceed greater than one-half of the total resources held by the young man. If (19) leads to nonpositive consumption then consumption's value is set to 100 dollars, and savings accordingly defined. If (19) exceeds half of total financial resources (the sum of wealth and current income) then savings is set to one-half of total resources, and consumption defined accordingly.

3.2. Distributional assumptions

The stochastic nonlabor income α, the identically and independently distributed wage innovations u, the first period wage shock ε_1, and the taste for leisure shocks ζ have distributions:

$$\alpha_{nt}\sim \mathrm{N}(0,\sigma^2_\alpha), \qquad (20)$$

$$u_{nt}\sim \mathrm{N}(0,\sigma^2_u), \qquad (21)$$

$$\varepsilon_{n1}\sim \mathrm{N}(0,\sigma^2_u/(1-\rho^2)), \qquad (22)$$

$$\{\zeta_{njt}\}_{j=1,4}\sim \mathrm{N}(0,\Sigma_\zeta), \qquad (23)$$

where each σ^2 denotes a variance, and Σ_ζ is a symmetric, positive definite matrix.

3.3. Identification

Labor supply decisions depend on relative alternative valuations. Consequently, the model is not identified in levels. Identification is achieved in the usual way by working with the following differenced system.

$$z_{njt}(s_{nt})=\tilde V_j(s_{nt})-\tilde V_4(s_{nt}) \quad j\in\{1,2,3\}, \qquad (24)$$

where $\tilde{V}_j(s_{nt}) = V_j(s_{nt})/(\Sigma_\varsigma(1,1)+\ \Sigma_\varsigma(4,4) - 2\Sigma_\varsigma(1,4))^{1/2}$. Now define

$$\tilde{\sigma}^2 = (\Sigma_\varsigma(1,1) + \Sigma_\varsigma(4,4) - 2\Sigma_\varsigma(1,4)). \tag{25}$$

Then, suppressing n and t subscripts, the corresponding utility function differences are, for $j = 1,2,3,$

$$\begin{aligned}
&(U_t^*(c_j, j, j_{-1}, \zeta_j) - U_t^*(c_4, 4, j_{-1}, \zeta_4))\tilde{\sigma}^{-1} \\
&\quad = \theta_0(j) - \theta_0(4) + \theta_1(\log(c_j) - \log(c_4)) + \theta_2(\log(c_j)^2 - \log(c_4)^2) \\
&\qquad + \theta_3 t \log(1 - 0.36h_j) + \theta_4 \log(c_j)\log(1 - 0.36h_j) \\
&\qquad + \theta_5 t \log(1 - 0.36h_j)^2 + \theta_6 t^2 \log(1 - 0.36h_j) + \eta_j,
\end{aligned} \tag{26}$$

where $\theta_i = \theta_i^* \tilde{\sigma}^{-1}$, $\eta_j = (\zeta_j - \zeta_4)\tilde{\sigma}^{-1}$, $\{\eta_j\} \sim \mathrm{N}(0, \Sigma_\eta)$, $var(\eta_1) = 1$, and

$$\theta_0(j) - \theta_0(4) = C_1\chi(\cdot) + C_2\chi(\cdot) + \phi_1 h_j(t+1) + \phi_2\chi_{j=1\,\&\,t=1} + \phi_3\chi_{j=2\,\&\,t=1}. \tag{27}$$

Here, $C_i = C_i^* \tilde{\sigma}^{-1}$ $(i = 1,2)$, and $\phi_i = \phi_i^* \tilde{\sigma}^{-1}$, for $i = 1,2,3$.

Nobody in my sample chooses overtime hours at age 17. Consequently, ϕ_2, ϕ_3 and C_2 are not globally separately identified by the available data. (C_2 needs to be sufficiently large to ensure nobody chooses overtime, and then ϕ_2 and ϕ_3 can be chosen to ensure the appropriate fraction chooses each of the three other alternatives.) I achieve identification by imposing a diffuse but proper prior on the transition cost C_2.

For notational ease, denote the differenced utility function's regressor structure by Ξ, and the entire set of utility function parameters by θ, so that

$$(U_t^*(c_{njt}, j, j_{n,t-1}, \zeta_{njt}) - U_t^*(c_{n4t}, 4, j_{n,t-1}, \zeta_{n4t}))\tilde{\sigma}^{-1} \equiv \Xi'_{njt}\theta + \eta_{njt}. \tag{28}$$

Next define

$$F_j(s_{nt}) \equiv (F^*(s_{nt}, j, a_{njt}) - F^*(s_{nt}, 4, a_{n4t}))\tilde{\sigma}^{-1}.$$

The arguments that remain in the differenced future components are described in Appendix A. Denote by π the coefficients of the differenced future component's regressors, and let Ψ be the associated regressor matrix (composed of polynomials in the state-variables,) so that

$$F_j(s_{nt}) = \Psi'_{njt}\pi \quad (j = 1,2,3). \tag{29}$$

It is worthwhile to emphasize again that the structure of Ψ is derived from the laws of motion, and that this feature distinguishes this framework from the usual selection model.

Using the above notation, we write (24) as

$$\begin{aligned}
z_{njt} &= \Xi'_{njt}\theta + \Psi'_{njt}\pi + \eta_{njt} \\
&= Q'_{njt}\Lambda + \eta_{njt},
\end{aligned} \tag{30}$$

where $Q_{njt} = [\Xi'_{njt}, \Psi'_{njt}]'$, and $\Lambda = [\theta', \pi']'$.

It is important to conclude this section with a comment on measurement error in the wage and wealth data. Measurement error variance is identified off the fact that measurement error realizations do not affect the actual wage but only affect the wage observed by the econometrician. Hence, measurement error does not affect decisions, while shocks to actual wages and wealth influence both decisions and observed wages. However, since it is consistently found that wages and wealth have a very small effect on choices, measurement error identification typically requires a substantial amount of data. The data set I use in this study, as noted in Section 4, is rather small and includes many missing observations (about half of the wealth data is missing). Consequently, measurement error variances are only very weakly separately identified from the variances of the structural disturbances. Accordingly, this paper presents results for the case when measurement error variance is set to zero. Failure of this assumption could lead to downward biased estimates of the wage and wealth substitution elasticities, and a masking of the effect of market experience, in particular, on wages.

3.4. The joint posterior distribution

Before developing the desired joint posterior distribution, upon which Bayesian inference is based, first substitute out consumption using (6) the budget constraint. Terms in $\log(c_{njt})$ then become terms in $\log(w_{nt}h_j + A_{nt} - a(s_{nt}, j|\gamma))$. Since these terms enter the Ξ regressor structure, it follows that Q_{njt} depends on γ in a nonlinear way. In addition, Q_{njt} depends on functions of wages and wealth. It is worthwhile to emphasize these facts as they play important roles in the development of the likelihood and the subsequent discussion of the Gibbs sampling algorithm. However, to reduce notational burden, I will suppress any direct indication of the dependence of Q_{njt} on functions of wages and wealth and the parameters γ. Also for notational ease, below I write R_{nt} instead of $R_{y(n,t)}$.

Next, it is convenient to work with precision matrices. Define these by[10]

$$h_u = \frac{1}{\sigma_u^2}, \quad h_\alpha = \frac{1}{\sigma_\alpha^2}, \quad H_\eta = \Sigma_\eta^{-1}.$$

Finally, define the choice indicator

$$d_{njt} = \begin{pmatrix} 1 & \text{if } j \text{ was chosen by } n \text{ at age } t \\ 0 & \text{otherwise} \end{pmatrix}.$$

The posterior distribution of interest is the joint distribution of the parameters, latent variables and unobserved variables conditional on all of the observed quantities. Let A^u indicate the vector of all unobserved wealth values, and let A^o denote the vector of all observed quantities. Define the analogous vectors for unobserved and observed wages. Then the joint posterior density of interest can be expressed as

$$p(A^u, W^u, \{\{z_{njt}\}_{j=1,2,3}\}_{nt}, \beta, \rho, \Lambda, \gamma, h_u, h_\alpha, H_\eta | A^o, W^o, \{x_{nt}, t, E_n, (d_{njt})_{j=1,2,3}\}_{nt}). \tag{31}$$

[10] Because the notation is standard, and because no confusion should arise, for the remainder of this section and the next I use h and H to denote precisions instead of market hours.

Using Bayes' theorem, it is easy to show that this density is proportional to

$$p(\{A_{nt}, w_{nt}, \{z_{njt}\}_{j=1,2,3}\}_{nt} | \beta, \rho, \Lambda, \gamma, h_u, h_\alpha, H_\eta, \{x_{nt}, t, E_n, (d_{njt})_{j=1,2,3}\}_{nt}) \times p(\beta, \rho, \Lambda, \gamma, h_u, h_\alpha, H_\eta). \quad (32)$$

To develop this joint posterior, consider the first distribution in the product that defines (32). Recalling that I treat the initial wealth observation as given, this distribution can be factored into the product of the following conditional distributions.

$$\begin{aligned} &p(\{A_{n,t}, w_{nt}, \{z_{nkt}\}_{k=1,2,3}\}_{nt} | \cdot) \\ &= p(\{\{z_{nkt}\}_{k=1,2,3}\}_{nt} | \{w_{nt}, A_{nt}, \cdot\}_{nt}) p(\{A_{n,t>1}\}_{nt} | \{w_{nt}\}_{nt}, \{A_{n1}\}_n, \cdot) \\ &\quad \times p(\{w_{nt}\}_{nt} | \cdot). \end{aligned} \quad (33)$$

Using the distributional assumptions made above, the first conditional distribution on the right-hand side is

$$\begin{aligned} &p(\{z_{njt}\}_{j=1,2,3} | \{w_{nt}, A_{nt}, \cdot\}) \\ &\propto \prod_{n,t} |H_\eta|^{1/2} \exp\left\{ -\frac{1}{2} \begin{pmatrix} \{z_{n1t} - Q'_{n1t}\Lambda\} \\ \{z_{n2t} - Q'_{n2t}\Lambda\} \\ \{z_{n3t} - Q'_{n3t}\Lambda\} \end{pmatrix}' H_\eta \begin{pmatrix} \{z_{n1t} - Q'_{n1t}\Lambda\} \\ \{z_{n2t} - Q'_{n2t}\Lambda\} \\ \{z_{n3t} - Q'_{n3t}\Lambda\} \end{pmatrix} \right\} \\ &\quad \times I(\{z_{njt}, d_{njt}\}_{j=1,2,3}), \end{aligned} \quad (34)$$

where

$$I(\{z_{njt}, d_{njt}\}_{j=1,2,3}) = \left\{ \begin{array}{c} 1 \text{ if } d_{njt} = 1 \text{ \& } z_{njt} > 0 \text{ \& } (z_{njt} > z_{nkt})(j \neq k) \\ \text{or} \\ 1 \text{ if } (\forall j)(d_{njt} = 0 \text{ \& } (z_{njt} < 0)) \\ \text{or} \\ 0 \text{ otherwise} \end{array} \right\}. \quad (35)$$

The indicator function takes value zero unless the values for all z_j are consistent with observed choices. If a market alternative k is chosen, z_k must be positive and greater than the other values $\{z_j\}_{j \neq k}$, while if 'home' is chosen all z_j must be negative.

The distribution of wealth given wages and choices is:

$$p(\{A_{n,t>1}\} | \{w_{nt}, \cdot\}) \propto \prod_{n,t>1} h_\alpha^{1/2} \exp\left\{ -\frac{h_\alpha}{2} (A_{nt} - R_{n,t-1} a(n, j^*_{t-1}, t-1 | \gamma))^2 \right\}, \quad (36)$$

where $a(n, j^*_{t-1}, t-1)$ denotes the level of savings by person n who chose alternative j^*_{t-1} in period $t-1$, and R is the real gross return on that savings.

Finally, the marginal distribution of wages is simply

$$p(\{w_{nt}\}|\{\cdot\}) \propto (h_u(1-\rho^2))^{N/2} \prod_n \frac{1}{w_{n1}} \exp\left[\frac{-h_u(1-\rho^2)}{2}(\log w_{n1} - X'_{n1}\beta)^2\right]$$
$$\times \prod_{n,t\geqslant 2} h_u^{1/2} \frac{1}{w_{nt}} \exp\left\{-\frac{h_u}{2}(\log w_{nt} - \rho \log w_{nt-1} - X'_{nt}\beta + \rho X'_{nt-1}\beta)^2\right\}.$$

Next, consider the second term in expression (32), which is the prior distribution of the model's parameters. I assume flat and unbounded priors, except as follows. I adopt uninformative priors $p(h_u) \propto h_u^{-1}$, $p(h_\alpha) \propto h_\alpha^{-1}$ and $p(H_\eta) \propto |H_\eta|^{-2}$ (see, e.g., Zellner, 1971, Section 8.1). Priors for the savings function's parameters are $\gamma_1 \sim U[0, 4000]$, $\gamma_2 \sim U[-0.25, 0.25]$, $\gamma_j \sim U[-500, 500]$ for $j = 3, 4, 5$, and $\gamma_6 \sim U[-0.6, 0.6]$. Finally, I assume ρ is uniformly distributed over $[-1, 1]$. Use of these boundary restrictions seemed to improve the performance of the numerical algorithm, and Table 10 shows that the marginal posterior distributions generally lie well within the prior. Finally, as an identifying assumption, the age 17 transition cost C_2 is posited to be distributed normally with a mean of zero and a standard deviation of 1000.

Hence, the joint posterior distribution of interest (31) is proportional to

$$|H_\eta|^{-2} h_\alpha^{-1} h_u^{-1} I_\gamma I_\rho \exp\{-(C_2/1000)^2/2\}$$
$$\times \prod_{n,t} |H_\eta|^{1/2} \exp\left\{-\frac{1}{2}\begin{pmatrix}\{z_{n1t} - Q'_{n1t}\Lambda\}\\ \{z_{n2t} - Q'_{n2t}\Lambda\}\\ \{z_{n3t} - Q'_{n3t}\Lambda\}\end{pmatrix}' H_\eta \begin{pmatrix}\{z_{n1t} - Q'_{n1t}\Lambda\}\\ \{z_{n2t} - Q'_{n2t}\Lambda\}\\ \{z_{n3t} - Q'_{n3t}\Lambda\}\end{pmatrix}\right\} I(\{z_{njt}, d_{njt}\}_j)$$
$$\times \prod_{n,t>1} h_\alpha^{1/2} \exp\left\{-\frac{h_\alpha}{2}(A_{nt} - R_{n,t-1}a(n, j^*_{t-1}, t-1|\gamma))^2\right\}$$
$$\times (h_u(1-\rho^2))^{N/2} \prod_n \frac{1}{w_{n1}} \exp\left[\frac{-h_u(1-\rho^2)}{2}(\log w_{n1} - X'_{n1}\beta)^2\right]$$
$$\times \prod_{n,t\geqslant 2} h_u^{1/2} \frac{1}{w_{nt}} \exp\left\{-\frac{h_u}{2}(\log w_{nt} - \rho \log w_{nt-1} - X'_{nt}\beta + \rho X'_{nt-1}\beta)^2\right\}. \quad (37)$$

I show in Appendix B that this distribution is proper.

3.5. The Gibbs sampler

To construct the marginal posterior distributions of the model's parameters analytically is difficult. One reason is the joint posterior is nonstandard, and to construct marginal distributions it would be necessary to carry out integrations over all unobserved wage and wealth values. Instead, I choose to approximate the marginal posteriors using Gibbs sampling [11] with data augmentation (see, e.g., Tanner and Wong, 1987). Tierney (1994) proved that, under weak regularity conditions, draws from an appropriate set of conditional posterior distributions may be conducted so that they converge to the joint posterior distribution. It is easy to show that the regularity conditions hold in the present case.

3.5.1. The algorithm

My Gibbs sampling-data augmentation algorithm involves 10 steps. For notational convenience, I assume the sample consists of N agents, each of whom is observed over T periods. The Gibbs sampler proceeds as follows. [12]

1. Draw unobserved wages.
2. Draw unobserved wealth.
3. Draw h_α, the precision of the wealth shock.
4. Draw Γ, the coefficients of the savings function.
5. Draw β, the log-wage equation coefficients.
6. Draw ρ, the correlation coefficient.
7. Draw h_u, the precision of the log-wage equation shocks.
8. Draw z_{njt}, the relative alternative valuations.
9. Draw π and θ, the coefficients of the future component and utility function.
10. Draw H_η, the preference shocks' precision matrix.

3.5.1.1. The wage draw.

The distribution of w_{nt} $(T > t > 1)$, conditional on all of the other parameters and variables, is given by

$$\begin{aligned} p(w_{nt}) \propto \frac{1}{w_{nt}} \exp\left\{-\frac{h_u}{2}(\log w_{nt} - \rho \log w_{nt-1} - X'_{nt}\beta + \rho X'_{nt-1}\beta)^2\right\} \\ \times \exp\left\{-\frac{h_u}{2}(\log w_{n,t+1} - X'_{n,t+1}\beta - \rho \log w_{nt} + \rho X'_{nt}\beta)^2\right\} \\ \times \exp\left\{-\frac{h_\alpha}{2}(A_{n,t+1} - R_{nt}a_{n,j^*_{nt},t})^2\right\} \\ \times \exp\left\{-\frac{1}{2}(\{z_{njt} - Q'_{njt}\Lambda\}_{j=1,2,3})' H_\eta(\{z_{njt} - Q'_{njt}\Lambda\}_{j=1,2,3})\right\}. \end{aligned} \tag{38}$$

The first two lines derive from the marginal distribution of wages, in the third line wages enter the period t savings decision, and the fourth line includes functions of wages in Q_{njt}. Recall that Q_{njt} is defined in (30) and discussed further in the first paragraph of Section 3.4.

[11] A reader not interested in the computational details can skip directly to section four.

[12] I coded the Gibbs sampler software in FORTRAN 77, and used the IMSL numerical libraries extensively. The software is available from the author on request.

To obtain the distribution for w_{n1} requires replacing the first line of (38) with $1/w_{n1}\exp[-(h_u(1-\rho^2)/2)(\log w_{n1}-X'_{n1}\beta)]$. The distribution of w_{nT} omits the second and third lines of (38).

The wage distribution is nonstandard, since wages enter (38) in both levels and logs. To sample from the wage distribution I use rejection methods (see, e.g., Geweke (1992), Appendix A). For instance, when $1<t<T$, my approach is to draw a candidate wage $\tilde{w}_{nt}$ from the Gaussian source distribution implied by the first two lines in (38). The draw is accepted with probability proportional to the product of the last two lines in this expression.[13]

3.5.1.2. The wealth draw. Given all other parameters and variables, A_{nt} $(T>t>1)$ has distribution:

$$p(A_{nt}) \propto \exp\left\{-\frac{h_\alpha}{2}(A_{n,t+1}-R_{nt}a_{nj^*_{nt}t})^2\right\}\exp\left\{-\frac{h_\alpha}{2}(A_{nt}-R_{nt}a_{nj^*_{n,t-1}t-1})^2\right\}$$
$$\times\exp\left\{-\frac{1}{2}(\{z_{njt}-Q'_{njt}\Lambda\}_{j=1,2,3})'H_\eta(\{z_{njt}-Q'_{njt}\Lambda\}_{j=1,2,3})\right\} \tag{39}$$

A_{nt} enters the first exponential as an argument of $a_{nj^*_{nt}t}$, and the third exponential in Q_{njt}. The distribution of A_{n1} omits the second exponential of (39), and the distribution of A_{nT} omits the first exponential.

I sample from this distribution with rejection methods. For instance, when $1<t\leqslant T$, I use the source distribution $\exp\{-\frac{h_\alpha}{2}(A_{nt}-R_{nt}a_{nj^*_{n,t-1}t-1})^2\}$, and accept each draw with probability proportional to the product of the remaining two terms when evaluated at the candidate draw.[14]

3.5.1.3. The γ draw (savings function parameters). Given values for all the other parameters and variables, the distribution of γ is given by

$$p(\gamma)\propto\bar{p}(\gamma)\exp\left\{-\frac{1}{2}\begin{pmatrix}\{z_{n1t}-Q'_{n1t}\Lambda\}_{n,t}\\ \{z_{n2t}-Q'_{n2t}\Lambda\}_{n,t}\\ \{z_{n3t}-Q'_{n3t}\Lambda\}_{n,t}\end{pmatrix}'(H_\eta\otimes I_{NT})\begin{pmatrix}\{z_{n1t}-Q'_{n1t}\Lambda\}_{n,t}\\ \{z_{n2t}-Q'_{n2t}\Lambda\}_{n,t}\\ \{z_{n3t}-Q'_{n3t}\Lambda\}_{n,t}\end{pmatrix}\right\}$$
$$\times\exp\left\{-\frac{h_\alpha}{2}\sum_{n,t>1}(A_{nt}-R_{n,t-1}a(n,j^*_{t-1},t-1|\gamma))^2\right\}, \tag{40}$$

where $\bar{p}$ denotes the uniform prior, and the first exponential term is included since, as noted above, the savings policy function appears in the regressors of Q_{njt}.

To sample from this distribution again use a rejection procedure. I draw the elements of γ individually from their prior uniform distributions. Each draw is accepted with

[13] An additional restriction imposed on unobserved wages was that they be no greater than 100,000. This boundary restriction was useful for improving the performance of the Gibbs algorithm.

[14] I imposed the additional restriction that unobserved wealth values are no less than −5000 and no greater than 50,000. I found that this boundary restriction improved the performance of the numerical algorithm.

probability proportional to the right-hand side of (40) when evaluated at the candidate draw.

3.5.1.4. The h_α draw (precision of wealth shock). The conditional posterior for h_α is

$$p(h_\alpha) \propto h_\alpha^{(N(T-1)-2)/2} \exp\left\{-\frac{h_\alpha}{2}\sum_{n,t>1}(A_{nt}-R_{n,t-1}a_{nj^*_{n,t-1}t-1})^2\right\} \tag{41}$$

so that

$$h_\alpha \sum_{n,t>1}(A_{nt}-R_{n,t-1}a_{nj^*_{n,t-1}t-1})^2 \sim \chi^2_{N(T-1)} \tag{42}$$

which can be drawn directly.

3.5.1.5. The β draw (log-wage equation coefficients). The conditional posterior distribution of β is:

$$p(\beta) \propto \exp\left\{-\frac{h_u}{2}\sum_{n,t}(Y^*_{nt}-X^*_{nt}\beta)^2\right\}, \tag{43}$$

where

$$Y^*_{n1} = (1-\rho^2)^{1/2}\log w_{n1}, \tag{44}$$

$$Y^*_{nt} = \log w_{nt} - \rho \log w_{n,t-1} \quad (t \geqslant 2), \tag{45}$$

$$X^*_{n1} = (1-\rho^2)^{1/2}X_{n1}, \tag{46}$$

$$X^*_{nt} = X_{nt} - \rho X_{n,t-1} \quad (t \geqslant 2). \tag{47}$$

I draw from the implied Gaussian distribution for β directly.

3.5.1.6. The ρ draw (productivity shock correlation coefficient). The conditional posterior distribution of ρ is

$$\begin{aligned} p(\rho) \propto{}& \chi_{\rho\in[-1,1]}(1-\rho^2)^{N/2}\exp\left[\frac{-h_u(1-\rho^2)}{2}\sum_n(\log w_{n1}-X'_{n1}\beta)^2\right] \\ &\times \exp\left\{-\frac{h_u}{2}\sum_{n,t\geqslant 2}(Y^*_{nt}-X^*_{nt}\beta)^2\right\} \\ &\times \exp\left\{-\frac{1}{2}(\{z_{njt}-Q'_{njt}\Lambda\})'_{j=1,2,3}\Sigma_\eta^{-1}(\{z_{njt}-Q'_{njt}\Lambda\})_{j=1,2,3}\right\}, \end{aligned} \tag{48}$$

where Y^*_{nt} and X^*_{nt} $(t \geqslant 2)$ are as defined above, the indicator function reflects the prior imposed in Section 3.4, and the final term enters because ρ enters the future component regressors (see Appendix A), and thus Q_{njt}. To draw from this distribution I used rejection sampling. I draw candidate $\tilde{\rho}$ from the Gaussian distribution

$$p(\tilde{\rho}) \propto \exp\left\{ -\frac{h_u}{2} \sum_{n,t \geqslant 2} (Y^*_{nt} - X^*_{nt}\beta)^2 \right\} \tag{49}$$

and accept candidate $\tilde{\rho}$ with probability proportional to the product of the remaining terms of (48) when evaluated at the candidate draw.

3.5.1.7. The h_u draw (log-wage error's precision). The conditional posterior for h_u is

$$p(h_u) \propto h_u^{(NT-2)/2} \exp\left\{ -\frac{h_u}{2} \left\{ \begin{array}{l} (1-\rho^2) \sum_n (\log w_{n1} - X'_{n1}\beta)^2 \\ + \sum_{n,t>1} (\log w_{nt} - \rho \log w_{nt-1} - (X'_{nt} - \rho X'_{nt-1})\beta)^2 \end{array} \right\} \right\} \tag{50}$$

so that

$$h_u S^* \sim \chi^2_{NT}, \tag{51}$$

where S^* is the sum of squares in the inner brackets of (50).

3.5.1.8. The $\{z_{njt}\}_{j=1,3}$ draw (relative alternative values). Given values for all the other parameters and variables, the conditional posterior distribution of $\{z_{njt}\}_{j=1,2,3}$ is as follows.

$$p(\{z_{njt}\}_{j=1,2,3}) \propto \prod_{n,t} |H_\eta|^{1/2} \exp\left\{ -\frac{1}{2} \begin{pmatrix} \{z_{n1t} - Q'_{n1t}\Lambda\} \\ \{z_{n2t} - Q'_{n2t}\Lambda\} \\ \{z_{n3t} - Q'_{n3t}\Lambda\} \end{pmatrix}' \times H_\eta \begin{pmatrix} \{z_{n1t} - Q'_{n1t}\Lambda\} \\ \{z_{n2t} - Q'_{n2t}\Lambda\} \\ \{z_{n3t} - Q'_{n3t}\Lambda\} \end{pmatrix} \right\} I(\{z_{njt}, d_{njt}\}_{j=1,2,3}) \tag{52}$$

Clearly, $\{z_{njt}\}_{j=1,2,3}$ follows a truncated multivariate Gaussian distribution. Following Geweke (1991), I draw the z_{njt} one-by-one. This reduces the problem to sampling from three truncated, univariate Gaussian distributions, which is easily accomplished using standard inverse CDF techniques.

3.5.1.9. The π and θ draw (future component and utility function parameters). Given values for all the other parameters and variables, π and θ can be drawn as a block. Recalling that $\Lambda=[\theta',\pi']'$, the draw is made from the following distribution:

$$p(\Lambda) \propto \exp\left\{-\frac{1}{2}(C_2/1000)^2\right\} \exp\left\{-\frac{1}{2}\begin{pmatrix}\{z_{n1t}-Q'_{n1t}\Lambda\}_{n,t}\\ \{z_{n2t}-Q'_{n2t}\Lambda\}_{n,t}\\ \{z_{n3t}-Q'_{n3t}\Lambda\}_{n,t}\end{pmatrix}'\right.$$

$$\left.\times(H_\eta\otimes I_{NT})\begin{pmatrix}\{z_{n1t}-Q'_{n1t}\Lambda\}_{n,t}\\ \{z_{n2t}-Q'_{n2t}\Lambda\}_{n,t}\\ \{z_{n3t}-Q'_{n3t}\Lambda\}_{n,t}\end{pmatrix}\right\}. \tag{53}$$

Hence, leaving aside the prior, the joint density of π and θ is Gaussian,

$$p(\Lambda) \propto \mathrm{N}((\mathbf{Q}'(H_\eta\otimes I_{NT})\mathbf{Q})^{-1}\mathbf{Q}'\mathbf{Z},(\mathbf{Q}'(H_\eta\otimes I_{NT})\mathbf{Q})^{-1}), \tag{54}$$

where $\mathbf{Z}=\left\{\begin{matrix}\{z_{n1t}\}_{nt}\\ \{z_{n2t}\}_{nt}\\ \{z_{n3t}\}_{nt}\end{matrix}\right\}$ and $\mathbf{Q}=\left\{\begin{matrix}\{Q_{n1t}\}_{nt}\\ \{Q_{n2t}\}_{nt}\\ \{Q_{n3t}\}_{nt}\end{matrix}\right\}$. I draw from this distribution, and accept with probability $\exp\{-\frac{1}{2}(\tilde{C}_2/1000)^2\}$, where $\tilde{C}_2$ is the current age 17 transition cost draw.

3.5.1.10. The H_η draw (precision matrix for preference shocks). In this step I draw H_η and impose the scale normalization that $\Sigma_\eta(1,1)=1$. An appropriate way to do this is described in Geweke et al. (1994) and is as follows. First, note that unscaled H_η has a distribution with kernel

$$|H_\eta|^{(NT-4)/2}\exp\left\{-\frac{1}{2}\begin{pmatrix}\{z_{n1t}-Q'_{n1t}\Lambda\}_{n,t}\\ \{z_{n2t}-Q'_{n2t}\Lambda\}_{n,t}\\ \{z_{n3t}-Q'_{n3t}\Lambda\}_{n,t}\end{pmatrix}'(H_\eta\otimes I_{NT})\begin{pmatrix}\{z_{n1t}-Q'_{n1t}\Lambda\}_{n,t}\\ \{z_{n2t}-Q'_{n2t}\Lambda\}_{n,t}\\ \{z_{n3t}-Q'_{n3t}\Lambda\}_{n,t}\end{pmatrix}\right\}. \tag{55}$$

Hence, H_η follows a Wishart distribution, or

$$H_\eta \sim W(S^{-1},NT),$$

where

$$S=\sum_{n,t}\begin{pmatrix}\{z_{n1t}-Q'_{n1t}\Lambda\}\\ \{z_{n2t}-Q'_{n2t}\Lambda\}\\ \{z_{n3t}-Q'_{n3t}\Lambda\}\end{pmatrix}\begin{pmatrix}\{z_{n1t}-Q'_{n1t}\Lambda\}\\ \{z_{n2t}-Q'_{n2t}\Lambda\}\\ \{z_{n3t}-Q'_{n3t}\Lambda\}\end{pmatrix}'.$$

It is easy to draw H_η from the Wishart distribution, and then invert this matrix to form Σ_η. Finally, Σ_η and the coefficients of the utility function and future component regressors Λ are appropriately normalized. The results I report are based on the normalized coefficient values.

4. The data

The data are from the 1979 youth cohort of the National Longitudinal Surveys of Labor Market Experience (NLSY). The NLSY consists of 12686 individuals, about half of whom are male, that were 14–21 years of age as of January 1, 1979. The sample consists of a core random sample, and an oversample of blacks, Hispanics, poor whites, and the military. This work is based on the whites in the core sample that were 16 years of age or less as of January, 1979. I follow each individual in the subsample from the time they leave school through 1993. Thus, there are at most 15 annual observations available for each individual in the sample.

Estimation of the model requires information on annual hours worked, annual income, and wealth for each individual. The NLSY includes the information needed to obtain these quantities. However, because the model involves discrete labor-supply decisions some definitional arbitrariness is unavoidable when assigning observations to the model's alternatives. I use the following definitions.

(1) *Work*: An individual was said to have worked part-time, 1000 annual hours, if they reported 1–1500 h worked during the year. An individual was said to have worked full-time, 2000 annual hours, if they reported 1501–2500 h worked during the year. An individual was said to have worked overtime, 3000 annual hours, if they reported over 2500 h of work during the year. He is assumed to have remained at home otherwise.

(2) *Real wage*: The real annualized wage is determined by dividing the reported real annual income by the number of reported annual hours and then multiplying by 2000.

(3) *Wealth*: Asset data is available only in years 1985–1990, and 1992–1993. When available, I constructed the individual's wealth position by subtracting a measure of total reported debts from total reported assets. Total debt was determined by adding the values of mortgages, back taxes, debts on farm or business, money owed on vehicles and other debts that exceed \$500. Total assets were found by adding the values of residential property, savings accounts, farms or businesses, vehicles, and other assets with worth that exceeded \$500. [15]

In my model labor market experience, education and age account for human capital. Therefore, it is important to censor individuals that spend time in human-capital enhancing activities that are not accounted for by the model. In particular, all individuals who spent any time in the military were censored from the sample. Also, because the model treats school as exogenous, I begin to record individual observations in the year following their final year of school. Finally, to reduce the computational burden of the numerical procedure, I censored from the sample all individuals with any missing wealth information within the years 1985–1990 or 1992–1993.

(4) *Real returns on savings*: I construct this series by finding $i_y/(\text{inflation}_y)$, where i_y is the 1-year t-bill rate available in year y (obtained from http://www.federalreserve.gov/

[15] It is well known that wealth data includes a substantial amount of noise. In my case, the large variation about age-specific means made effective implementation of the estimation procedure difficult. To help reduce the influence of extreme values on the results I manipulated the asset data as follows. If individual n at age t had an asset value A_{nt}, and $A_{nt} \notin [-250t, 10000t]$, then I eliminated the observation from the sample.

releases/H15/data/a/tbaa1y.txt), and inflation$_y$ is calculated from the cpi-urban price index.

The final panel includes 309 individuals and a total of 2524 observations. This panel is small relative to other panels that have been used in the labor supply literature. There are two points to note in this regard. First, I employ a Bayesian estimation procedure so that inference is exact in small samples. Second, my panel is not atypical, in the sense that, as I point out below, the behavior of the young men I observe is consistent with well known stylized facts in male labor supply (see, e.g., Pencavel, 1986).

4.1. Summary statistics

Table 1 describes the choice distribution by age for the individuals in my panel. Full-time work accounts for about 64% of overall choices, part-time another 13% and

Table 1
Hours choice distributions for white males ages 17–30

Choice	Part-time	Full-time	Overtime	Home	Total
Age					
17	6	2	0	1	9
	66.67%	22.22%	0.00%	11.11%	100.00%
18	12	14	1	3	30
	40.00%	46.67%	3.33%	10.00%	100.00%
19	39	52	5	4	100
	39.00%	52.00%	5.00%	4.00%	100.00%
20	51	94	15	3	163
	31.29%	57.67%	9.20%	1.84%	100.00%
21	33	119	29	1	182
	18.13%	65.38%	15.93%	0.55%	100.00%
22	31	138	36	0	205
	15.12%	67.32%	17.56%	0.00%	100.00%
23	24	150	44	2	220
	10.91%	68.18%	20.00%	0.91%	100.00%
24	22	160	54	2	238
	9.24%	67.23%	22.69%	0.84%	100.00%
25	19	169	56	2	246
	7.72%	68.70%	22.76%	0.81%	100.00%
26	23	170	66	1	260
	8.85%	65.38%	25.38%	0.38%	100.00%
27	13	181	76	2	272
	4.78%	66.54%	27.94%	0.74%	100.00%
28	23	179	78	1	281
	8.19%	63.70%	27.76%	0.36%	100.00%
29	15	128	67	1	211
	7.11%	60.66%	31.75%	0.47%	100.00%
30	9	65	31	2	107
	8.41%	60.75%	28.97%	1.87%	100.00%
Aggregate	320	1621	558	25	2524
	12.68%	64.22%	22.11%	0.99%	100.00%

Table 2
Transition rates of white males ages 17–30

	Choice (t)			
Choice ($t-1$)	Part-time	Full-time	Overtime	Home
Part-time	0.381	0.527	0.068	0.025
Fulltime	0.070	0.813	0.115	0.003
Overtime	0.025	0.271	0.704	0.000
Home	0.565	0.087	0.087	0.261
Standard deviation of change in log annual hours:				0.31

overtime accounts for 22%, which is all but about 1% of the remainder. The frequency of part-time choices declines monotonically to age 25, and then levels off at around 8%. The frequency of 'Home' also declines sharply after age 18, and never exceeds 2% from age 20 to 30. Working full-time and overtime become more common with age. Full-time work is chosen by 22% of young men at age 17, moving to a high of 68% at age 23, and staying above 60% through age 30. Overtime is not an observed choice at age 17, but overtime choices increase by between two and three percentage points per year, leveling off at about 30% at ages 29 and 30.

Previous empirical studies have found evidence of substantial idiosyncratic, year-to-year hours changes (see, e.g., Card, 1994). The current panel reflects this finding: the standard deviation of the change in log annual hours between consecutive years is 0.31.[16] Such year-to-year changes are described in Table 2, which reports the transition rates from the row alternative at age $t-1$ to the column alternative at age t. On average, full-time and overtime are chosen across consecutive years at rates of 81% and 70%, respectively. This persistence is not found in part-time or home choices, which are chosen consecutively with rates 38% and 26%, respectively. Among those who choose part-time, 53% choose to work full-time the following year. Note that no one transitions from overtime to home.

Table 3 records mean annual wages in terms of 1987 dollars. Wage data is not available for the few individuals who are observed to work at age 17, and note this implies that wage data is available in my panel beginning in 1981. Mean wages rise monotonically from \$9565 at age 18 to \$21995 at age 30. Table 4 reports the aggregate wealth statistics in terms of 1987 dollars.[17] As noted above, asset and debt data has been collected only since 1985. Hence, the youngest men in the sample (age 14 in 1979) are age 20 when initial wealth information is collected. Mean wealth is smallest at age 20 and has value \$2257. It increases through age 24, where it peaks at \$7549. After age 24 mean wealth varies between \$5400 and \$7500 without any obvious pattern.

[16] The standard deviation of the change in log annual hours was constructed by first omitting all 'home' observations, and then forming the pooled series $\log(h_{nt}/h_{n,t-1})$ for all young men n and ages t for which the term was defined.

[17] Statistics were calculated after the asset screen, as that is the data that my model attempts to explain.

Table 3
Mean real[a] wage of white males ages 18–30

Age	Mean wage	Standard error	Observations
18	9565	1445	25
19	10663	531	95
20	11849	783	158
21	13004	567	178
22	14089	729	203
23	16812	1454	217
24	18016	1174	235
25	18774	628	239
26	19133	597	255
27	19777	586	264
28	20418	606	272
29	21081	680	204
30	21995	1015	100

[a]Base year = 1987.

Table 4
Mean real[a] wealth of white males ages 18–30

Age	Mean wealth	Standard error	Observations
18	—	—	—
19	—	—	—
20	2257	404	36
21	3772	413	93
22	4259	389	166
23	5386	454	170
24	7549	682	187
25	6631	743	175
26	7421	1022	146
27	5860	1054	124
28	7209	1182	123
29	5450	1268	151
30	7132	1919	77

[a]Base year = 1987.

4.2. Statistics by educational attainment

It is of interest to compare the choices, wages and wealth of individuals with different levels of education. I choose to group individuals according to educational attainment as follows: 8–11 years, 12–15 years or 16 + years. This grouping was chosen with an eye towards highlighting the college graduation effect (16 + years of education) and the high school graduation effect (12–15 years of education). My panel contains 633 person–period observations at the lowest level, 1649 observations for those who graduated from high school but not college, and 242 for those who received at least an undergraduate degree.

Table 5A–C describe the life-cycle choice distributions for each educational group. Although the sample sizes are quite small, one feature of the data is that over 10%

Table 5
Hours choice distributions for white males ages 17–30

Choice	Part-time	Full-time	Overtime	Home	Total
Age					
(a) *8–11 years of education*					
17	6	2	0	1	9
	66.67%	22.22%	0.00%	11.11%	100.00%
18	12	12	1	3	28
	42.86%	42.86%	3.57%	10.71%	100.00%
19	21	18	1	2	42
	50.00%	42.86%	2.38%	4.76%	100.00%
20	21	25	5	1	52
	40.38%	48.08%	9.62%	1.92%	100.00%
21	12	36	8	0	56
	21.43%	64.29%	14.29%	0.00%	100.00%
22	12	36	8	0	56
	21.43%	64.29%	14.29%	0.00%	100.00%
23	11	39	6	0	56
	19.64%	69.64%	10.71%	0.00%	100.00%
24	5	38	10	2	55
	9.09%	69.09%	18.18%	3.64%	100.00%
25	6	36	11	1	54
	11.11%	66.67%	20.37%	1.85%	100.00%
26	9	34	11	1	55
	16.36%	61.82%	20.00%	1.82%	100.00%
27	2	39	12	1	54
	3.70%	72.22%	22.22%	1.85%	100.00%
28	7	37	10	0	54
	12.96%	68.52%	18.52%	0.00%	100.00%
29	5	27	11	0	43
	11.63%	62.79%	25.58%	0.00%	100.00%
30	2	13	4	0	19
	10.53%	68.42%	21.05%	0.00%	100.00%
Aggregate	131	392	98	12	633
	20.69%	61.93%	15.48%	1.90%	100.00%
(b) *12–15 years of education*					
17	0	0	0	0	0
	0.00%	0.00%	0.00%	0.00%	100.00%
18	0	2	0	0	2
	0.00%	100.00%	0.00%	0.00%	100.00%
19	18	34	4	2	58
	31.03%	58.62%	6.90%	3.45%	100.00%
20	30	69	10	2	111
	27.03%	62.16%	9.01%	1.80%	100.00%
21	21	83	21	1	126
	16.67%	65.87%	16.67%	0.79%	100.00%
22	19	100	28	0	147
	12.93%	68.03%	19.05%	0.00%	100.00%
23	12	103	35	2	152
	7.89%	67.76%	23.03%	1.32%	100.00%

Table 5 (*continued*)

Choice	Part-time	Full-time	Overtime	Home	Total
24	15	103	42	0	160
	9.38%	64.38%	26.25%	0.00%	100.00%
25	12	109	41	1	163
	7.36%	66.87%	25.15%	0.61%	100.00%
26	13	107	47	0	167
	7.78%	64.07%	28.14%	0.00%	100.00%
27	10	111	55	1	177
	5.65%	62.71%	31.07%	0.56%	100.00%
28	11	111	55	1	178
	6.18%	62.36%	30.90%	0.56%	100.00%
29	9	80	44	1	134
	6.72%	59.70%	32.84%	0.75%	100.00%
30	6	44	22	2	74
	8.11%	59.46%	29.73%	2.70%	100.00%
Aggregate	176	1056	404	13	1649
	10.67%	64.04%	24.50%	0.79%	100.00%
(c) *16 or more years of education*					
17	0	0	0	0	0
	0.00%	0.00%	0.00%	0.00%	100.00%
18	0	0	0	0	0
	0.00%	100.00%	0.00%	0.00%	100.00%
19	0	0	0	0	0
	0.00%	0.00%	0.00%	0.00%	100.00%
20	0	0	0	0	0
	0.00%	0.00%	0.00%	0.00%	100.00%
21	0	0	0	0	0
	0.00%	0.00%	0.00%	0.00%	100.00%
22	0	2	0	0	2
	0.00%	100.00%	0.00%	0.00%	100.00%
23	1	8	3	0	12
	8.33%	66.67%	25.00%	0.00%	100.00%
24	2	19	2	0	23
	8.70%	82.61%	8.70%	0.00%	100.00%
25	1	24	4	0	29
	3.45%	82.76%	13.79%	0.00%	100.00%
26	1	29	8	0	38
	2.63%	76.32%	21.05%	0.00%	100.00%
27	1	31	9	0	41
	2.44%	75.61%	21.95%	0.00%	100.00%
28	5	31	13	0	49
	10.20%	63.27%	26.53%	0.00%	100.00%
29	1	21	12	0	34
	2.94%	61.76%	35.29%	0.00%	100.00%
30	1	8	5	0	14
	7.14%	57.14%	35.71%	0.00%	100.00%
Aggregate	13	173	56	0	242
	5.37%	71.49%	23.14%	0.00%	100.00%

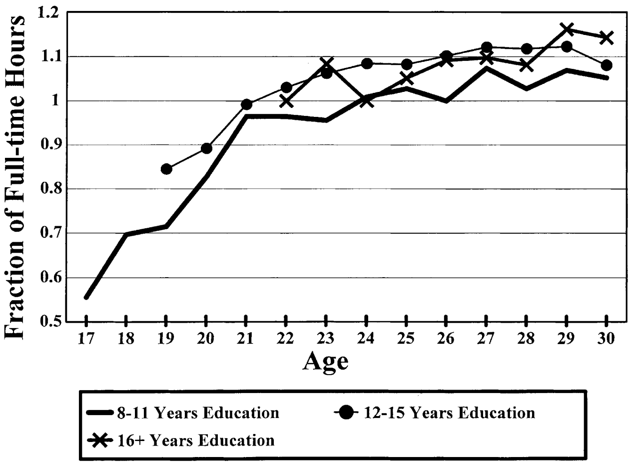

Fig. 1. Labor supply by age and education.

of men aged 17 and 18 who have neither completed nor are attending high school choose not to work, and still 5 remain home at age 19. For young men who receive their high school (but not college) diploma, over 3.5% choose not to work at age 19. On the other hand, every young man with at least an undergraduate degree chooses to work after graduation. Table 5A–C suggest that younger and less educated men tend to provide less labor than older, more educated men. Fig. 1 is a plot of Table 5A–C, and makes clear that within each educational group aggregate hours increase with age.

Table 6A–C describe the life-cycle wage path within each educational group. Fig. 2 is a plot of Table 6A–C. After age 22, average wages are greater for those with a high school education, and greater yet for those holding a college diploma. Although college graduates may have higher wage growth than others, most of the wage disparity seems attributable to a level effect: at both ages 22 and 30, a college graduate earns about $10000 more than a nongraduate. Interestingly, at age 30, wages of high school graduates are only $400 greater than nongraduates. This difference is $1000, in favor of the nongraduates, at age 22.

Table 7A–C describe the life-cycle wealth profile for young men by educational group. Fig. 3 is a plot of Table 7A–C. As with wages, wealth is positively related to education. Average wealth is positive over the sample period for each group. However, only the high school (no college) group exhibits evidence of wealth growth. Average wealth for high school—but not college—graduates grows from $2500 at age 18 to $7500 at age 22, and then remains near that value. The mean wealth of college graduates fluctuates around $10000, while the mean wealth of the least educated group evolves about a mean of $2500.

5. Marginal posterior distributions and fit

The model includes 62 parameters: 18 coefficients for the wage equation, 11 for the utility function, six for the savings function, nine variance or covariance terms and

Table 6
Mean real[a] wages of white males ages 18–30

Age	Mean wage	Standard error	Observations
(a) *8–11 years of education*			
18	9747	1560	23
19	10254	777	39
20	13154	2174	50
21	13504	1260	55
22	14667	2214	56
23	14612	957	56
24	14339	963	53
25	16037	869	52
26	16481	981	53
27	16343	1123	50
28	17631	1300	53
29	18380	1007	42
30	20246	2488	18
(b) *12–15 years of education*			
18	7465	2412	2
19	10949	724	56
20	11244	552	108
21	12780	599	123
22	13726	555	145
23	17435	2076	149
24	18625	1667	159
25	18422	762	158
26	18580	674	164
27	19430	676	173
28	19426	681	171
29	20303	865	129
30	20656	1057	69
(c) *16+ years of education*			
18	—	—	—
19	—	—	—
20	—	—	—
21	—	—	—
22	24219	781	2
23	19335	2478	12
24	22274	2114	23
25	25600	2288	29
26	25221	2138	38
27	25426	1787	41
28	27029	1669	48
29	27561	1766	33
30	31521	3303	13

[a]Base year = 1987.

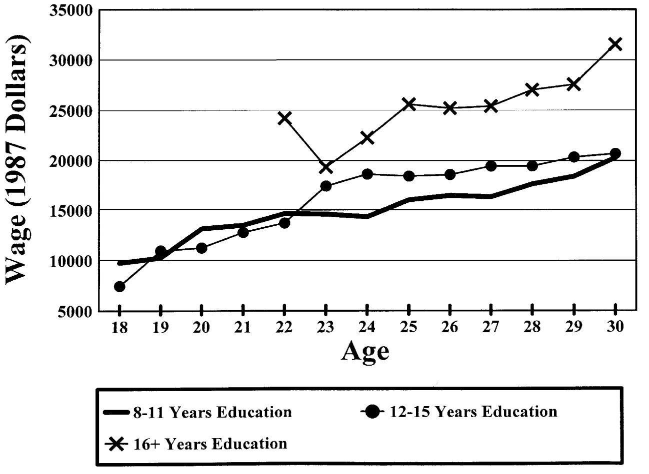

Fig. 2. Wage by age and education.

18 coefficients for the future component. Only 61 coefficients were free to vary, since $\Sigma_\eta(1,1)$ was fixed at one.

The marginal posterior distributions described below are based on the last 8000 cycles of a 14,000 cycle Gibbs sampling run. I assessed convergence by looking at plots of the draw sequences, and by using a split-sequence convergence diagnostic suggested by Gelman (1996). Both the visual and more formal statistical evidence suggest that convergence had occurred by cycle 6000.

5.1. *Marginal posterior distributions*

The full set of Gibbs draws for selected wage equation coefficients, the productivity shock variance and the wage equation's correlation coefficient are reported in Fig. 4. The marginal posterior densities derived from the final 8000 draws are provided in Fig. 5. The mean of each posterior has the expected sign. In particular, education, experience, lagged hours, and age each have positive mean, while age squared has a negative posterior mean. The age coefficient's distribution lies entirely to the right of zero and has a plausible mean of 0.12. The density of age squared has substantial mass approaching zero from the left, and none to the right due to the prior.

The marginal posterior distribution of the coefficient on total previous work experience lies strictly to the right of zero with mean 0.034. About 76% of the distribution of once-lagged hours lies to the right of zero, and has mean 0.015. Hence, evaluated at posterior means, the wage elasticity of recent work experience is about 50% greater than that of other work experience. This finding is qualitatively consistent with the results from studies of female labor supply reported by Altug and Miller (1998) and Miller and Sanders (1997). Quantitatively, Altug and Miller (1998), for example, found a wage elasticity of about 0.2 for once-lagged hours, and 0.05 for hours twice lagged.

Table 7
Mean real[a] wealth of white males ages 18–30

Age	Mean wealth	Standard error	Observations
(a) *8–11 years of education*			
18	—	—	—
19	—	—	—
20	1508	581	9
21	4561	806	30
22	3596	744	47
23	2783	467	41
24	5873	1169	42
25	3584	829	41
26	1829	1270	35
27	2638	1935	26
28	6052	2119	26
29	1309	1614	35
30	448	1816	15
(b) *12–15 years of education*			
18	—	—	—
19	—	—	—
20	2506	493	27
21	3396	466	63
22	4485	461	117
23	6194	589	120
24	7484	871	126
25	7392	1021	118
26	7639	1264	93
27	6374	1307	90
28	6292	1431	78
29	6483	1714	99
30	8513	2437	57
(c) *16+ years of education*			
18	—	—	—
19	—	—	—
20	—	—	—
21	—	—	—
22	6600	1344	2
23	6477	1727	9
24	11689	1971	19
25	8823	1682	16
26	17172	3349	18
27	10557	1979	8
28	12559	3723	19
29	7962	3644	17
30	11451	6080	5

[a]Base year = 1987.

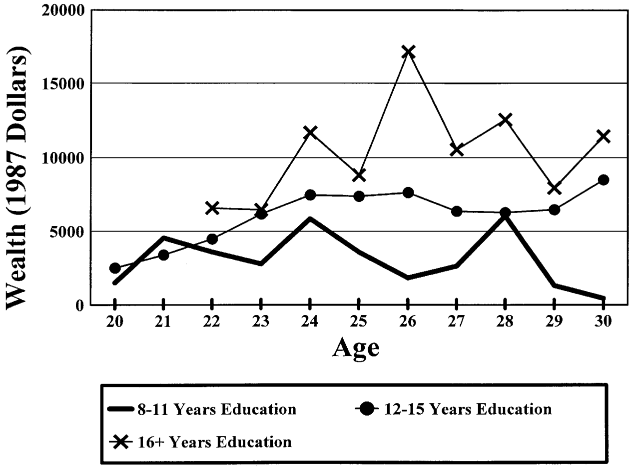

Fig. 3. Wealth by age and education.

This suggests that recent work experience might have a relatively greater impact on female's wage offers.

The means and standard deviations of each wage equation parameter's marginal posterior distribution are reported in Table 8. Note that the wage equation includes an intercept and 12 year dummies, corresponding to the 13 years for which wage observations exist (see Section 4.1). These year dummies are estimated very imprecisely. This table also reports the results of an OLS regression on observed log-wages. The OLS regression uses the specification as described in Section 2.1, except that the residual is assumed to be independently and identically distributed across time and individuals. Also, there are no restrictions placed on the equation's coefficients. It is well known that such OLS estimates can be biased and inconsistent. There are two primary sources of this bias. The first is selection due to incidentally truncated wages. The second can arise if the residual is serially correlated. To see this, note that market experience at t is a function of the decision at $t-1$, and because the decision at $t-1$ depends on the value of the residual, serial correlation in the error process could lead to endogeneity. For example, if individuals with higher wages tend to work more, then the partial correlation between a secularly-persistent skill shock and market experience might be positive. Consequently, the coefficient of experience β_1 might be estimated with upward bias. It turns out, as Table 8 shows, that the estimates from the different procedures are close, in the sense that each posterior mean is within two standard errors of the OLS point estimates.

The empirical posterior densities of selected utility function parameters are described in Fig. 6. Coefficients of terms involving the product of age and leisure have distributions that lie largely on one side of zero. Moreover, the posterior means of these distributions are plausible economically, as they jointly imply that, at every age, the nonstochastic part of utility is strictly increasing in leisure. The distributions of log

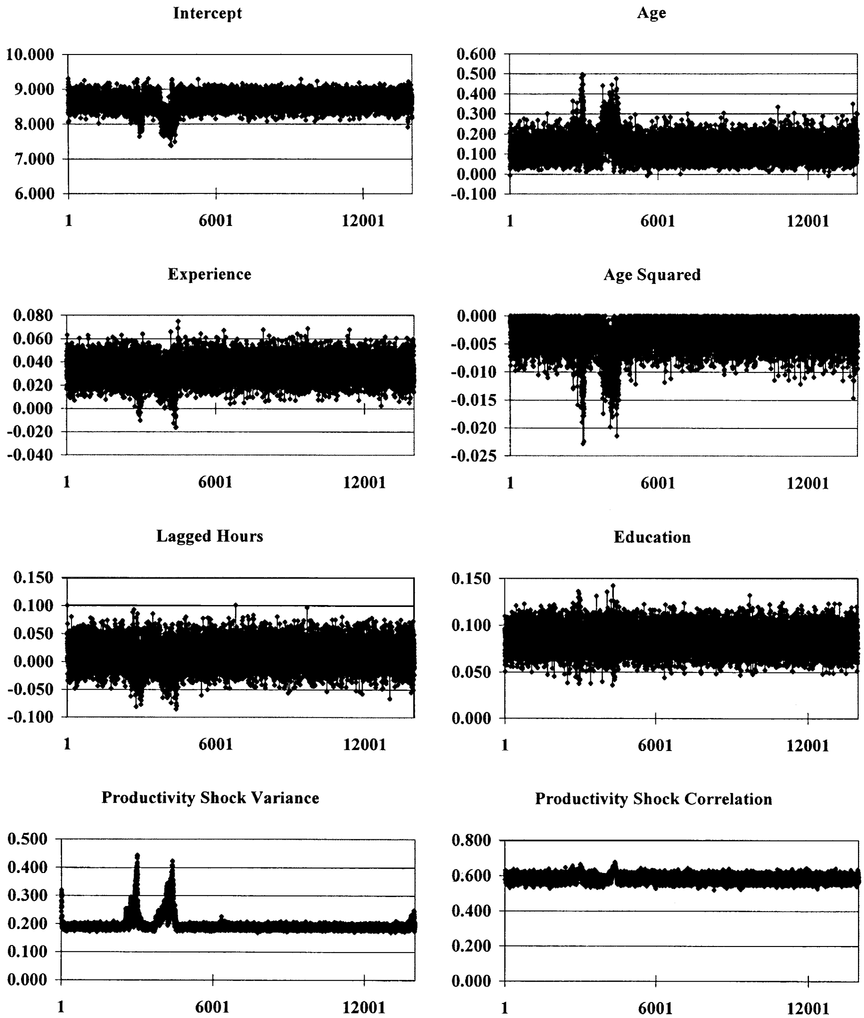

Fig. 4. Draw sequences for selected parameters.

consumption and log consumption squared have means near zero and posterior standard deviations that are large relative to these means. In fact, these parameters are jointly insignificant in the sense that simulations of the model from posterior means, but excluding these two consumption terms, have essentially no effect on simulated life-cycle outcomes (e.g., at each age, average hours change only beyond the third or fourth decimal point). Omitting terms in age and leisure, on the other hand, affects outcomes dramatically. It is worthwhile to note that the coefficient of the product of log consumption and log leisure has a positive mean, 0.026, which suggests that consumption and leisure are complements. However, its posterior standard deviation is relatively large at 0.073.

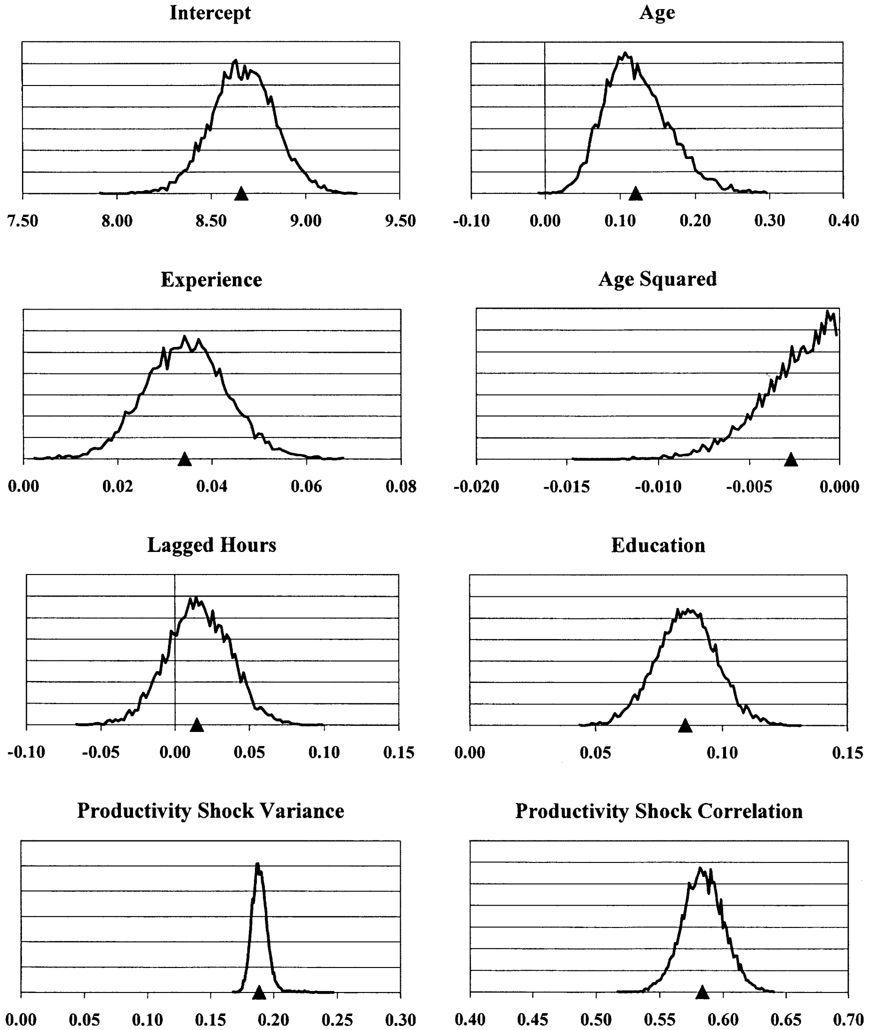

Fig. 5. Marginal posterior distributions for selected log-wage equation parameters. The triangles mark posterior means.

The posterior means for all of the utility function parameters are reported in Table 9. It is interesting to note that the mean of C_1, the transition cost affecting decisions after age 17, is greater than two posterior standard deviations from zero. In addition, its mean, -0.79, is large in magnitude relative to the posterior means of the standard deviations of the second and third alternatives' preference shocks, which are 0.55 and 0.62, respectively. This suggests that transition costs may mitigate hours variation across adjacent years. The correlations between the first and second, and second and third, preference shocks are -0.25 and 0.65, with relatively

Table 8
Posterior means and standard deviations and OLS estimates for log wage equation parameters

Parameter[a]	Dynamic model		OLS	
	Mean	SD	Estimate	SE
β_1: Intercept	8.662	0.168	8.991	0.199
β_2: Experience	0.034	0.009	0.029	0.007
β_3: Lagged hours	0.015	0.021	−0.005	0.017
β_4: Age	0.121	0.043	0.057	0.051
β_5: Age^2	−0.003	0.002	0.002	0.003
β_6: Education	0.085	0.012	0.087	0.008
β_7: 1982	0.078	0.151	−0.095	0.196
β_8: 1983	−0.058	0.163	−0.197	0.193
β_9: 1984	−0.089	0.175	−0.201	0.200
β_{10}: 1985	−0.128	0.189	−0.223	0.210
β_{11}: 1986	−0.134	0.205	−0.222	0.221
β_{12}: 1987	−0.101	0.221	−0.182	0.232
β_{13}: 1988	−0.106	0.237	−0.212	0.240
β_{14}: 1989	−0.191	0.252	−0.294	0.247
β_{15}: 1990	−0.231	0.268	−0.373	0.252
β_{16}: 1991	−0.311	0.283	−0.473	0.256
β_{17}: 1992	−0.397	0.299	−0.605	0.260
β_{18}: 1993	−0.472	0.315	−0.738	0.263
ρ: Log wage error correlation	0.584	0.016		
$\sigma^2(u)$: Error process variance	0.189	0.016		

[a] ‘Age’ is the age of the respondent minus 16. ‘Education’ is years of education minus 10.

small empirical standard deviations. The posterior mean correlation between the first and third shocks is largest at −0.89, which is large relative to its posterior standard deviation of 0.02. These precise posterior distributions, based on vague priors and a fairly small sample, suggest that preference shocks are important determinants of labor supply decisions. This lends support to specifications used by Keane and Wolpin (1997), among others, in the structural empirical labor supply literature.

Table 10 gives the posterior means and standard deviations for the parameters associated with the savings function and future component. The partial correlation of income with savings decisions seems reasonable. Evaluated at its posterior mean, an additional \$1000 in income leads to \$105 in additional savings. I find that an additional year of education is associated with \$310 in additional savings. Also, individuals with higher wealth seem to save more. All else equal, when evaluated at posterior means the savings rule requires increasing savings by 31 cents for every additional dollar of wealth carried into the period.

Table 10 also reports the posterior means and standard deviations of the future component parameters. The future component includes 18 terms in the agent’s state variables. Seven of the 18 coefficients have posterior distributions that lie largely to

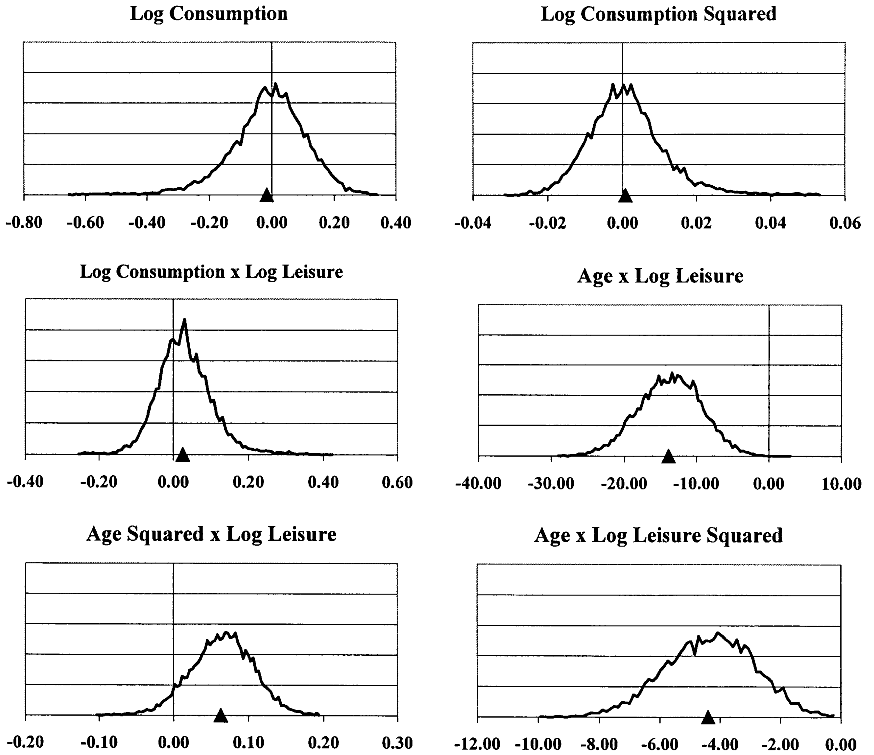

Fig. 6. Marginal posterior distributions for selected utility function parameters. The triangles mark posterior means.

one side of zero. Three of these are the alternative specific constants, and each of the others involves combinations of experience, education, age and the current hours decision. The coefficients on terms involving wages and savings are jointly significant, however, in the sense that setting them to zero substantially worsens the fit of the model, as described next.

5.2. Fit

To assess goodness of fit I simulated the model from the means of its parameters' marginal posterior distributions, and compared the resulting choice, wage, wealth and transition profiles to those derived from my panel. I valued age 17 wealth at $\$2941/(1.02)^3$ in each simulation. The reason is that wealth data is first available at age 20, and then has aggregate mean of \$2941. Since the average real interest rate over the sample period was about 2%, and since my simulated life-cycles begin at age 17, my choice of endowment followed. The years of schooling of the agents in my panel ranges from eight to 20. Hence, I simulated 14-year histories for 500 agents of each educational type from eight to 20. This results in 6500 simulated

Table 9
Posterior means and standard deviations of utility function parameters

Parameter[a]	Mean	SD
θ_1: Log consumption	−0.0162	0.1216
θ_2: (Log consumption)^2	0.0008	0.0104
θ_3: Age × log leisure	−13.8591	4.3428
θ_4: Log consumption × log leisure	0.0258	0.0732
θ_5: Age × (log leisure)^2	−4.3863	1.5071
θ_6: Age^2 × log leisure	0.0630	0.0417
C_1: Transition cost (age > 17)	−0.7929	0.0837
C_2: Transition cost (age = 17)	−0.7773	0.1511
ϕ_1: Age × hours	−5.4364	1.3515
ϕ_2: Part-time and 17	0.4879	0.4341
ϕ_3: Full-time and 17	0.4475	0.3323
$\Sigma\eta(1,1)$ SD of alternative one's preference shock	1.00	PEGGED
$\Sigma\eta(2,2)$ SD of alternative two's preference shock	0.63	0.10
$\Sigma\eta(3,3)$ SD of alternative three's preference shock	0.59	0.07
$\Sigma\eta(1,2)$ Correlation of preference shocks one and two	−0.43	0.17
$\Sigma\eta(1,3)$ Correlation of preference shocks one and three	0.61	0.10
$\Sigma\eta(2,3)$ Correlation of preference shocks two and three	−0.60	0.18

[a]'Age' is the age of the respondent minus 16.

14-year histories. I ran 6500 simulations for each of the three age cohorts in my panel, resulting in a total of 19,500 simulated 14-year histories. Then, choice frequencies, wage profiles, wealth paths and transition frequencies were generated as weighted averages of the simulations. The weights were set according to the proportions in my panel. For example, the simulated choices of college graduates aged 17 to 19 have no weight in the age 17–19 aggregate statistics, because I do not observe any such person.

The age-hours profile is derived from:

$$\text{Hours}_t \equiv 0.5N_{t,\text{part-time}} + N_{t,\text{full-time}} + 1.5N_{t,\text{overtime}} \tag{56}$$

where $N_{t,x}$ is the observed fraction of age t agents that provide x hours of market labor. Fig. 7 plots the simulated and NLSY hours profile. The simulated profile tracks the actual profile quite closely. There is an apparent difference at age 19, where the actual profile rises slightly less quickly than predicted from the simulations. Shedding light on the source of this difference, and performing a more stringent comparison to the data, requires disaggregating choices. Fig. 8A–D plot choice frequencies by alternative. Again, the simulated profiles provide a close visual match. The overpredicted aggregate hours are revealed to stem primarily from too few predicted part-time, and too many full-time, choices at age 19. Both home and overtime choices seem to be well matched over the entire life-cycle.

The model's fit can also be gauged by examining predicted transition frequencies. Fig. 9 suggests the model captures the main features of the data along this dimension

Table 10
Posterior means and standard deviations of savings function and future component parameters

Parameter[a]	Mean	SD
Savings function		
γ_1: Intercept	1838.84	777.84
γ_2: Income	0.105	0.028
γ_3: Education	310.71	137.41
γ_4: Experience	−16.74	81.06
γ_5: Age	69.91	91.74
γ_6: Wealth	0.31	0.04
σ_α: SD of wealth shock	11832	852
Future component[a b]		
π_1: Savings	1.6399E − 05	2.0936E − 05
π_2: Savings^2	1.4020E − 10	1.3884E − 10
π_3: Experience^2	6.8069E − 02	3.7356E − 02
π_4: Age × savings	1.1690E − 06	2.0700E − 06
π_5: Experience × savings	−8.0700E − 07	1.4630E − 06
π_6: Wage × savings	1.7370E − 07	5.8110E − 06
π_7: Education × savings	−1.0610E − 06	2.1110E − 06
π_8: Wage × experience	0.0165	0.0564
π_9: Education × experience	0.1225	0.0569
π_{10}: Experience^3	0.0043	0.0013
π_{11}: Age^2 × experience	0.0340	0.0228
π_{12}: Age × experience^2	−0.0110	0.0042
π_{13}: Education × experience^2	0.0022	0.0040
π_{14}: Education^2 × experience	0.0009	0.0064
π_{15}: Age × education × experience	−0.0067	0.0089
π_{16}: Part-time	3.66	0.99
π_{17}: Full-time	6.75	1.79
π_{18}: Overtime	8.43	2.66

[a] 'Age' is the age of the respondent minus 16, and 'Education' is years of education minus 10.
[b] The parameter labels correspond to the discussion in Appendix A.

as well. The model generates substantial choice persistence in full-time and overtime work. The own transition to full-time and overtime hours is 0.81 and 0.70 in the NLSY, respectively, compared to 0.78 and 0.63 in the simulations. The transition rate from part-time to part-time hours is 0.38 in both the NLSY and the simulations. The transitions from home to part-time work are not matched well, but there are very few such transitions in either the NLSY or simulated data. Finally, the volatility of hours between consecutive years is 0.34 in the simulated data, as compared to the NLSY value of 0.31.

Fig. 10 describes the in sample predictions for the aggregate wage and wealth profiles. The main features of the wage profile are reflected in the simulations. The aggregate wage path is somewhat lower than the actual path in the teens and early 20's, and a bit higher between the ages of 26 and 30. The model predicts that mean wealth will increase over the life-cycle from about $2000 in the late teens and early 20's, to about $9000 by age 30. This is broadly consistent with what occurs in the NLSY data.

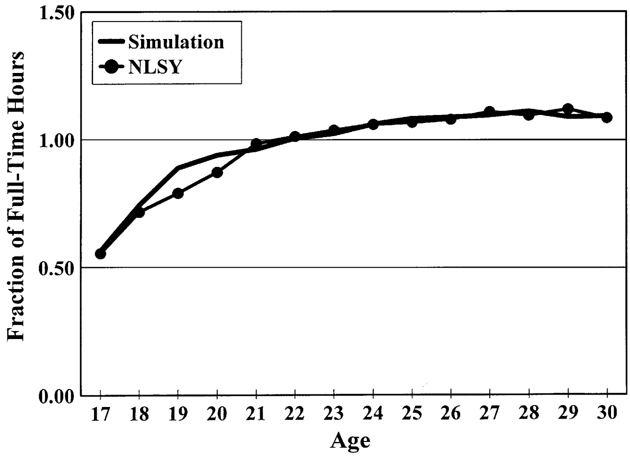

Fig. 7. Simulated and NLSY labor supply by age.

6. Wage, wealth and human capital simulations

6.1. Wages and wealth

Consider first the effect of an uncompensated single-period change in the wage on labor supply.[18] My procedure is to compare all simulations to the baseline simulation reported in Figs. 7 and 8A–D. The new simulations are identical to the baseline, except that at a single age the wage for each individual is increased by 10%. Since the realizations of the stochastic variables for all other ages are held fixed across simulations, any change in labor supply is attributable to the effect of the single-period wage change. I measure the elasticity of the labor supply response by dividing the percentage change in hours (as defined by (56)) by the percentage change in the wage, in this case ten percent.[19]

Fig. 11 describes the results of the wage elasticity simulations. The figure plots the simulated uncompensated wage elasticity for all ages between 17 and 30, and note that each age corresponds to a unique simulation. Consistent with many others' findings, the simulated elasticities are small in magnitude, with the greatest deviation from zero being −0.036 at age 30. Theory, of course, does not predict a sign for the uncompensated elasticity's value. I do not find much evidence of an age effect, which some have argued might exist (see, e.g., Shaw, 1989).

The analogous wealth elasticities for each age are reported in Fig. 12. In these simulations, wealth was incremented at a single age by 10% of its absolute value and elasticities calculated accordingly. The magnitudes are about the same as the wage

[18] The simulated elasticity I report is not wealth- or utility-constant.

[19] Changes in a single-period wage lead to changes in labor supply which lead to changes in subsequent wages through experience effects. In principle, this could have important consequences for subsequent hours decisions. In fact, it turns out that such effects are negligible, so I do not report them here.

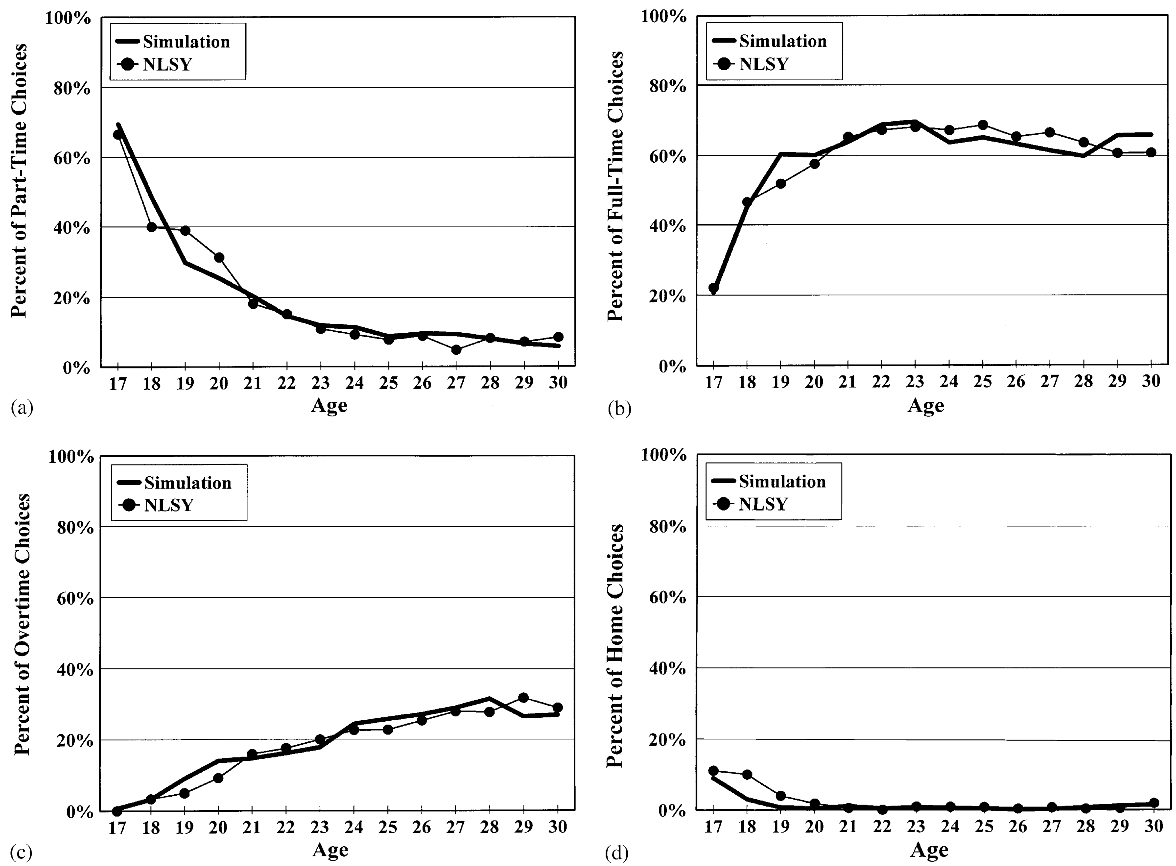

Fig. 8. (a) Fraction of simulated and NLSY agents who choose to work part-time by age. (b) Fraction of simulated and NLSY agents who choose to work full-time by age. (c) Fraction of simulated and NLSY agents who choose to work overtime by age. (d) Fraction of simulated and NLSY agents who choose to provide no market work by age.

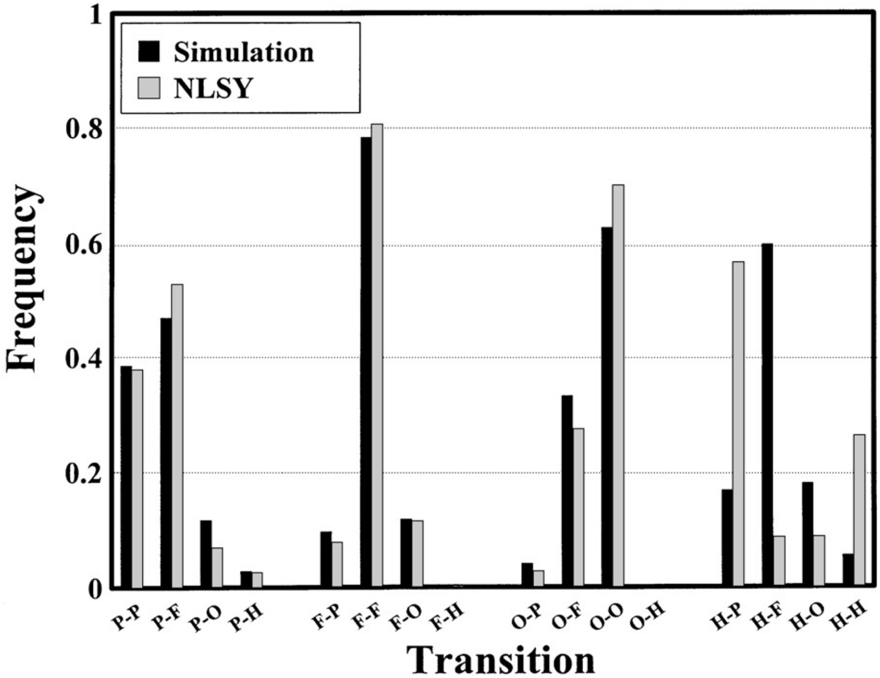

Fig. 9. Aggregate transition frequencies for simulated and NLSY agents. The two columns above ‘P–P’ indicate the frequency of transitions from part-time to part-time hours between consecutive years. ‘P–F’ indicates transitions from part-time to full-time between consecutive years, ‘P–O’ indicates part-time to overtime and ‘P–H’ indicates part-time to no market work. The remaining labels are defined analogously.

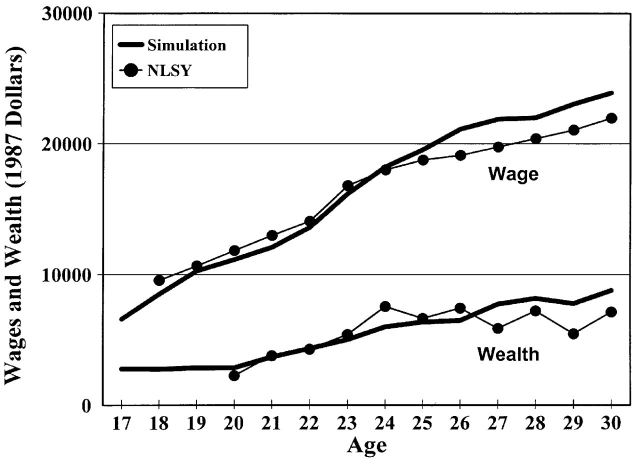

Fig. 10. Simulated and NLSY wage and wealth profiles by age.

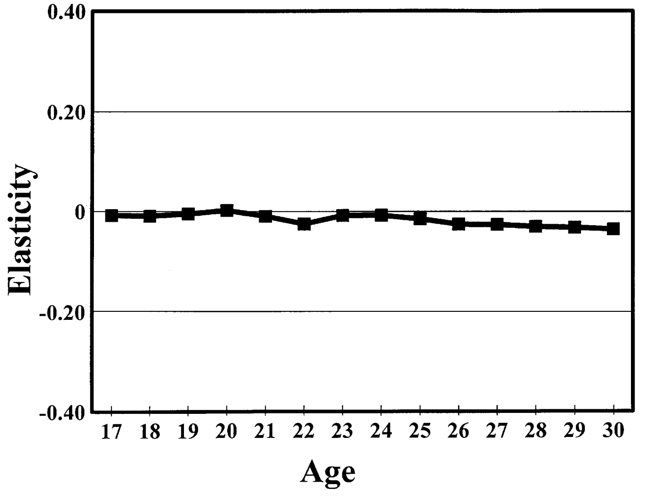

Fig. 11. Estimated uncompensated wage elasticity of labor supply by age.

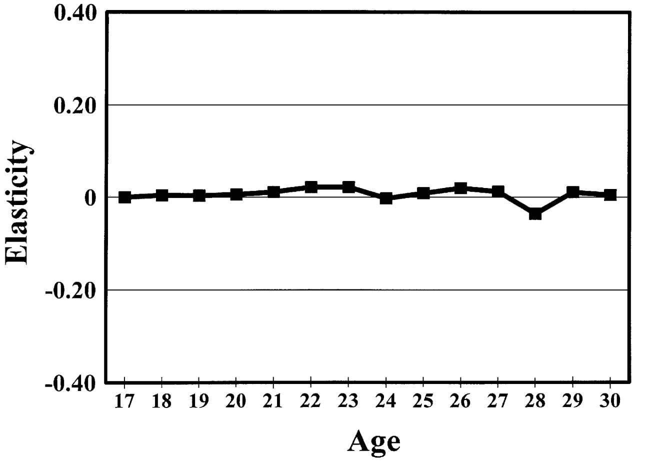

Fig. 12. Estimated uncompensated wealth elasticity of labor supply by age.

elasticities, and again there is no obvious age effect. Overall, Figs. 11 and 12 provide evidence that, for the young men in my panel, changes in wages and wealth are not able to account for observed labor supply variation.

6.2. Human capital

At least since the work of Mincer (1958), theoretical models of human capital and labor supply have described implications of the fact that labor supply's value involves both a contemporaneous payoff and a human capital effect that stems from the investment value of work. There remains, however, much to be learned about the effect that the investment value of work has on actual (or observed) hours outcomes. Shaw (1989) and Altug and Miller (1998) each find indirect evidence for an investment effect on hours. Eckstein and Wolpin (1989) argue that human capital investment is a very important factor in female labor market participation, while Keane and Wolpin (1997) examine the effect it can have on occupational choices.

In the current model, the future component accounts for the expected value of human capital investments. It also accounts for the expected benefit of investment in physical capital. In the model, human capital investments are reflected by total accumulated work experience. Because the law of motion for market experience is $x_{t+1} = x_t + h_{j_t^*}$, I assume that terms in the future component (18) that involve this law of motion account for part of labor supply's investment value. The other part of labor supply's value stems from the fact that previous work experience affects wages differently than other work experience. To account for this, I also include the future component's alternative specific intercepts as determinants of the investment value of current hours.

Let $H(s, j)$ denote the part of the relative future component that captures the value of investment in human capital, where s is the state-vector and j is the hours alternative. Then from the above assumption and Appendix A it is seen that $H(s, j)$ is the relative future component $F(\cdot)$ under the restriction that coefficients π_1, π_2, π_4, π_6 and π_7 are zero. When evaluated at posterior means this has economically reasonable implications. For one, it turns out that

$$H(\bar{s}_t, \text{overtime}) > H(\bar{s}_t, \text{full-time}) > H(\bar{s}_t, \text{part–time}) > 0 \tag{57}$$

for all $t = 17, \ldots, 30$, where $\bar{s}_t$ is the mean observed value of the state vector. The qualitative implication is that part-time work has greater human capital investment value than remaining home (because H is net of the investment return to 'home'), full-time has more than part-time, and overtime more than full-time. Moreover, again evaluated at each age's mean state vector, the "marginal" return to human capital investment is diminishing, in the sense that

$$H(\text{part-time}) > H(\text{full-time}) - H(\text{part-time}) > H(\text{overtime}) - H(\text{full-time}) > 0.$$

To assess the human capital accumulation incentive effect on labor supply I conduct a simulation that investigates the impact of changing the law of motion from $x_{t+1} = x_t + h_j$ to $x_{t+1} = x_t + 0.99h_j$. Under the modified law of motion the expected benefit of human capital accumulation is slightly reduced for every positive hours alternative. Since the contemporaneous costs to labor supply remain unchanged, one might expect this to reduce hours at all ages.

Fig. 13 describes the results of this simulation exercise. Each agent's state-variables were, at each age, set to their value in the baseline simulation. This reveals the effect of the modified law of motion on labor supply while holding everything else fixed. Labor supply falls at all ages, but the effect is largest from age 28 to 30. The effect at age 30 is to reduce labor supply about 3.5%, while before age 28 labor supply is reduced by between 0.5 and 1.0%. Loosely speaking, since the rate of human capital accumulation has been decreased by 1% in the simulations, one might say that this

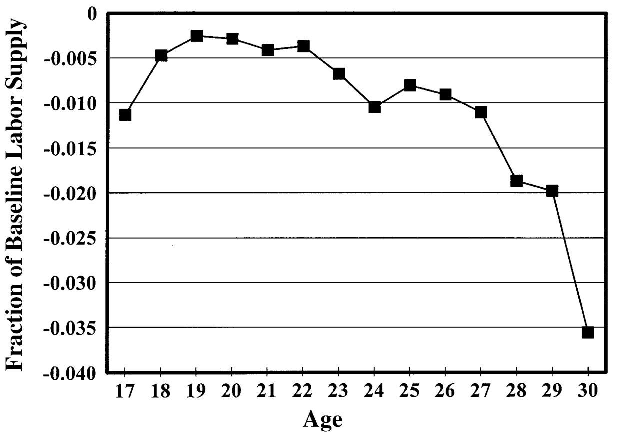

Fig. 13. Effect of a reduction in the human capital investment value of market participation on labor supply by age.

"human capital investment" elasticity of labor supply is between one-half and one until the late 20's, at which point it increases. Because the effect on relatively older men is greater than that on men in their late teens and early 20's, my results suggest that a reduction the expected investment benefit of contemporaneous labor supply might shift down and tend to flatten the age profile of hours.

The age pattern associated with the effect of human capital on labor supply could stem from several sources. One is that young men come to understand slowly over time that contemporaneous work experience has a future benefits. Alternatively, my finding could be driven by age-dependent variation in my panel's characteristics. For example, there are no college graduates at age 17. It is possible that occupations differ in the extent to which they reward accumulated work experience, and that occupations that provide relatively high returns to human capital investment tend to be open primarily to those with higher levels of education. The results of Keane and Wolpin (1997), which suggest blue and white collar occupations reward previous work experience differently, provide some evidence that this might be the case.

7. Conclusion

This paper analyzed a dynamic, stochastic model of life-cycle labor supply that incorporates both savings and human capital investment. I used a Bayesian approach, based on Geweke and Keane (1999), to draw inferences about the model's parameters. A feature of this procedure is that it does not require one to make strong assumptions about the way individuals form expectations. I drew inferences with respect to wage and wealth elasticities and human capital investment effects on labor supply.

The results of my analysis are consistent with past research in life cycle labor supply that assumes rational expectations. Bayesian point estimates of the log-wage equation parameters suggest that education, age and previous market experience each affect wages positively and with reasonable magnitude. Moreover, consistent with results found in studies of female labor supply (see, e.g., Altug and Miller, 1998), Bayesian point estimates indicated recent work experience has a greater effect on wages than other work experience. Estimates of preference parameters implied that utility is increasing in leisure. Although consumption effects were not well pinned down, the estimated effect of the interaction of consumption and leisure implies they are complements.

Simulations of the model from posterior means revealed that it provides a reasonable fit to the data. An additional set of simulations shed light on the effects of wages, wealth and the incentive to accumulate human capital on life-cycle labor supply. I found that wages and wealth seem to have little effect on labor supply decisions, while the incentive to accumulate human capital was found to be relatively more important, particularly for men in their late 20's. My results suggest that reducing the expected benefit of human capital accumulation might shift down and flatten the age-hours profile.

One strong assumption I maintained throughout this research is that all individuals use the same decision rule. Accumulating evidence from laboratory experiments suggests that, in fact, behavior in dynamic decision problems is not well explained by models that allow for only a single decision rule (see, e.g., Houser and Winter, 2001). Houser et al. (2001) extend the approach to inference described in this paper to accommodate multiple decision rules.

Acknowledgements

I thank Michael Keane, John Geweke and Susumu Imai for very valuable input on this research. I also thank two anonymous referees for providing useful comments that improved the paper. Any remaining errors are my own.

Appendix A. The relative future component

The coefficients π that appear in the relative future component are determined by differencing the level future component $F(\cdot)$ with respect to "home", which is the fourth alternative. In the notation of equation (18), it is easy to show that the following parameters drop out due to differencing:

$$\pi_3^*, \pi_4^*, \pi_7^*, \pi_8^*, \pi_9^*, \pi_{12}^*, \pi_{13}^*, \pi_{16}^*, \pi_{17}^*, \pi_{23}^*, \pi_{24}^*, \pi_{26}^*, \pi_{28}^*, \pi_{30}^*.$$

In addition, it is straightforward to verify that the differenced terms corresponding to π_6^* and π_{15}^* are perfectly collinear with hours terms that appear in the utility function. Thus, letting a_k and h_k denote the savings and hours decision associated with alternative k (note the n and t subscripts have been suppressed for clarity), the following terms make up the relative future component estimated in this paper. Note that the coefficient labels correspond to the order in which they are reported in Table 10, and that the labels differ from (18).

$$
\begin{aligned}
F^*(j) - F^*(4) = {} & \pi_1 R(a_j - a_4) + \pi_2 R^2(a_j^2 - a_4^2) + \pi_3((x+h_j)^2 - x^2) \\
& + \pi_4(t+1)R(a_j - a_4) + \pi_5 R(a_j(x+h_j) - a_4 x) \\
& + \pi_6 \rho\varepsilon R(a_j - a_4) + \pi_7 RE(a_j - a_4) + \pi_8 \rho\varepsilon h_j \\
& + \pi_9 E h_j + \pi_{10}[(x+h_j)^3 - x^3] + \pi_{11}(t+1)^2 h_j \\
& + \pi_{12}(t+1)[(x+h_j)^2 - x^2] + \pi_{13} E[(x+h_j)^2 - x^2] \\
& + \pi_{14} E^2 h_j + \pi_{15}(t+1)E h_j \\
& + \pi_{16}\chi(\text{part-time}) + \pi_{17}\chi(\text{full-time}) + \pi_{18}\chi(\text{overtime}).
\end{aligned}
$$

Appendix B. Existence of joint posterior distribution

In the following I use notation as it was developed in the body of the text. In addition, let N_w, N_y and N_z denote the cardinal number of unobserved wages, unobserved wealth and latent relative alternative values, respectively. Then let $\Omega \equiv [A_\Omega, B_\Omega]^{N_w}$, $\Gamma \equiv [A_\Gamma, B_\Gamma]^{N_y}$ and $\Delta \equiv [A_\Delta, B_\Delta]^{N_z}$ denote the domain of unobserved wages, unobserved wealth and latent relative alternative values, respectively. The bounds on the domains are such that $0 < A_\Omega < B_\Omega < \infty$, $-\infty < A_\Gamma < B_\Gamma < \infty$ and $-\infty < A_\Delta < B_\Delta < \infty$. Next, let B denote the domain of $\beta \times h_u$, so that $B = \Re^6 \times \Re_{++}$, and set $V \equiv [-1, 1]$ as the domain of ρ, which reflects the prior specified in Section 3.5. Recall the utility function and future component coefficients are denoted jointly by Λ, and denote the domain of $\Lambda \times H_\eta$ by T, where T is chosen so that each element of Λ lies along the real line, and $H_\eta \in T \Rightarrow H_\eta$ is positive definite. Let $G = G_1 \times \Re_{++}$ represent the domain of $\gamma \times h_\alpha$, and recall that G_1 is the compact subset of $\Re^6$ implied by the priors imposed in Section 3.4. Let W be the vector of observed and unobserved wages, Υ the vector of observed and unobserved wealth, Z the vector of all latent relative utilities and D the vector of all choice indicators. With this notation it is convenient to define the following three functions.

$$
\begin{aligned}
& f(\Lambda, H_\eta, \gamma, W, \Upsilon, Z, D) \\
& \quad = |H_\eta|^{NT-4/2} \exp\left\{ -\frac{1}{2} \sum_{n,t} \begin{pmatrix} \{z_{n1t} - Q'_{n1t}\Lambda\}_{n,t} \\ \{z_{n2t} - Q'_{n2t}\Lambda\}_{n,t} \\ \{z_{n3t} - Q'_{n3t}\Lambda\}_{n,t} \end{pmatrix}' \right. \\
& \quad \left. \times H_\eta \begin{pmatrix} \{z_{n1t} - Q'_{n1t}\Lambda\}_{n,t} \\ \{z_{n2t} - Q'_{n2t}\Lambda\}_{n,t} \\ \{z_{n3t} - Q'_{n3t}\Lambda\}_{n,t} \end{pmatrix} \right\} I(Z, D),
\end{aligned}
$$

$$g(\gamma, h_\alpha, W, \Upsilon, D) = h_\alpha^{(N(T-1)-2)/2} \exp\left\{-\frac{h_\alpha}{2}\sum_{n,t>1}(A_{nt} - Ra(n, j_{t-1}^*, t-1|\gamma))^2\right\},$$

$$h(\beta, h_u, \rho, W) = (1-\rho^2)^{N/2} h_u^{(NT-2)/2}$$
$$\times \exp\left[-\frac{h_u(1-\rho^2)}{2}\left\{\begin{array}{l}\sum_n\{(\log w_{n1} - X_{n1}'\beta)^2 + \sum_{t>1}(\log w_{nt} \\ \quad -\rho\log w_{nt-1} - X_{nt}'\beta + \rho X_{nt-1}'\beta)^2\}\end{array}\right\}\right]\prod_{n,t}\frac{1}{w_{nt}}.$$

The goal is to show

$$\int_\Omega\left\{\int_{\Gamma,G}\left[g(\gamma, h_\alpha, W, \Upsilon, D)\int_{T,\Lambda} f(\Lambda, H_\eta, \gamma, W, \Upsilon, Z, D)\right]\int_{B,V} h(\beta, h_u, \rho, W)\right\} < \infty.$$

Begin with the inner integral

$$\int_{B,V} h(\beta, h_u, \rho, W).$$

Note that $h(\cdot)$ can be expressed

$$(1-\rho^2)^{N/2} h_u^{(NT-k-2)/2}\exp\left[-\frac{h_u}{2}S_{\hat\beta}\right] h_u^{k/2}\exp\left[-\frac{h_u}{2}(\beta-\hat\beta)'X^{*'}X^*(\beta-\hat\beta)\right],$$

where k is the dimension of β, $\hat\beta = (X^{*'}X^*)^{-1}X^{*'}Y^*$, $S_{\hat\beta} = (Y^* - X^{*'}\hat\beta)'(Y^* - X^{*'}\hat\beta)$ (using notation from (44)–(47)). It follows that, for any value of $\rho \in V$, and for any unobserved wage values in Ω, the joint distribution of h_u and β is proportional to the normal-Gamma distribution (Bernardo and Smith, 1995, p. 136), hence finitely integrable over B. Let $h^*(\rho, W) = \int_B h(\beta, h_u, \rho, W)$.

Turn next to $\int_{T,\Lambda} f(\Lambda, H_\eta, \gamma, W, \Upsilon, Z, D)$. We can express

$$f(\Lambda, H_\eta, \gamma, W, \Upsilon, Z, D)$$
$$= |H_\eta|^{(NT-4)/2}\exp\left\{-\frac{1}{2}tr(S(\hat\Lambda)H_\eta) - \frac{1}{2}(\Lambda-\hat\Lambda)'(\mathbf{Q}'(H_\eta\otimes I_{NT})\mathbf{Q})(\Lambda-\hat\Lambda)\right\} I(Z, D),$$

where

$$\hat\Lambda = (\mathbf{Q}'(H_\eta\otimes I_{NT})\mathbf{Q})^{-1}\mathbf{Q}'\mathbf{Z}$$

and

$$\mathbf{Z} = \left\{\begin{array}{c}\{z_{n1t}\}_{nt}\\ \{z_{n2t}\}_{nt}\\ \{z_{n3t}\}_{nt}\end{array}\right\},\quad \mathbf{Q} = \left\{\begin{array}{c}\{Q_{n1t}\}_{nt}\\ \{Q_{n2t}\}_{nt}\\ \{Q_{n3t}\}_{nt}\end{array}\right\}.$$

Hence, given any constellation of (γ, W, Υ, Z) such that $\gamma \in G_1$, $Z \in \Lambda$, unobserved wages lie in Ω, unobserved wealth values lie in Γ and $I(Z, D) > 0$, we have that the joint density of H_η and Λ is normal-Wishart, hence finitely integrable over T. Let $f^*(\gamma, W, \Upsilon, Z, D) = \int_T f(\Lambda, H_\eta, \gamma, W, \Upsilon, Z, D)$.

Turning to $g(\gamma, h_\alpha, W, \Upsilon, D)$, it is clear that for any constellation of wages in Ω, wealth in Γ and parameters $\gamma \in G_1$, the precision h_α has a gamma distribution. It follows that for any $\gamma \in G_1$, $g^*(\gamma, W, \Upsilon, D) \equiv \int_{h_\alpha > 0} g(\gamma, h_\alpha, W, \Upsilon, D) < \infty$.

Therefore, we need only show that

$$\int_{\Omega,\Gamma,G_1,\Delta,V} g^*(\gamma, W, \Upsilon, D) f^*(\gamma, W, \Upsilon, Z, D) h^*(\rho, W) < \infty.$$

Since $\Omega, \Gamma, \Delta, G_1$, and V are each compact, and since g^*, f^* and h^* are each bounded everywhere within these sets, it follows that the integral is exists and is finite.

References

Altonji, J.G., 1986. Intertemporal substitution in labor supply: evidence from micro data. Journal of Political Economy 94, S176–S215.

Altug, S., Miller, R., 1990. Household choices in equilibrium. Econometrica 58, 543–570.

Altug, S., Miller, R., 1998. The effect of work experience on female wages and labour supply. Review of Economic Studies 65, 45–85.

Bernardo, J., Smith, A., 1995. Bayesian Theory. John Wiley and Sons, New York.

Blinder, A., Weiss, Y., 1976. Human capital and labor supply: a synthesis. Journal of Political Economy 84, 449–472.

Browning, M., Deaton, A., Irish, M., 1985. A profitable approach to labor supply and commodity demands over the life-cycle. Econometrica 53, 503–543.

Card, D., 1994. Intertemporal labor supply: an assessment. In: Sims, C. (Ed.), Advances in Econometrics. Cambridge University Press, Cambridge, pp. 49–78.

Eckstein, Z., Wolpin, K., 1989. Dynamic labor force participation of women and endogenous work experience. Review of Economic Studies 56, 375–390.

Epstein, L., Zin, S., 1989. Substitution, risk aversion and the temporal behavior of consumption and asset Returns: a theoretical framework. Econometrica 46, 185–200.

Gelman, A., 1996. Inference and monitoring convergence. In: Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), Markov Chain Monte Carlo in Practice. Chapman & Hall, London, pp. 131–144.

Geweke, J., 1991. Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints. In: Keramidas, E.M. (Ed.), Computing Science and Statistics: Proceedings of the Twenty-Third Symposium on the interface. Interface Foundation of North America, Inc., Fairfax, pp. 571–578.

Geweke, J., 1992. Priors for macroeconomic time series and their application. Working Paper, Federal Reserve Bank of Minneapolis.

Geweke, J., Keane, M.P., 1999. Bayesian inference for dynamic discrete choice models without the need for dynamic programming. In: Mariano, R.S., Schuermann, T., Weeks, M. (Eds.), Simulation Based Inference and Econometrics: Methods and Applications. Cambridge University Press, Cambridge.

Geweke, J.F., Keane, M.P., Runkle, D., 1994. Alternative computational approaches to inference in the multinomial probit model. The Review of Economics and Statistics 76, 609–632.

Geweke, J., Houser, D., Keane, M.P., 2001. Simulation based inference for dynamic multinomial choice models. In: Baltagi, B. (Ed.), A companion to Theoretical Econometrics. Blackwell, Oxford, pp. 466–493.

Heckmann, J.J., Sedlacek, G., 1985. Heterogeneity, aggregation and market wage functions: An empirical model of self-selection in the labor market. Journal of Political Economy 93, 1077–1125.

Hotz, J., Kydland, F., Sedlacek, G., 1988. Intertemporal substitution and labor supply. Econometrica 56, 335–360.

Hotz, V.J., Miller, R.A., 1993. Conditional choice probabilities and the estimation of dynamic models. Review of Economic Studies 60, 497–529.

Hotz, V.J., Miller, R.A., 1988. An empirical analysis of life cycle fertility and female labor supply. Econometrica 56, 91–118.

Houser, D., Winter, J., 2001. How do Behavioral Assumptions Affect Structural Inference? Evidence from the Laboratory. University of Arizona, Arizona.

Houser, D., Keane, M., McCabe, K., 2001. Behavior in a Dynamic Decision Problem: Evidence from the Laboratory. University of Arizona, Arizona.

Keane, M., 1994. A computationally practical simulation estimator for panel data. Econometrica 62, 95–116.

Keane, M.P., Wolpin, K.I., 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. Review of Economics and Statistics 76, 648–672.

Keane, M.P., Wolpin, K.I., 1997. The career decisions of young men: a dynamic model of human capital investment over the life-cycle. Journal of Political Economy 105, 473–522.

MaCurdy, T., 1981. An empirical model of labor supply in a life-cycle setting. Journal of Political Economy 89, 1059–1085.

MaCurdy, T., 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.

Manski, C.F., 1993. Dynamic choice in social settings: Learning from the experience of others. Journal of Econometrics 58, 121–136.

Miller, R.A., Sanders, S.G., 1997. Human capital development and welfare participation. Carnegie–Rochester Conference Series on Public Policy 46, 1–45.

Mincer, J., 1958. Investment in human capital and personal income distribution. Journal of Political Economy 66, 281–302.

Mincer, J., 1974. Schooling Experience and Earnings. National Bureau of Economic Research, New York.

Pencavel, J., 1986. Labor supply of men: a survey. In: Ashenfelter, O., Layard, R. (Eds.), Handbook of Labor Economics, Vol. 1. North Holland, Amsterdam, pp. 3–44.

Roy, A.D., 1951. Some thoughts on the distribution of earnings. Oxford Economic Papers 3, 135–146.

Rust, J., 1987. Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. Econometrica 55, 999–1033.

Rust, J., 1994. Estimation of dynamic structural models, problems and prospects: Discrete decision processes. In: Sims, C. (Ed.), Advances in Econometrics, Sixth World Congress, Vol. II. Cambridge University Press, Cambridge, pp. 119–170.

Shaw, K., 1989. Life-cycle labor supply with human capital accumulation. International Economic Review 30, 431–456.

Tanner, M.A., Wong, W.H., 1987. The calculation of posterior distributions by data augmentation. Journal of the American Statistical Association 82, 528–550.

Tierney, L., 1994. Markov chains for exploring posterior distributions. The Annals of Statistics 22, 1701–1762.

Willis, R., 1986. Wage determinants: a survey and reinterpretation of human capital earnings functions. In: Ashenfelter, O., Layard, R. (Eds.) Handbook of Labor Economics, Vol. 1. North Holland, Amsterdam, pp. 525–602.

Willis, R., Rosen, S., 1979. Education and self-selection. Journal of Political Economy 87, S7–S36.

Wolpin, K., 1984. An estimable dynamic stochastic model of fertility and child mortality. Journal of Political Economy 92, 852–874.

Zellner, A., 1971. An Introduction to Bayesian Inference in Economics. John Wiley and Sons, New York.