Intertemporal Labor Supply and Human Capital Accumulation*

Susumu Imai†
The Pennsylvania State University

April 15, 2000

Abstract

There has been considerable interest in both labor economics and macro economics in estimating the intertemporal elasticity of substitution for labor supply. In this paper, I solve and estimate a dynamic model that allows agents to optimally choose their labor hours and consumption over a continuum of positive real numbers, and that allows for both human capital accumulation and savings. Estimation results and simulation exercises indicate that the intertemporal elasticity of substitution is much higher than those estimated by MaCurdy (1981) or Altonji (1985), and that their estimates are biased downwards because of the omission of the human capital accumulation eœect. The human capital accumulation eœect renders the life-cycle path of the shadow wage relatively Æat, even though wages increase signiøcantly with age. Hence, a rather Æat lifecycle labor supply path can be reconciled with a high intertemporal elasiticity of substitution.

*This paper is taken from my dissertation at the University of Minnesota. I am grateful to my thesis advisor Mike Keane, as well as to John Geweke, Richard Rogerson, and Daniel Houser for many useful comments.

†Address:608, Kern Graduate Building, The Pennsylvania State University, University Park, PA, 16802, e-mail: sxi5@psu.edu

1 Introduction

The degree of intertemporal substitution in labor supply has been a topic of considerable interest in both labor and macro economics for at least the past 30 years (see, e.g., Lucas and Rapping (1969)). Recently, there have been several studies that address the question using micro panel data. Classic examples are MaCurdy (1981) and Altonji (1985). They focused on estimating the intertemporal elasticity of substitution in labor supply, using marginal utility of wealth constant labor supply functions. In their work they assume that the utility function is time separable and that wages are exogenous.

But if current labor supply leads to human capital accumulation, then estimates of the intertemporal elasticity of substitution in labor supply under a false assumption of no human capital accumulation are likely to be biased towards zero. The reason is as follows: As the wage increases over the life-cycle, the substitution effect induces labor supply to increase, thus providing an incentive for people to supply more labor in older age. On the other hand, both concavity of the value function with respect to human capital and the approaching retirement period lower the marginal rate of return to human capital investment, thus reducing the incentive to supply labor. If these two factors roughly cancel, then even if wages increase over the life-cycle, labor supply will be little changed (see Fig. 1). But if we only allow for the substitution effect and not the human capital effect, then we will falsely conclude that the intertemporal elasticity of labor supply is low simply from the fact that labor supply remains roughly constant over the life-cycle even though wages increase.

In this paper, I address the issue of human capital accumulation in two steps. In the first step, I estimate the life-cycle labor supply model using maximum likelihood estimation based on a full solution of agents' dynamic programming problem that allows human capital accumulation. In the estimation, I use the male sample in the NLSY data. My estimates of the disutility of labor parameter implies that the elasticity of the intertemporal substitution of labor supply is 3.718, which is quite comparable to results obtained and used in the macro literature. In the second step, I simulate data from the estimated model. Using the simulated data, I estimate consumption and labor supply Euler equations like those used in MaCurdy (1981) and Altonji (1986), which do not allow for human capital accumulation. The elasticity estimates I obtain from the simulated data using the MaCurdy and Altonji estimation methods range from -0.3 to 2.4 when I use all the simulated data, and from -0.1

to 2.4, after I removed the outliers from the simulated data using a similar procedure to MaCurdy (1981). These estimates are signiøcantly lower than the ML estimates. OLS and IV results from the NLSY data are also reported and there the intertemporal elasticity of labor supply is again estimated to be low. The high elasticity obtained by full solution estimation and the contrasting low elasticity implied by the conventional estimates indicate that the latter are signiøcantly biased towards zero, and one of the main reasons behind this is the omission of human capital accumulation. Thus, my results may explain the apparent contradiction between the macro and micro literature noted above.

Notice that in conventional methods of estimation such as MaCurdy (1981) or Altonji (1986), the intertemporal elasticity of substitution in labor supply is deøned and estimated as the elasticity of substitution when workers change labor supply along the anticipated life cycle wage path. But in macroeconomics, the discussion is about how labor supply responds to unanticipated business cycle shocks. In my analysis, I explicitly solve the dynamic programming model including unanticipated wage shocks. Therefore, my estimate is more relevant in providing micro evidence for use in calibrating real business cycle models. After the estimation, I also simulate the hours response to the 2 % increase in temporary wage payment. Even thouth the elasticity of labor supply is 3.718,

Some recent work on estimating dynamic labor supply models has also included human capital accumulation. Shaw (1989) estimates the Euler equations of optimal consumption and labor supply using the nonlinear GMM method. But estimating the model using maximum likelihood has beneøts compared to using the common Euler equation GMM framework. First, in the labor supply literature it is well known (see Altonji (1986)) that the instruments used for GMM estimation are poor, and the problems (including bias) arising from use of poor instruments are pointed out by Nelson and Startz (1991) and others.

Furthermore, there are some shortcomings of GMM estimation that are particularly related to the estimation of an intertemporal labor supply model that explicitly incorporates human capital accumulation. First, because nonlinear GMM must be used, measurement error in the regressors biases the results. And labor market data are known to be contaminated with signiøcant measurement error (for a detailed discussion about measurement error in panel data, see Altonji (1986)). Also, potential nonconcavities in agents' problem as a result of human capital investment may create problems for GMM. If we think of labor supply as an input to income production, then an increase in labor supply by Δ percent in every period results in an increase in wage income of more than Δ percent. That is because an increase in labor supply also raises future wages via human capital accumulation. Hence, there are increasing returns to scale in the income generating process and potential nonconvexities in the model. In such a case, just looking at the ørst order condition may not be suŒcient to claim that agents are solving the intertemporal optimal labor supply problem. In my estimation process, I explicitly solve the continuous variable dynamic programming problem, and embed the solution in the maximum likelihood estimation. Hence, my solution truly assumes that individuals choose optimal labor supply and consumption over the life-cycle.

Another key advantage of the ML is that it requires us to fully specify the shock distributions in the model, so that the model may be simulated. In contrast, when we estimate Euler conditions what we recover is the sample ıexpectation errorȷ, and not the distribution of the underlying shocks in the model. It is misleading to try to simulate the model using the sample expectation error, since the model usually imposes complicated dynamic relationships between the sequence of sample expectation errors, which we cannot recover unless we estimate the model by explicitly solving it. Researchers are interested in parameter estimates in intertemporal labor supply models because they use those parameter estimates for simulating the impact of various policy experiments. Here, I not only provide parameter estimates, but also simulate the model and conduct some policy experiments based on the model. The simulation is done by drawing from the error distribution of the shock in the model, which was explicitly incorporated in the likelihood function which I maximized.

Eckstein and Wolpin (1991) also include human capital in a life cycle labor supply model that they estimate by ML. But they restrict the labor hours to zero hours or full time. If most people intertemporally substitute labor but still stay within the same classiøcation (that is, either full time or zero hours), the estimate of the intertemporal elasticity of labor in the discrete choice case will be downwardly biased. If most people work around the borders of two categories, then small changes in labor supply will be classiøed as changes between two categories; thus the labor supply elasticity will be overestimated. Since the former case seems to be more likely in a discrete choice model of labor supply, one is likely to get downwardly biased estimates of the intertemporal elasticity of substitution of labor. Also, since they use a linear utility function for consumption, there is no wealth eœect in their model, even though the wealth eœect may be an important factor linking peoples' decisions intertemporally.

As noted earlier, the major obstacle to maximum likelihood estimation is that the full solution of the continuous variable dynamic programming problem is extremely computationally demanding. There are two reasons. First, the state space of the dynamic programming problem is now inønite. Even in a discrete choice dynamic programming problem, where there is only a ønite state space, researchers are usually plagued with the problem of having too many state space points to evaluate the value function (see, e.g. Keane-Wolpin (1994)). In the continuous choice case, explicit evaluation of the value function at each state space point is impossible. Secondly, in contrast to the discrete choice dynamic programming problem where solving for the control variables is a trivial optimization over a ønite set of choices, in the continuous choice problem solving for the control variables is the main source of computational burden. It requires a two dimensional nonlinear Newton search algorithm to ønd optimal consumption and labor supply at each state point.

Here I develop an algorithm which approximates the solution to the DP problem. The algorithm successively solves the Bellman equation backwards from the last period. First, I choose a ønite set of grid points over assets and human capital at which to evaluate the value function. Then, at each human capital and asset grid point, taste shock quadrature points are chosen, and I solve for optimal consumption and labor supply using a Newton search algorithm at each point. Then, I integrate the value function over the taste shocks using Gaussian quadrature integration. Next, at each asset grid point, I use Chebychev polynomials to get a value function surface over the human capital grid points. Given the Chebychev polynomial approximation, integration of the value function with respect to the human capital shock can be done analytically. The Emax function derived as above is used for the next Bellman backward iteration. The main feature of the algorithm is that I avoid two dimensional quadrature integration by exploiting the fact that there is a one-to-one mapping from human capital to wages. Reducing the dimensionality of quadrature integration decreases the number of computationally demanding Newton searches by an order of magnitude. Also, by reducing the range of human capital points at which the value functions must be evaluated, it makes the Newton search algorithm itself easier and more accurate.

I apply my model to the 1979 youth cohort of the National Longitudinal Survey of Labor Market Experience (NLSY). Among the features of the NLSY, which distinguishes it from the commonly used Panel Study of Income Dynamics (PSID), is detailed asset data for individuals from 1985. Instead of using food consumption to derive the marginal value of wealth at some period, which is commonly done by researchers using the PSID, I can use asset data directly. I can then derive total consumption by using the asset data and the intertemporal budget constraint.

The organization of the paper is as follows: Section 2 presents the model, section 3 describes the algorithm for solving the DP problem, and section 4 describes the algorithm for forming the likelihood function. Section 5 describes the data, and section 6 discusses the estimates and some model simulations. Section 7 concludes.

2 The Model Speciøcation

In this section, I present a life-cycle model for an individual agent who rationally chooses his optimal life-cycle path for consumption and hours of labor supply.

At a given calendar time period s and age t, the agent's period utility of consumption is a concave function of the consumption of market goods, C, and the disutility of labor is a convex function of the hours of labor supply h and the taste shock ϵ_2. Preferences are additively separable over time. Agents choose the optimal consumption and labor supply by maximizing their discounted expected life-cycle utility over the working horizon T, which is

$$E_t \sum_{\tau=t}^{T} \beta^{\tau} [U(C_\tau, \tau) - V(h_\tau, \epsilon_{2,\tau}, \tau)] \tag{1}$$

Agents also face an intertemporal budget constraint and a human capital production constraint.

The intertemporal budget constraint is

$$A_{t+1} = (1 + r)A_t + W_{t,s}h_t - C_t. \tag{2}$$

where A_t is the agent's asset holdings at age t and r is the interest rate. The observed wage $W_{t,s}$ at age t, time s is deøned as the product of the human capital stock K_t times the rental rate on a unit of human capital, R_s.

$$W_{t,s} = R_s K_t \tag{3}$$

The rental rate R_s is the market price of the services of a unit of human capital. I assume a perfect market for human capital. Hence, at any time s, all agents face the same rental rate R_s.

Human capital evolves according to the human capital production function, which is a deterministic function of current labor supply hours h_t, current human capital K_t, and age t, along with the multiplicative wage shock, ε_t. That is,

$$K_{t+1} = g(h_t, K_t, t)\epsilon_{1,t+1}. \quad (4)$$

K_{t+1} is the age t human capital after the wage shock $\epsilon_{1,t+1}$ is realized.

At age t, time period s, the agent's decision process can be described as the following maximization of the value function.

$$\begin{aligned} V_{t,s}(A_t, K_t, \epsilon_t) &= max_{C_t,h_t}\{U(C_t, t) - V(h_t, \epsilon_{2,t}, t) \\ &+\beta E_t V_{t+1,s+1}[(1 + r)A_t + R_s K_t h_t - C_t, f(h_t, K_t, \epsilon_{1,t+1}, t), \epsilon_{t+1}]\} \quad (5) \end{aligned}$$

For the utility function for consumption C, I choose a CRRA form augmented to include age eœects:

$$U(C, t) = C(t)\frac{C^{a_1}}{a_1}$$

where $C(t)$ is a spline in age.[1]

[1] In the data, wages and labor supply are relatively small on average when people are young compared to when they are old. Hence, annual labor income is small when young. Therefore, if consumption is roughly constant, people should be in debt when young, and repay it in later years of the life cycle. But this is not the case in the data. (See Table 3 for mean age prøles of wage, hours and assets) Age eœects in consumption are necessary to explain the positive asset holdings when young observed in the data. To capture this, the utility of consumption should be smaller when young, so that consumption rises over time, and individuals won't go into heavy debt early in life.

The function $C(t)$ is spline in age with break points at age 25, 33, and 47. Speciøcally,

$$C(t) = C_0[C_1 + C_2\frac{(t-20)}{5}]$$

if

$$t - 20 \leq 5$$

and

$$C(t) = C_0[C_1 + C_2 + (C_3)\frac{(t-25)}{8}]$$

if

$$5 < t - 20 \leq 13$$

and

$$C(t) = C_0[C_1 + C_2 + C_3 + (1 - C_1 - C_2 - C_3)\frac{(t-33)}{14}]$$

The disutility of labor, which is a function of hours h, is assumed to have the following functional form.

$$V(h, \epsilon_1) = \epsilon_1 b(t) \frac{h^{a_2}}{a_2} \tag{6}$$

where $b(t)$ is an age eœect, which I assume to be as follows.

$$b(t) = b_0(1.0 + b_1 \min\{t - 20, 4\} + b_2\sqrt{t - 20})$$

Since, as we will see in section 6 , the estimate of b_1 is negative, this implies that the marginal disutility of labor initially decreases with age until the age of 24. This helps explain the slight observed increase in labor supply hours until the age 24, which the model was unable to explain based on the wage path alone. The marginal disutility of labor then increases over time since the estimates of b_2 turns out to be positive. Keane and Wolpin (1997) also ønd that some age eœects are necessary to reconcile life cycle labor supply patterns with the human capital investment model that they estimate.

Except for the added age eœects, the functional forms are adopted from MaCurdy (1981) and Altonji (1986). That will enable me to compare my results and theirs in later estimation and simulation exercises. Furthermore, in this parameter speciøcation, the degree of intertemporal substitution of labor supply without human capital eœects can be summarized by a single intertemporal elasticity of labor supply parameter, which is

$$b_2 = \frac{1}{a_2 - 1}$$

By introducing intratemporal nonseparability of consumption and labor supply, one could explain the consumption proøles in the data without resorting to the age eœects in consumption. But that would make the results of the estimation and simulation exercises less comparable to the results by MaCurdy (1981) and Altonji (1986).

---

if

$$13 < t - 20 \leq 27$$

and

$$C(t) = C_0$$

if

$$27 < t - 20$$

The age eœect C_t starts at $(C_0 C_1)$ at age 20, and moves up to C_0 at age 47, with a kink at age 28,and 33.

Many empirical papers analyzing intertemporal labor supply behavior, such as Shaw (1989) or Hotz, et. al. (1988), use a translog function of consumption and leisure as the utility function. While this approach has the advantage of being locally Æexible, none of the parameters can be straightforwardly interpreted as describing the intertemporal elasticity of labor supply. Hence, from their estimation results, it is diŒcult to draw any conclusions about how much people substitute labor intertemporally, unless one simulates their estimated models. But as discussed above, their estimation results are not well suited for simulation exercises because they use Euler equation methods that do not involve full solution of the model.

I assume the human capital production function to be as follows.

$$g(K, h, t) = k_0 + \delta K + G(K, h, t) \tag{4}$$

where $G()$ is a function of current human capital, labor hours and age t. Fig. 2 gives some evidence on the shape of the human capital production function from the NLSY data. The øgure shows the relationship between current labor supply hours and the next period wage, within diœerent cells for the current wage. Each cell has length 2 dollars. From the graph, we can see that the function relations of future wages to labor hours has a higher slope when the current wage is higher. That implies there is a signiøcant complementarity between current wages and current hours in terms of learning by doing. We can capture this complementarity by parameterizing $G(K, h, t)$ to be

$$G(K, h, t) = A_0(B_1 + K)h^{\alpha}$$

Also notice that the slope of the curves relating future wages to current hours are bounded at $h = 0$. That is, the derivative of the human capital production function with respect to hours around $h = 0$ is likely to be bounded. We can capture that by introducing the intercept term d_1.

$$G(K, h, t) = A_0(B_1 + K)(h + d_1)^{\alpha}$$

Furthermore, for very large hours, the slopes of the lines on the graph seem to be close to zero or even negative. Thus, there is a possibility that for very large hours, increases in hours have no eœect on future human capital or even decrease future human capital. We can account for this possibility by adding another term to $G()$ as follows.

$$G(K,h,t) = A_0(B_1 + K)[(h+d_1)^\alpha - B_2(h+d_1)]$$

For estimation, I also include a pure age effect in the human capital production function, similar to the age effect included in the human capital earnings function by Keane and Wolpin (1997). Thus, I adopt the following functional form.

$$G(K,h,t) = (A_0 + A_1(t-20))(B_1 + K)[(h+d_1)^\alpha - B_2(h+d_1)] \quad (4)$$

The wage and taste shocks are assumed to have i.i.d. mean 1 log normal distributions. That is,

$$ln(\epsilon_i) \sim N(-\frac{1}{2}\sigma_i^2, \sigma_i), i = 1,2 \quad (7)$$

I also allow for the measurement errors in wages, labor supply hours and assets. I defer the discussion of the measurement error functional forms until section 4.

I set the working horizon T at age 65. At the terminal period, I have the following parameterization for the end of period value function.

$$V_{T+1}(A_{T+1}) = \begin{cases} (3log(A_{T+1}+\phi) - 1 - 3log(\phi) & \text{if } A_{T+1} > 0 \\ (\frac{A_{T+1}-\phi}{\phi})^3 & \text{otherwise} \end{cases} \quad (8)$$

I chose this specification because this function is continuously differentiable in assets and the derivative is decreasing in assets. It turns out that the coefficient ϕ is difficult to estimate, because the NLSY only has data on individuals until the age 35. After some experimentation, I set ϕ to be 100,000. The reasons that agents get positive value from having assets at age 65 are, for example, that they are able to enjoy consumption until their death, and possibly can leave bequests to their heirs.

Now, to understand the effect of introducing human capital accumulation on the hours response to wage changes, I consider the first order conditions of the above problem with respect to consumption and labor. These are:

$$U_C(C_t,t) - \beta E_t V_{A,t+1,s+1}(A_{t+1}, K_{t+1}, \epsilon_{2,t+1}) = 0$$

$$-V_h(h_t, \epsilon_{2,t}, t) + R_s K_t U_C(C_t,t) + \beta E_t f_h V_{t+1,s+1,K}(A_{t+1}, K_{t+1}, \epsilon_{2,t+1}) = 0.$$

Notice that the current marginal disutility of labor equals the wage times the marginal utility of consumption, which is the marginal return to increases in current wage income due to increases in labor supply, plus an extra term that captures the marginal return to increases in future human capital. As the wage increases over the life-cycle, the substitution eœect induces labor supply to increase, thus providing an incentive for people to supply more labor in older age. On the other hand, both concavity of the value function with respect to human capital and the approaching retirement period lower the marginal rate of return for human capital investment, thus reducing the incentive to supply labor. If these two factors roughly cancel, then even if wages increase over the life cycle, labor supply will be little changed. (See Fig. 1).

Readers who are not interested in the algorithms for solving the dynamic programming problem and forming the likelihood can skip the next two sections and go directly to section 5.

3 Solving the Continuous Stochastic Dynamic Programming Problem

As discussed before, the problem agents solve in each period is as follows:

$$\begin{aligned} V_{t,s}(A_t, K_t, \epsilon_{2,t}) &= max_{C_t,h_t}\{U(C_t, t) - V(h_t, \epsilon_{2,t}, t) \\ &\quad + \beta E_t V_{t+1,s+1}(A_{t+1}, K_{t+1}, \epsilon_{2,t+1})\} \end{aligned} \tag{5'}$$

subject to the intertemporal budget constraint,

$$A_{t+1} = (1 + r)A_t + R_s K_t h_t - C_t \tag{2}$$

and human capital production function

$$K_{t+1} = g(h_t, K_t, t)\epsilon_{1,t+1} \tag{4}$$

where ϵ_{t+1} is the human capital shock realized at age $t + 1$.

Notice that the next period's human capital K_{t+1} is not known to the individual at age t. Let us rewrite the Bellman equation in terms of variables that the agent knows at age t. Deøne $\tilde{K}_{t+1}$ to be the next period human capital before the human capital shock is realized. That is,

$$\tilde{K}_{t+1} = g(h_t, K_t, t)$$

$$K_{t+1} = \tilde{K}_{t+1}\epsilon_{1,t+1}.$$

Deøne the value function in terms of $\tilde{K}$ as follows

$$\begin{aligned} V_{t,s}(A_t, \tilde{K}_t, \epsilon_{1,t}, \epsilon_{2,t}) &= max_{C_t,h_t}(U(C_t,t) - V(h_t, \epsilon_{2,t}, t) \\ &\quad +\beta E_t V_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1}, \epsilon_{1,t+1}, \epsilon_{2,t+1})). \end{aligned}$$

Also, deøne the Emax function V^E as follows:

$$V^E_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1}) = E_t V_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1}, \epsilon_{1,t+1}, \epsilon_{2,t+1})$$

Then, the above problem can be rewritten as follows:

$$V_{t,s}(A_t, \tilde{K}_t, \epsilon_{1,t}, \epsilon_{2,t}) = max_{C_t,h_t}\{U(C_t,t) - V(h_t, \epsilon_{2,t}, t) + \beta V^E_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1})\}$$

subject to the intertemporal budget constraint,

$$A_{t+1} = (1+r)A_t + R_s K_t h_t - C_t$$

and human capital production function

$$\tilde{K}_{t+1} = g(h_t, K_t, t)$$

There are several computational obstacles to solving the continuous stochastic dynamic programming problem that I assume the agents are facing. In order to numerically solve the above problem, in general, we have to start at the terminal period, T, and backsolve to $t = \theta$, where θ is the start of the planning period (assumed to be age 20).

Now, let the state space for the Emax function be $(A_t, \tilde{K}_t)$. Suppose I have already solved for the Emax function for age $t+1$. That is, I have already calculated the Emax function

$$V^E_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1}) = E_t V_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1}, \epsilon_{1,t+1}, \epsilon_{2,t+1}).$$

for all possible values of A_{t+1} and $\tilde{K}_{t+1}$. The next step in the backsolving process is to ønd the $V^E_{t,s}(A_t, \tilde{K}_t)$. Given the state space point, $(A_t, \tilde{K}_t)$, I need to derive the integral of $V_{t,s}(A_t, \tilde{K}_t, \epsilon_{1,t}, \epsilon_{2,t})$ with respect to $\epsilon_{1,t}$ and $\epsilon_{2,t}$. Furthermore, in integrating, for each shock (ϵ_1, ϵ_2), I need to ønd the optimal consumption and labor supply to derive the value $V_{t,s}(A_t, \tilde{K}_t, \epsilon_{1,t}, \epsilon_{2,t})$.

To get an idea of the magnitude of the computational problem involved in solving this model, assume that agents only have 2 possible choices of consumption and labor supply, that is, a total of 4 discrete choices in each period. Also, suppose that there is no taste shock or human capital shock that I need to integrate over. Then, because the state variables in future periods depend on these choices, for a discrete choice dynamic programming problem with $T - 20$ time periods, I need to evaluate the value function at least $4^{(T-20)}$ state space points. I have 20 time periods, this amounts to at least $1.099511D + 12$ points. Suppose on the other hand, I discretize the state space of assets and human capital into n_A times n_K grid points. Further suppose that at each grid point, I evaluate the value function with respect to n_1 times n_2 combinations of human capital shocks and taste shocks, and integrate over the shocks to get the Emax function. Then the required number of evaluations of the value function is at least $n_A \times n_K \times n_1 \times n_2 \times (T - 20) \times 4$, which is again extremely computationally demanding even with modest numbers of grid and quadrature points. In continuous choice dynamic programming problems, the state space is continuous, and hence the number of state space points is inønite. Therefore, explicit evaluation of the value function at each state space point $(A_t, \tilde{K}_t)$ is impossible.

Furthermore, compared to discrete choice dynamic programming models, where optimization over the control variable only involves maximizing over a discrete set of choices, in the continuous choice problem I examine here, ønding optimal consumption and labor supply requires a two dimensional non-linear search at each state point.

Here, in order to cope with the computational problem, I will use a set of approximation and interpolation methods. First, I only explicitly solve for the value functions at a ønite set of asset and human capital grid points. The value functions at the remaining points are derived by Chebychev polynomial least squares interpolation.

At each asset and human capital grid point $(A_i, \tilde{K}_j)$ the value function is an integral over the taste shock and the wage shock. A straightforward way of performing these integrations is to use two dimensional quadrature. That means, given the grid point $(A_i, \tilde{K}_j)$, calculate the quadrature points for the shocks,

$$(\epsilon_{1,i_q}, \epsilon_{2,j_q})$$

$$i_q = 1, ..., n_q, j_q = 1, ..., n_q$$

and then, solve for the optimal consumption and labor supply to get $V_{t,s}(A_t, \tilde{K}_t, \epsilon_{1,i_q}, \epsilon_{2,j_q})$.

Then, I can use quadrature to integrate over the value function to get the Emax function. That is, form

$$V_{t,s}^{E}(A_i, \tilde{K}_j) = \int\int V(A_i, \tilde{K}_j, \epsilon_1, \epsilon_2) d\epsilon_1 d\epsilon_2 \approx \sum_{i_q, j_q} V_{t,s}(A_i, \tilde{K}_j, \epsilon_{i_q}, \epsilon_{j_q}) w_{i_q} w_{j_q}$$

where w_i are the weights for the quadrature integration.

But this approach is still computationally extremely demanding for several reasons, which are mainly due to the fact that I need to apply two dimensional quadrature integration. First, I still need to evaluate the value function over asset and human capital grid points, and wage and taste shock quadrature points. That means, I need to conduct 2 dimensional Newton search routines at $n_A \times n_K \times n_q \times n_q \times (T - 20)$ points, where n_A, n_K are the number of asset and human capital grid points, and n_q is the number of quadrature points for taste shocks and human capital shocks. And it is the two dimensional Newton search routine to ønd optimal consumption and labor supply which is by far the most computationally demanding part of the algorithm. Secondly, as I increase the dimension of the quadrature integration, even if I keep the quadrature points per dimension the same, I experience a decrease of the accuracy of the integration. Hence, if I wish to integrate over two dimensions and still have comparable accuracy to one dimensional quadrature integration with n quadrature points, it is in general likely that I will need more than $n \times n$ quadrature points.

But these diŒculties are minor compared to the problem of controlling for the range of K. Human capital $K = \tilde{K}\epsilon_{1,i_q}$ can take on very small values if both $\tilde{K}$ and ϵ_{1,i_q} are small, and can take on very large values if both $\tilde{K}$ and ϵ_{1,i_q} are large. The Newton search routine at very low or high values of human capital is both disproportionately time consuming and inaccurate compared to Newton search at other points.

To avoid the quadrature integration with respect to the human capital shocks, I add another interpolation and approximation step that exploits the fact that there is a one-to-one mapping from human capital to wages. I explain the algorithm in detail in the appendix.

4 Maximum Likelihood Estimation

I use NLSY data to estimate the parameters of the model. There are several features of the data which I need to consider when I estimate the model. First, as in most other panel data, variables such as wages, labor supply and assets are measured with error. Hence, the estimation procedure should incorporate a measurement error component in those variables. Second, there are periods where assets are missing. Hence, during the estimation process, I need to account for the missing asset data. In my likelihood function, I take into account both of these problems.

Suppose that for age t, period s, the true wage $W_{t,s} = R_s K_t$, labor supply and assets $(W_{t,s}, h_t, A_t)$ are observed with measurement error. I denote by $v_t = (v_{1,t}, v_{2,t}, v_{3,t})$ the vector of the measurement error in labor income, hours of labor supply, and assets, respectively. I assume that the wage measurement error is log normally distributed with mean 1. That is,

$$ln(v_{1,t}) \sim N(-\frac{1}{2}\sigma_{v,1}^2, \sigma_{v,1})$$

I assume that the hours measurement error is normally distributed as follows:

$$v_{2,t} \sim N(0, \sigma_{v,2})$$

Furthermore, for the measurement error of assets, I assume the following.

$$v_{3,t} \sim N(0, \sigma_{v,3})$$

$$\sigma_{v,3} = \sigma_{v,3,1} + \sigma_{v,3,2} t$$

And in order to øll in the missing ørst period assets, I assume they are distributed as follows.

$$A_{t_0}^m \sim N(\bar{A}, V_{\bar{A}})$$

Also, I assume that

$$R_s = 1$$

for all periods s. Hence,

$$W_{t,s} = W_t = K_t$$

Now, let us denote the observed variables as $\{W_t^D, h_t^D, A_t^D\}_{t=t_0}^T$ where t_0 is the starting period. Then, $W_t^D = K_t^D$. That is,

$$K_t^D = K_t \frac{v_{1,t}}{v_{2,t}}$$

$$h_t^D = h_t v_{2,t}$$

$$A_t^D = A_t + v_{3,t}$$

Finally, for the initial period wage, I assume the following measurement error.

$$K_t^D = K_t v_0$$

where

$$\ln v_0 \tilde{} N(-\frac{1}{2}\sigma_{v_0}^2, \sigma_{v_0}^2)$$

Also, the interest rate r is set to be 5%.

Here, I adopt the simulated ML method. Denote by $\{K_t^m, h_t^m, C_t^m, A_t^m\}_{t=t_0}^T$ the sequence of the true human capital, true labor supply, true consumption and true assets at the m th simulation draw. I repeat the simulation M times and derive the likelihood in the following steps.

Step 1 Simulate $\{K_t^m, h_t^m, C_t^m, A_t^m\}_{t=t_0}^T$ starting from the initial period as follows.

1) Draw the true initial true human capital $K_{t_0}^m$.

First, draw the initial period measurement error v_{0,t_0} and then, derive

$$K_{t_0}^m = \frac{K_{t_0}^D}{v_0}$$

2) Draw the true initial asset $A_{t_0}^m$.

If the initial asset is observed, then draw the measurement error v_{3,t_0} , and derive

$$A_{t_0}^m = A_{t_0}^D - v_{3,t_0}$$

If the initial asset is not observable, then draw $A_{t_0}^m$ from $N(\bar{A}, V_{\bar{A}})$.

3) Draw the taste shock

$$ln(\epsilon_2) \sim N(-\frac{1}{2}\sigma_2^2, \sigma_2),$$

and solve for the optimal consumption and labor supply. That is,

$$\begin{aligned}\{C_{t_0}^m, h_{t_0}^m\} &= argmax_{\{C_{t_0}, h_{t_0}\}}\{U(C_{t_0}, t_0) - V(h_{t_0}, \epsilon_{2,t_0}, t_0) \\ &\quad +\beta E_{t_0} V_{t_0+1}[(1+r)A_{t_0} + K_{t_0}^m h_{t_0} - C_{t_0}, \\ &\quad f(h_{t_0}, K_{t_0}^m, \epsilon_{1,t_0+1}, t_0), \epsilon_{t_0+1}]\}\end{aligned}$$

subject to

$$A_{t_0+1} = (1+r)A^m_{t_0} + K^m_{t_0}h_{t_0} - C_{t_0}.$$

$$\tilde{K}_{t_0+1} = g(h_{t_0}, K^m_{t_0}, t_0).$$

Notice that I already have the polynomial approximation of the Emax function

$$V^E(A_{t_0+1}, \tilde{K}_{t_0+1}) = E_{t_0}V_{t_0+1}(A_{t_0+1}, \tilde{K}_{t_0+1}, \epsilon_{1,t_0+1}, \epsilon_{2,t_0+1})$$

from the DP solution, which I will use in this case.

4) Draw the human capital shock ϵ_{1,t_0+1} and derive the next period state variables. That is,

$$A^m_{t_0+1} = (1+r)A^m_{t_0} + K^m_{t_0}h^m_{t_0} - C^m_{t_0}.$$

$$\tilde{K}_{t_0+1} = g(h^m_{t_0}, K^m_{t_0}, t_0).$$

$$K^m_{t_0+1} = \tilde{K}_{t_0}\epsilon_{1,t_0+1}.$$

5) Now, repeat 3) and 4) until the end period T to derive the sequence of variables $\{K^m_t, h^m_t, C^m_t, A^m_t\}^T_{t=t_0}$.

Step 2 Given the simulated sequence of variables $\{K^m_t, h^m_t, C^m_t, A^m_t\}^T_{t=t_0}$, we then derive the measurement error. Then, we calculate the log likelihood increment for person i at the m th simulation draw as follows.

Let us denote

$$\tilde{v}^m_0 = logK^D_{t_0} - logK^m_{t_0}$$

$$\tilde{v}^m_{1,t} = logK^D_t - logK^m_t + logh^D_t - logh^m_t$$

$$\tilde{v}^m_{2,t} = h^D_t - h^m_t$$

$$\tilde{v}^m_{3,t} = A^D_t - A^m_t$$

Then, we can derive $\tilde{v}^m_{1,t}$, $t = t_0+1, T$, $\tilde{v}^m_{2,t}, t = 1, T$, $\tilde{v}_{3,t}, t = 2, T$. The log likelihood increment is

$$l^m_i = \sum_{t=t_0+1}^{T}[\frac{(v^m_{1,t} + \frac{1}{2}\sigma_1)^2}{-2\sigma_1^2} - log\sigma_1 - logK^D_t]I(K^D_t, h^D_t \ observable)$$

$$+\sum_{t=t_0}^{T}[\frac{v_{2,t}^{m\ 2}}{-2\sigma_1^2} - log\sigma_1]I(h_t^D observable)$$
$$+\sum_{t=t_0+1}^{T}[\frac{(v_{3,t}^m)^2}{-2\sigma_3^2} - log\sigma_3]I(A_t^D observable)$$
$$+[\frac{(A_{t_0}^m - \bar{A})^2}{-2\sigma_{\bar{A}}^2} - log\sigma_{\bar{A}}]I(A_t^D observable)$$
$$+[\frac{(v_0^m + \frac{1}{2}\sigma_0)^2}{-2\sigma_0^2} - log\sigma_0 - logK_{t_0}^D]I(K_{t_0}^D, h_{t_0}^D\ observable)$$

We set the starting time t_0 such that both $K_{t_0}^D, h_{t_0}^D$ are observable.

Step 3 I repeat the simulation and likelihood increment calculation for $m = 1, .., M$ and derive the simulated log likelihood increment for individual i as follows.

$$l_i = log[\sum_{m=1}^{M} exp(l_i^m)]$$

The total log likelihood is

$$l = \sum_{i=1}^{N} l_i$$

As discussed earlier, the major obstacle to the maximum likelihood estimation of the continuous choice dynamic programming problem is the iterative solution of the Bellman equation, which requires a Newton Search routine for optimal consumption and labor supply at each grid point of assets and human capital and quadrature point of the taste shock. In a standard maximum likelihood routine, a single iteration requires evaluation of the likelihood and its partial derivative with respect to all the model parameters. The usual practice is to calculate the derivatives of the likelihood function numerically as follows: Suppose that the parameter vector is $\theta = (\theta_1, \theta_2, ..., \theta_n)$. Then I solve the entire DP problem to evaluate the log likelihood for θ , which is $l(\theta, X^D)$. Then, for each $i = 1, \ldots, n$, solve the DP problem and evaluate the likelihood at parameter values $(\theta_1, \theta_2, ..., \theta_i + \Delta_i, ..., \theta_n)$ where Δ_i is a small positive number. Then, the numerical derivative is:

$$\frac{\partial l(\theta, X^D)}{\partial \theta_i} = \frac{l(\theta_1, \theta_2, ..., \theta_i + \Delta_i, ..., \theta_n, X^D) - l(\theta, X^D)}{\Delta_i}$$

Now, the eœect on the value function of a marginal change in outside conditions can be decomposed into two components. The ørst component results from the change in the value function with the choice variables held constant, and the second component results from the change in the value function due to changes in the choice variables. From the envelope theorem, we know that the magnitude of the second component is of second order. Hence, as long as the changes in the parameter are small, a valid approximation for the likelihood function under the parameter value

$$\theta'^i = (\theta_1, \theta_2, ..., \theta_i + \Delta_i, ..., \theta_n)$$

is to evaluate the value functions and the likelihood, with the optimal consumption and labor supply held øxed at the values derived for the parameter value θ. Because the approximation error in the value functions is of second order, the approximation error in the likelihood function evaluation is also of second order. Hence, for one evaluation of likelihood and all its partial derivatives, the Newton search algorithm over optimal consumption and labor supply at each grid point and quadrature point only needs to be done once. Since the Newton search algorithm is the most computationally demanding part of the whole likelihood evaluation, this signiøcantly reduces the computational burden. In fact, it makes the computational cost of estimating the continuous choice model comparable to that of estimating discrete choice dynamic programming models.

5 Data

The data are from the 1979 youth cohort of the National Longitudinal Survey of Labor Market Experience (NLSY). The NLSY consists of 12,686 individuals, approximately half of them men, who were 14 to 21 years old as of January 1, 1979. The sample consists of a core random sample and an oversample of blacks, Hispanics, poor whites, and the military. One unique characteristic of the NLSY is that from 1985, it has comprehensive asset information for each respondent. In any intertemporal labor supply model, the shadow price of assets, or marginal value of assets, plays an important role as linking period by period decisions intertemporally. In the past, researchers either ørst diœerenced away the shadow price of assets, as in MaCurdy (1981), or used the marginal utility of food consumption as a proxy for the shadow price of wealth. It was necessary for researchers analyzing the PSID to use food consumption data because that is the only consumption data it contains. Here I use the asset data directly to either measure the shadow price of wealth, or, using the intertemporal budget constraint, back out total consumption.

I use the male sample of the NLSY data. In my analysis, I treat schooling as exogenous. Since people can either accumulate human capital by on the job experience or schooling, omission of the schooling decision can be an important source of bias. In order to minimize the potential bias, for each person, I only use the years beginning from the year after he last attended school. Also, I censor anybody who served in the military from the sample. Appendix A describes in more detail how I constructed the data.

Since the NLSY only has asset data beginning in 1985, and the asset data in 1991 is missing, I recover the missing assets using the intertemporal budget constraint as I discussed in the previous section.

Table 1 gives the sample means of wages, hours of labor supply, and total wealth of individuals. Also, table 2 gives the quantiles of the wage and labor supply distribution. Notice that the sample mean of the wage far exceeds the median. This indicates that there are some very high wage values. In order to remove the eœect of outliers, I removed the top and bottom 2.5 percent of the wage and hours distributions. Also, I only used assets that satisfy the following formula:

$$-1500 \times (t - 10) \leq A(t) \leq 1500 \times (t - 10)$$

where t is the age of the individual. This was necessary because there were some assets whose values were either extremely high or low. After censoring the data, the sample means are closer to the medians (see table 3). Also, in the data, the percentage of the individuals with zero hours supplied is less than 10%. Hence, for the estimation of this paper, I adopt the conventional approach of the prime age male labor supply literature and assume only interior solutions, and I exclude individuals having zero hours of work in the data from the sample. The estimation of the intertemporal labor supply model with corner solutions using the dynamic programming maximum likelihood approach is left for future research. Because of the computational burden of the dynamic programming and estimation routine, I restrict the heterogeneity to be only whether the individual graduated from high school or not. And I restricted the sample size for the estimation by randomly choosing 500 people from the data.

6 Estimation Results and some Simulation Experiments

I report the parameter estimates in table 4. The key result is the estimate of the disutility of labor parameter, which is 1.2691. The implied elasticity of intertemporal substitution is

$$b_2 = \frac{1}{a_2 - 1} = 3.718.$$

This elasticity estimate is reasonably close to the elasticity parameter macroeconomists typically use to calibrate business cycle models (Eichenbaum, Hansen and Singleton (1984) obtain an elasticity estimate that is around 5, and Prescott (1986) uses 2 in his calibration exercise.).

To evaluate the øt of the model, I artiøcially generated 500 individual life-cycle paths from ages 20 to 65 using the estimated parameters. The various age proøles in the simulated data are reported in table 5. Fig. 3 to Fig. 5 compare simulated age proøles of wages, labor supply and assets with those of the data. The simulated age proøles closely resemble the actual age proøles. Notice that the age-hours proøle before retirement is rather Æat compared to the signiøcant humped shape of the simulated age-wage proøle. The model is able to reconcile this fact with a large intertemporal elasticity of substitution because of the human capital eœect, as I have discussed earlier.

The simulated asset paths indicate that, even though labor income is small when people are young, individuals do not go into sizeable debt early in their life. In a model with perfect capital markets, that implies that for the same consumption level, the value of consumption is smaller when people are young. Alternative explanations could be the existence of some ønance constraints, or some nonseparability between consumption and labor. I did not pursue these explanations in this paper. The out of sample predictions of the model for wages, hours and assets look quite reasonable. The simulated age-hours path predicts retirement behavior of the agents at older ages, although that occurs somewhat sooner in the simulation than in reality. But later cohorts have been retiring at younger ages. It is also worthwhile to notice that the model predicts large asset accumulation around the age 60, and dissaving afterwards, which is the actual pattern of savings and dissavings (See Carrol (1997)).

The performance of the approximate DP solution can be indirectly inferred from the age proøle of the discounted marginal utility of consumption. In table 5, I report the simulated mean age proøle of $[(1.0+r)\beta]^{t-20}U_C(C_t,t)$, which should be constant over age. The proøle is roughly constant. This gives indirect evidence that the solution algorithm seems to work well overall.

Using the simulated data, I conduct OLS and IV exercises to estimate the elasticity of intertemporal substitution using the methods of MaCurdy and Altonji. That is, I estimate the following equation via OLS and IV:

$$\Delta ln(h_t) = Const + b_2 \Delta ln(W_t) + v_t$$

where v_t is the error term. From the values of estimated coeŒcient b_2, I also recover the disutility of labor parameter a_2 as follows:

$$a_2 = \frac{1}{b_2} + 1.$$

In table 6, I report the results of estimation on simulated data, and in table 7, I report the results of estimation on the simulated data that I cleaned by using an outlier elimination procedure that is similar to MaCurdy (1981)[2]. The instruments for my IV exercise include a constant term, experience, experience squared and the twice lagged wage. All the OLS and IV results are the average of 10 repetitions with independently simulated data. I used various age groups for the exercise, starting with simulation from age 20 to 65. Then, I obtained results with ages from 20 to 55, from 20 to 45, and from 20 to 35, which is the age group in the NLSY.

First, notice that for the age group $20-35$, the outlier removed simulation and data give quite similar OLS estimates of the log wage change coeŒcient. The OLS estimate from the simulated data being -0.280 and that from the actual data being -0.196. Thus, the model seems to reproduce not only the average age proøles of the data, but also the negative correlation between log wage change and log hours change in the data.

[2] The outlier elimination rules I adopt are:

1. Annual hours worked must be less than 4,680 hours.
2. For the calculation of changes in log earnings, the absolute value of the diœerence in a person's average hourly earnings in adjacent years cannot exceed \$16 or a change of 200 percent.
3. For the calculation of the changes in log hours, the absolute value of the diœerence in the annual hours of work in adjacent years cannot exceed $3,000$ hours or a change of 190 percent.

The OLS estimates of the intertemporal elasticity b_2 are smaller than the IV estimates, and both are much smaller than the ML estimate. That is, the implied intertemporal elasticity of substitution b_2 given the value for a_2 estimated by ML is much higher than the conventional estimates. The IV estimates vary with the age composition of the simulated data. It seems that if old individuals are heavily represented in the data, then the elasticity estimates tend to increase. In particular, I obtain 2.445 for the raw simulated data, and 2.385 for the simulated data without outliers, when the age group is from 20 to 65 years. For younger age groups, the elasticity estimates are much lower, for example, 0.17 for ages 20 to 35 for the original simulated data, and -0.06 for ages 20 to 35 for the outlier removed simulated data. Those results underscore the fact that the largest changes in the age-hours profile occur during retirement.

It is worth noticing that the elasticity estimates from the simulated data are low, compared to the true value regardless of whether we remove outliers or not, which indicates that the downward bias does not depend much on the outliers. Overall, these results confirm the point that we get biased estimates of the intertemporal elasticity of labor supply if we do not explicitly allow for human capital accumulation in the model. The theory implies that in the OLS case, there is another reason for downward bias because, in this case, the error term is correlated with the regressor due to the income effect. This is confirmed in my numerical example.

I also report the OLS and IV results obtained from the NLSY data that were used to obtain the ML estimates. The elasticity estimates are: 0.116 for the original NLSY data and 0.198 for the cleaned data. All estimates of the intertemporal elasticity of labor supply are much smaller than the one derived from the ML estimation.

To assess the importance of human capital accumulation for the labor supply decision, I also report in table 8 the age profile of the mean marginal rate of substitution between consumption and labor, and the marginal rate of substitution divided by the wage. Notice that the marginal rate of substitution is significantly higher than the real wage early in life. At age 20 it is 2.1 times greater than the real wage. Then, the marginal rate of substitution becomes closer to the actual wage rate at later stages of the career. The bias in the MaCurdy and Altonji estimation method arises from the fact that they do not recover the marginal rate of substitution, or the "effective wage", which is higher than the observed wage when there is human capital accumulation.

The reason the standard IV results are biased can best be described as follows. Deøne $\tilde{W}_t$ as the marginal rate of substitution between labor supply and consumption in period t. Also, deøne

$$\xi_t = ln(\tilde{W}_t) - ln(W_t).$$

Then, from the deønition of the marginal rate of substitution,

$$ln(\tilde{W}_t) = (a_2 - 1)ln(h_t) - (a_1 - 1)ln(C_t).$$

If I ørst diœerence away the consumption term, then I get the following expression:

$$\begin{aligned} \Delta ln(h_t) &= Const + \frac{1}{a_2 - 1}\Delta ln(\tilde{W}_t) + v_t \\ &= Const + \frac{1}{a_2 - 1}\Delta ln(W_t) + \frac{1}{a_2 - 1}\Delta \xi_t + v_t \end{aligned}$$

where $\frac{1}{a_2-1}\Delta\xi_t + v$ is the error term. Conventional instruments are correlated with $\Delta\xi_t$. For example, table 8 shows how $\Delta\xi_t$ is negatively correlated with age because the human capital eœect decreases with age. In this case, the elasticity estimates from IV estimation using age as one of the instruments will likely be negatively biased.

Furthermore, I also report the results based on Altoji (1985) type IV estimation on the simulated data, where consumption is used as a proxy for marginal utility of wealth. That is, I estimate the equation

$$ln(h_t) = \frac{1}{a_2 - 1}ln(W_t) - \frac{a_1 - 1}{a_2 - 1}ln(C_t).$$

Since I did not estimate measurement error of consumption, I used the simulated consumption value without measurement error as the regressor, together with simulated wage with measurement error. The intertemporal elasticity is again estimated to be much lower than the true value, conørming our claim that omission of human capital eœects biases the elasticity estimates downwards.

Next I consider the implications of my model for the elasticity of hours with respect to wages. In Fig 6., I plot the change in age t hours due to a 2 percent temporary increase in the age t wage payment, holding human capital øxed. This experiment can be interpreted as a temporary increase in rental rate for human capital by 2 percent. For a person at age 20 the 2 percent increase only increases hours by around 0.6 percent. But the elasticity of hours with respect to the current wage becomes larger with age. At around age 60, hours increase by more than 4 percent. That is, with age, hours become more responsive to wage changes. This supports the claim that when young, because of the high human capital investment returns, labor supply is insensitive to the wage change. But when old, since the human capital eœect is relatively insigniøcant, hours respond more to wage changes. Shaw (1989) noted that the human capital model implies this pattern.

7 Summary and Conclusions

In this paper, I estimate the intertemporal elasticity of substitution of labor supply and other parameters in a framework where people explicitly take into account human capital accumulation when they make labor supply decisions. I explicitly solve the continuous variable dynamic programming problem for optimal consumption and labor supply decisions and use the derived Emax function in a maximum likelihood routine. Using the estimated parameters, I conduct simulation experiments to generate age-wage, and age-labor hours proøles. I also use the simulated data to estimate the intertemporal elasticity of substitution parameter using the conventional OLS and the IV methods.

The results indicate that the maximum likelihood method based on the full solution of the continuous stochastic dynamic programming problem gives an estimated elasticity of intertemporal substitution parameter of 3.718, which is comparable to the elasticity results discussed in the macro literature. In contrast, in the micro literature, MaCurdy and Altonji have obtained IV estimates using the PSID that range from roughly 0.37 to 0.88. Using the NLSY, and applying the same IV procedure as MaCurdy and Altonji, I obtain elasticity estimates of 0.116 and 0.198. The results presented here suggest that the slightly higher elasticity estimates for the NLSY stem from its greater representation of young men. Clearly, however, the main reason for my much higher estimate of the intertemporal elasticity of substitution when I use the full solution procedure is my explicit inclusion of human capital accumulation in the model.

The simulated age proøle of wages and the marginal rate of substitution between labor supply and consumption implies that in the early stage of the agents' careers, the eœective wage, which I deøne as the marginal rate of substitution, is as much as 2.1 times higher than the real wage, implying that at younger ages, even if observed wages are low, the high eœective wage resulting from high returns to human capital accumulation induces agents to have high labor supply. However, as agents acquire experience, the ratio of the marginal rate of substitution to the wage falls. Through this mechanism, the labor supply model with human capital accumulation is able to reconcile a high elasticity of substitution with the fact that wages have a pronounced hump shape over the life cycle while the hump in hours is much more modest.

Lastly, a word of caution is in order when one interprets the above results. Simulation results show that the elasticity of intertemporal substitution being estimated around 4 does not imply that the model predicts individuals to change labor supply 4 times the wage change. On the contrary, the simulated hours response to the temporary wage increase is far less dramatic, especially amoung the young individuals. That is because when young, human capital accumulation is an important factor in determining labor supply, which temporary wage change has little eœect on.

7.1 References

Altug Sumru and Miller Robert A. (1998) ıThe Eœect of Work Experience on Female Wages and Labor Supply.ȷ Review of Economic Studies, 65, 45-85

Altonji Joseph G. (1982) ıThe Intertemporal Substitution Model of Labor Market Fluctuations: An Empirical Analysis.ȷ Review of Economic Studies 783-824.

Altonji Joseph G. (1986] ȷIntertemporal Substitution in Labor supply : Evidence from Micro Data.ȷ Journal of Political Economy 94 (3) , Part 2, T176-T215.

Becker, Gary T (1967) Human Capital. New York: Columbia University Press.

Blinder, Alan and Yoram Weiss (1976) ıHuman Capital and Labor Supply : A Synthesis.ȷ Journal of Political Economy 84: 449-72.

Ben-Porath, Yoram (1967) ıThe Production of Human Capital and Labor Supply: A Synthesis.ȷ Journal of Political Economy 75: 352-65.

Browning, M. A. Deaton and M. Irish (1985) ıA Proøtable Approach of Labor Supply and Commodity Demands over the Life-Cycle.ȷ Econometrica, 53, 503-544.

Browning, M. A. and C Meghir (1991) ıTesting for Separability of Commodity Demands from Male and Female Labor Supply.ȷ Econometrica ,59, 925-952.

Carrol Christopher D. (1997) ȷBuœer-Stock Saving and the Life Cycle/Permanent Income Hypothesis.ȷ Quarterly Journal of Economics, 112, 1-55

Eckstein, Zvi and Kenneth I. Wolpin (1989) ıDynamic Labor Force Participation of Married Women and Endogenous Wage Growth.ȷ Review of Economic Studies, 56, 375-90.

Eckstein, Zvi and Kenneth I. Wolpin (1989) ıThe Specioøcation and Estimation of Dynamic Stochastic Discrete Choice Models.ȷ Journal of Human Resources, 24, 562-98.

Eichenbaum, M.S., Hansen, L.P., and Singleton, K.S. (1988) ȷA Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty,ȷ Quarterly Journal of Economics, 103, 51-78

Hansen, L.P. (1982) ıLarge Sample Properties of Generalized Method of Moments Estimators.ȷ Econometrica, 50, 1029-1054

Hansen, L.P. and K.J. Singleton (1982) ıGeneralized Instrumental Variables Estimation of Non Linear Rational Expectations Models.ȷ Econometrica , 50, 1269-1286.

Heckman, James J. (1976) ȷA Life Cycle Model of Earnings, Learning and Consumption.ȷ Journal of Political Economy 84 (Supplement), 511-44.

Heckman, James, J. and Guillaume Sedlacek (1985) ıHeterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market.ȷ Journal of Political Economy 93, 1077-1125.

Hotz, V. Joseph, Finn Kydland and G.Sedlacek (1988) ıIntertemporal Preferences and Labor Supply.ȷ Econometrica, 56 (2), 335-60.

Keane Michael P. (1994) ıA Computationally Practical Simulation Estimator for Panel Data.ȷ Econometrica 62:1, 95-116

Keane Michael P and Kenneth, J. Wolpin (1994) ıThe Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence.ȷ Review of Economics and Statistics,76:4, 648-72

Keane Michael P. and Kenneth, J. Wolpin (1997) ıThe Career Decisions of Young Men.ȷ Journal of Political Economy 105:3, 473-521

Kydland Finn, and Edward C. Prescott. (1982) ıTime to Build and Aggregate Fluctuations.ȷ Econometrica, 1345-70

Lucas Robert E., and Leonard A. Rapping (1969) ıReal Wages, Employment, and InÆation.ȷ in Phelps, E. et al., Microeconomic Foundations of Employment and InÆation Theory (W. W. Norton and Company).

MaCurdy Thomas E . (1981) ıAn Empirical Model of Labor Supply in a Life Cycle Setting.ȷ Journal of Political Economy. 88 (9): 1059-85

MaCurdy Thomas E ıA Simple Scheme for Estimating an Intertemporal Model of Labor Supply and Consumption in the Presence of Taxes and Uncertainty.ȷ International Economic Review. 24(2), 265-89.

Mincer, Jacob (1958) ıInvestment in Human Capital and the Personal Income Distribution.ȷ Journal of Political Economy, 66,281-302

Prescott, Edward, C. (1986) ȷTheory Ahead of Business Cycle Measurement.ȷ Federal Reserve Bank of Minneapolis Quarterly Review.

Rogerson, Richard and Rupert Peter (1991) ıNew Estimates of Intertemporal Substitution: The Eœect of Corner Solutions for Year-Round Workers.ȷ Journal of Monetary Economics, 27 (2), 255-69.

Roy, A.D. (1951) ıSome Thoughts on the Distribution of Earnings.ȷ Oxford Economic Papers 3, 135-46.

Rust, John (1992) ıStructural Estimation of Markov Decision Process.ȷ Handbook of Labor Economics Vol. 4 Amsterdam: North Holland.

Shaw, Kathryn L. (1989) ıLife-Cycle Labor Supply with Human Capital Accumulation.ȷ International Economic Review 30, 431-456

Willis, Robert J. and Sherwin Rosen (1979) ıEducation and Self-Selection.ȷ Journal of Political Economy 87 (Supplement), T7-T36.

Table 1: Mean Age Proøle

(Sample sizes are in parentheses)

age	hourly wage	hours	total wealth
20	5.133 (445)	1763.1 (460)	3285.7 (66)
21	5.985 (574)	1924.4 (583)	3166.9 (158)
22	6.038 (631)	1956.1 (644)	3630.4 (252)
23	6.987 (721)	1991.1 (736)	2536.7 (395)
24	7.144 (810)	2072.6 (828)	3093.2 (546)
25	7.511 (836)	2121.1 (863)	-627.5 (702)
26	8.359 (889)	2166.9 (913)	1051.0 (805)
27	8.161 (914)	2217.7 (938)	-1009.5 (836)
28	8.667 (941)	2250.9 (964)	-1283.9 (885)
29	11.712 (961)	2270.0 (986)	4328.3 (883)
30	9.414 (850)	2277.8 (882)	-6124.3 (778)
31	12.418 (761)	2290.1 (790)	3820.8 (684)
32	14.492 (602)	2319.9 (632)	4479.8 (541)
33	10.373 (474)	2284.8 (501)	2382.6 (381)
34	28.861 (331)	2197.0 (355)	13434.7 (292)
35	15.761 (216)	2261.1 (225)	-3494.4 (242)

Table 2: Quantile Age Proøle

	wage quantiles			hours quantiles		
age	25 %	50 %	75 %	25 %	50 %	75 %
20	3.147	4.512	6.198	1113	1870	2086
21	3.687	5.090	7.050	1320	2000	2220
22	4.001	5.472	7.522	1440	2080	2240
23	4.289	5.932	8.206	1565	2080	2280
24	4.621	6.275	8.811	1710	2080	2340
25	4.940	6.771	8.964	1820	2080	2375
26	5.290	7.413	10.16	1920	2080	2396
27	5.429	7.677	10.16	2002	2080	2455
28	5.779	7.943	10.72	2080	2104	2510
29	5.992	8.436	11.23	2072	2116	2562
30	6.137	8.789	11.80	2080	2125	2544
31	6.177	8.817	12.83	2080	2135	2580
32	6.260	8.983	12.36	2080	2184	2600
33	6.025	9.331	12.51	2047	2140	2524
34	6.593	9.316	13.63	1980	2080	2455
35	6.758	9.437	13.64	2080	2080	2445

Table 3: Mean Age Proøles After Censoring

age	mean wage	sample size	mean hours	sample size	mean assets	sample size
20	4.823	423	1691.6	441	2313.7	64
21	5.480	542	1860.0	561	2354.0	148
22	5.840	596	1897.9	620	2903.3	229
23	6.286	682	1929.9	710	3577.0	353
24	6.839	772	2018.3	804	2959.4	468
25	7.157	796	2075.1	841	2582.6	581
26	7.830	844	2109.7	886	2378.8	626
27	7.912	870	2163.2	911	3339.8	636
28	8.345	894	2201.5	937	2042.5	677
29	8.818	916	2218.8	958	815.27	640
30	9.166	808	2226.8	857	-514.6	570
31	9.268	722	2242.6	768	-1411.8	513
32	9.556	575	2274.1	616	-2307.6	403
33	9.664	450	2232.5	487	-2896.8	277
34	10.55	314	2146.6	345	-493.6	219
35	10.68	205	2213.0	218	-1598.7	175

Table 4: Estimation Results

(Standard Errors are in parentheses.)

Utility Function Parameters		
a_1	marginal utility of consumption	0.1990 (5.51E-4)
a_2	marginal disutility of labor	1.2691 (1.06E-3)
b_{0n}	disutility of labor, nonhighschool	1.689D-5 (1.59E-7)
b_{0h}	disutility of labor, highschool grad.	1.819D-5 (5.22E-8)
b_{1n}	disutility of labor, nonhighschool	-0.0978 (1.51E-3)
b_{1h}	disutility of labor, highschool grad.	-0.0712 (4.35E-4)
b_{2n}	disutility of labor, nonhighschool	1.8300E-3 (2.07E-3)
b_{2h}	disutility of labor, highschool grad.	0.0245 (1.98E-4)
σ_1	std.error of disutility shock	0.0416 (4.58E-3)
C_0	consumption utility weight	0.0354 (1.69E-4)
C_{1n}	consumption utility weight, nonhighschool	0.6762 (3.79E-3)
C_{1h}	consumption utility weight, highschool grad.	0.4983 (2.61E-3)
C_{2n}	consumption utility weight, nonhighschool	0.2265 (3.73E-3)
C_{2h}	consumption utility weight, highschool grad.	0.2316 (0.0163)
β	discount factor	0.9480 (2.92D-4)
$\overline{A}$	initial assets when the starting age is before 21	2509.1 (122.3)
$\overline{A}$	initial assets when the starting age is after 21, nonhighschool	3718.6 (598.7)
$\overline{A}$	initial assets when the starting age is after 21, nonhighschool	4011.4 (95.6)
$V_{\overline{A}}$	std. error, initial assets	1435.5.0 (26.9)

Production Function Parameters[3]		
δ_n	linear human capital term, nonhighschool	0.4053 (1.08E-3)
δ_h	linear human capital term, highschool grad.	0.3545 (4.99E-4)
A_{0n}	øxed coeŒcient in term $G(K,h,t)$, nonhighschool	0.1280 (3.32E-4)
A_{0h}	øxed coeŒcient in term $G(K,h,t)$, highschool	0.1373 (1.49E-4)
A_{1n}	age varying coeŒcient in term $G(K,h,t)$, nonhighschool	-1.694D-3 (1.02E-5)
A_{1h}	age varying coeŒcient in term $G(K,h,t)$, highschool	-2.156D-3 (1.15E-5)
α_n	exponent term on $(h+d_1)$, nonhighschool	0.2282 (1.84E-4)
α_h	exponent term on $(h+d_1)$, highschool	0.2307 (3.77E-3)
B_1	additive constant in capital term B_1+K	0.0732 (2.59E-3)
B_2	coeŒcient in negative term of hours. $-B_2(h+d_1)$	4.131D-4 (9.13E-7)
σ_0	std.error of wage shock.	0.051 (7.56E-4)
d_1	additive constant in hours term $h+d_1$	280.5 (15.2)
k_0	intercept term, nonhighschool	0.0361 (2.76E-3)
k_0	intercept term, highschool	0.0361 (2.43E-4)

Measurement Error Parameters		
σ_{v0}	std. error, initial period wage meas. error	0.4622 (3.30E-3)
σ_{v1}	std. error, wage meas. error	0.3388 (1.95E-3)
σ_{v2}	std. error, hours meas. error	560.7 (2.68)
σ_{v31}	std. error, asset meas. error	3782.8 (98.8)
σ_{v32}	std. error, asset meas. error	687.4 (8.37)

[3]

$$g(K,h,t) = (A_0 + A_1(t-20))(B_0+K)[(h+d_1)^{\alpha} - B_1(h+d_1)] + \delta K + k_0. \tag{4}$$

Table 5: Simulated Mean Age Proøles

age	mean wage	mean hours	mean assets	$[(1+r)\beta]^{t-20}$ $U_C(t)$
20	6.014	1924.5	2787.9	1.29D-5
21	6.193	1946.2	2887.3	1.31D-5
22	6.485	1990.8	2705.6	1.30D-5
23	6.836	2041.6	2256.4	1.29D-5
24	7.183	2097.1	2312.6	1.28D-5
25	7.474	2122.8	1393.0	1.27D-5
26	7.779	2149.7	217.9	1.27D-5
27	8.071	2176.1	-831.9	1.27D-5
28	8.371	2202.9	-2067.6	1.27D-5
29	8.672	2227.9	-3205.3	1.27D-5
30	8.967	2250.6	-4368.9	1.27D-5
31	9.251	2274.4	-5575.2	1.27D-5
32	9.529	2295.9	-6894.7	1.27D-5
33	9.816	2317.0	-8224.7	1.27D-5
34	10.09	2335.2	-8690.9	1.27D-5
35	10.35	2350.6	-8235.0	1.27D-5
36	10.61	2365.7	-6909.3	1.27D-5
37	10.86	2378.2	-4695.3	1.27D-5
38	11.11	2387.8	-1587.7	1.27D-5
39	11.33	2393.5	2400.9	1.27D-5
40	11.54	2396.7	7243.0	1.27D-5
41	11.74	2397.8	12914.2	1.27D-5
42	11.92	2394.0	19389.9	1.27D-5

Table 5 (Continued): Simulated Mean Age Proøles

age	mean wage	mean hours	mean assets	$[(1+r)\beta]^{t-20}$ $U_C(t)$
43	12.08	2383.0	26622.8	1.27D-5
44	12.23	2368.9	34508.0	1.27D-5
45	12.35	2346.9	42982.3	1.27D-5
46	12.44	2318.7	51939.9	1.27D-5
47	12.52	2281.6	61270.4	1.27D-5
48	12.58	2239.5	70829.1	1.27D-5
49	12.60	2184.0	80523.4	1.27D-5
50	12.60	2116.6	90146.0	1.27D-5
51	12.56	2036.9	99507.3	1.27D-5
52	12.48	1939.7	108408	1.27D-5
53	12.34	1828.8	116516	1.27D-5
54	12.16	1705.2	123630	1.27D-5
55	11.93	1570.6	129613	1.27D-5
56	11.62	1427.0	134357	1.28D-5
57	11.25	1272.9	137776	1.28D-5
58	10.81	1110.5	139757	1.29D-5
59	10.28	940.18	140227	1.30D-5
60	9.670	775.38	139094	1.31D-5
61	8.989	617.48	136371	1.31D-5
62	8.250	479.71	132088	1.32D-5
63	7.458	355.69	125758	1.33D-5
64	6.648	296.54	118095	1.39D-5
65	5.889	181.70	110232	1.31D-5

Table 6: ML, OLS, IV Results

(Std. errors are in parentheses)

Instruments:

const, experience (which is age-16), experience squared, twice lagged wage

	b_2	a_2
ML	3.718 (1.46D-2)[1]	1.269 (1.06D-3)
simulated data from age 20 to 65.		
OLS	-0.462 (0.004)	-1.165 (0.019)[1]
3SLS	2.445 (0.278)	1.409 (0.047)[1]
simulated data from age 20 to 55.		
OLS	-0.401 (0.004)	-1.494 (0.025)[1]
3SLS	1.614 (0.638)	1.620 (0.245)[1]
simulated data from age 20 to 45.		
OLS	-0.408 (0.005)	-1.451 (0.030)[1]
3SLS	0.640 (0.870)	2.563 (2.124)[1]
simulated data from age 20 to 35.		
OLS	-0.429 (0.006)	-1.331 (0.033)[1]
3SLS	0.170 (1.073)	6.882 (37.2)[1]
NLSY data. (From age 20 to 35)		
OLS	-0.175 (0.098)	-4.714 (3.200)[1]
3SLS	0.116 (0.087)	9.621 (6.466)[1]

[1]Delta method is used to calculate the standard errors.

Table 7: OLS, IV Results with cleaned data.

(Std. errors are in parentheses)

Instruments:

const, experience (which is age-16), experience squared, twice lagged wage.

	b_2	a_2
ML	3.718 (1.46D-2)[1]	1.269 (1.06D-3)
simulated data from age 20 to 65.		
OLS	-0.244 (0.006)	-3.098 (0.101)[1]
3SLS	2.385 (0.461)	1.419 (0.081)[1]
simulated data from age 20 to 55.		
OLS	-0.269 (0.007)	-2.718 (0.097)[1]
3SLS	1.880 (0.996)	1.532 (0.282)[1]
simulated data from age 20 to 45.		
OLS	-0.278 (0.008)	-2.597 (0.104)[1]
3SLS	0.785 (1.955)	2.274 (3.173)[1]
simulated data from age 20 to 35.		
OLS	-0.280 (0.011)	-2.571 (0.140)[1]
3SLS	-0.057 (0.528)	-16.544 (162.5)[1]
NLSY data. (From age 20 to 35)		
OLS	-0.196 (0.090)	-4.102 (2.343)[1]
3SLS	0.198 (0.103)	6.051 (2.627)[1]

[1] Delta method is used to calculate the standard errors.

Table 6: ML, OLS, IV Results of Altonji Estimation

(Std. errors are in parentheses)

Instruments:

const, experience (which is age-16), experience squared, experience cubed

	b_2	a_2
ML	3.718 (1.46D-2)[1]	1.269 (1.06D-3)
simulated data from age 20 to 65.		
OLS	-0.025 (0.012)	-39.0 (19.20)[1]
3SLS	1.908 (0.062)	1.524 (0.017)[1]
simulated data from age 20 to 55.		
OLS	-0.377 (0.005)	-1.653 (0.035)[1]
3SLS	-0.315 (0.021)	-2.175 (0.212)[1]
simulated data from age 20 to 45.		
OLS	-0.355 (0.005)	-1.817 (0.040)[1]
3SLS	0.168 (0.034)	6.952 (1.205)[1]
simulated data from age 20 to 35.		
OLS	-0.414 (0.007)	-1.416 (0.041)[1]
3SLS	0.248 (1.073)	5.032 (17.4)[1]

[1]Delta method is used to calculate the standard errors.

Table 8: Simulated MRS Over the Life Cycle

age	MRS	MRS/Wage
20	12.91	2.146
21	12.98	2.096
22	13.11	2.021
23	13.32	1.949
24	13.44	1.871
25	13.65	1.826
26	13.86	1.782
27	14.02	1.737
28	14.19	1.695
29	14.36	1.656
30	14.52	1.620
31	14.64	1.583
32	14.76	1.549
33	14.89	1.517
34	15.02	1.488
35	15.12	1.461
36	15.22	1.435
37	15.31	1.410
38	15.41	1.388
39	15.49	1.368
40	15.56	1.349
41	15.62	1.331
42	15.67	1.315

age	MRS	MRS/Wage
43	15.71	1.301
44	15.75	1.288
45	15.77	1.277
46	15.76	1.267
47	15.76	1.258
48	15.71	1.249
49	15.64	1.241
50	15.54	1.234
51	15.41	1.227
52	15.25	1.221
53	15.01	1.216
54	14.71	1.209
55	14.31	1.200
56	13.82	1.189
57	13.24	1.177
58	12.56	1.162
59	11.77	1.144
60	10.88	1.125
61	9.932	1.105
62	8.962	1.086
63	7.914	1.061
64	7.123	1.071
65	5.915	1.004

Appendix: Backward Solution of the Bellman Equations. Interpolation and Integration Steps.

To avoid quadrature integration with respect to the human capital shocks, I add another interpolation and approximation step. The basic logic which I apply is that the value function $V_{t,s}(A_t, \tilde{K}_t, \epsilon_{1,t}, \epsilon_{2,t})$ is only a function of assets, A_t, human capital, K_t, and the taste shock $\epsilon_{2,t}$. Once K_t is known, in order to calculate the value function, I do not need to know the values of the wage shock. So I go back to the original deønition of the value function, in terms of A_t, $K_t = \tilde{K}_t \epsilon_{1,t}$ and $\epsilon_{2,t}$. The solution steps are as follows:

Step 1 Integrating the value function with respect to the taste shock. Assume that the age $t+1$ Emax function

$$V^E_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1}) = E_t[V_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1}, \epsilon_{1,t+1}, \epsilon_{2,t+1})]$$

is already calculated. I use Gaussian-Hermite quadrature to integrate the expected value function over the taste shocks at asset and human capital grid points (A_i, K_j).

First, I calculate the quadrature points and weights for the taste shock $\epsilon_{2,t}$. Since $\epsilon_{2,t}$ has a log normal distribution with parameters $\mu_2 = -\frac{1}{2}\sigma_2^2$ and σ_2, $log(\epsilon_{2,t+1})$ is normally distributed with mean μ_2 and standard error σ_2. Let

$$x_{h,l}, \; l = 1, ..., n_2$$

be the points for Gauss-Hermite quadrature. Then,

$$\epsilon^q_{2,l} = exp(\sqrt{2}\sigma_2 x_{h,l} + \mu_2)$$

$$l = 1, ..., n_2$$

are the Gauss Hermite quadrature points for the above log normal distribution. Given $(A_i, K_j, \epsilon^q_{2,l})$, and the next period emax function, calculate the value function for each quadrature point of the taste shock $\epsilon^q_{2,l}$ as follows:

$$V_{t,s}(A_i, K_j, \epsilon^q_{2,l})$$

$$= max_{\{C_t, h_t\}}\{U(C_t, t) - V(h_t, \epsilon_{2,l}, t) + \beta V^E_{t+1,s+1}(A_{t+1}, \tilde{K}_{t+1})\}$$

subject to the intertemporal budget constraint and human capital production function. Notice that this step requires a two dimensional Newton search for optimal (C_t, h_t) at only $n_A \times n_K \times n_q$ grid points, so the factor of n_q arising from the human capital shock is eliminated. Now, we can approximate the integration as follows using the quadrature procedure, with w^q being the weights for Gauss-Hermite quadrature..

$$\begin{aligned} E_{\epsilon_2} V_{t+1,s+1}(A_i, K_j.\epsilon_2) &= \int_{-\infty}^{\infty} V_{t+1,s+1}(A_i, K_j, exp(\theta)) \\ &\quad \frac{1}{\sqrt{2\pi\sigma_1}} exp[-\frac{1}{2\sigma_1^2}(\theta-\mu_1)^2]d\theta \\ &= \int_{-\infty}^{\infty} V_{t+1,s+1}(A_i, K_j, exp(\sqrt{2}\sigma_1 x + \mu_1)) \frac{1}{\sqrt{\pi}} exp[-x^2]dx \\ &\approx \frac{\sum_{i=1}^{n_q} V_{t+1,s+1}(A_i, K_j, \epsilon_{2,l}^q) w_l^q}{\sqrt{\pi}} \end{aligned}$$

Step 2 Integrating the value function with respect to wage shock.

Next, I integrate the value function with respect to the wage shocks to derive the Emax function at the grid points $(A_i, \tilde{K}_j)$. From step 1, for each grid point (A_i, K_j), I already have the value function integrated with respect to the taste shocks. Now, for each given asset grid value A_i, I øt Chebychev polynomials of log human capital to n_2 values of the integrated value function. That is, I derive

$$\hat{E}_{\epsilon_2} V_{t,s}(A_i, K, \epsilon_{2,t}) = \sum_{l=0}^{n_c} \pi_l T_l(log(K))$$

where $T_l(log(K))$ is the l th order Chebychev polynomial of log human capital. The coeŒcients π are derived by least squares with dependent variables being

$$E_{\epsilon_2} V_{t,s}(A_i, K_j, \epsilon_{2,t}),\ j = 1, ..., n_2$$

Transform the Chebychev polynomials to the polynomials of $log(K)$. Then,

$$\hat{E}_{\epsilon_2} V_{t,s}(A_i, K, \epsilon_{2,t}) = \sum_{i=0}^{n_c} \pi_i' log(K)^i.$$

Notice that for any realized human capital shock $\epsilon_{1,t}$, the value function integrated over the taste shock is the sum of polynomials of the log wage shock and $log(\tilde{K})$.

$$\hat{E}_{\epsilon_2} V_{t,s}(A_i, \tilde{K}\epsilon_{1,t}, \epsilon_{2,t}) = \sum_{i=1}^{n_c} \pi_i' [log(\tilde{K}) + log(\epsilon_{1,t})]^i.$$

$$= \sum_{i=1}^{n_c} \pi'_i \sum_{j=0}^{i} \binom{n_c}{j} [log(\epsilon_{1,t})]^j [log(\tilde{K})]^{i-j}$$

Hence, we can integrate the value function with respect to the wage shock by integrating each wage shock polynomial separately as the following equation shows.

$$\begin{aligned} E\{V_{t,s}(A_i, \tilde{K}\epsilon_{1,t}, \epsilon_{2,t})\} &= E_{\epsilon_1} \hat{E}_{\epsilon_2} V_{t,s}(A_i, \tilde{K}\epsilon_{1,t}, \epsilon_{2,t}) \\ &= \sum_{i=1}^{n_c} \pi'_i \sum_{j=0}^{i} \binom{n_c}{j} \{E_{\epsilon_1}[log(\epsilon_{1,t})]^j\} [log(\tilde{K})]^{i-j} \end{aligned}$$

Here, since $log(\epsilon_{1,t})$ is normally distributed, integration of $[log(\epsilon_{1,t})]^j$ can be done analytically.

Then, I again approximate the above equation over $A_t, \tilde{K}_t$ using Chebychev polynomials to derive the Emax function at age t, which I use for solving the age $t-1$ Bellman equation.

Notice that using the above algorithm, the problems of applying two dimensional quadrature are avoided. First, the Newton nonlinear search over consumption and labor needs only be applied to $n_A \times n_K \times n_q \times (T-20)$ points, hence reducing the computational time by a factor of n_q. Also, since only one dimensional quadrature integration is involved, there is no accuracy loss due to high dimensional quadrature integration. Thirdly and most importantly, I only need to calculate the value function at the Chebychev grid point values for human capital and not at grid points for the human capital shock. Hence, I do not need to calculate value functions at extremely low or high human capital levels.

Appendix: Data Generation

I derived the wages, hours and asset data from the NLSY as follows:

hours data : I use the variable “Number of hours worked in past calendar year” from 1979 to 1993.

wage data : I first get the total wage income data from the variable “Total Income from wages and salary income past calendar year” from 1979 to 1993. And after adjusting for inflation using the GDP deflator, divide the variables with the hours data to get the hourly wage rate.

asset data : I added up the following variables in the NLSY to construct total positive assets: “Total market value of vehicles including automobiles r/spouse own”, “Total market value of farm/business/other property r/spouse own”, “Market value of residential property r/spouse own”, “Total market value of stocks/bonds/mutual funds”, “Total amount of money assets like savings accounts of r/spouse”, “Total market value of all other assets each worth more than $500”

I then added up the following variables to construct total negative assets: “Total amount of money r/spouse owe on vehicles including automobiles”, “Total amount of debts on farm/business/other property r/spouse owe”, “Amount of mortgages & back taxes r/spouse owe on residential property”, “Total amount of other debts over $500 r/spouse owe”

The total amount of assets is calculated by subtracting the total amount of negative assets from the total amount of positive assets.

Fig 1: Optimal life cycle labor supply

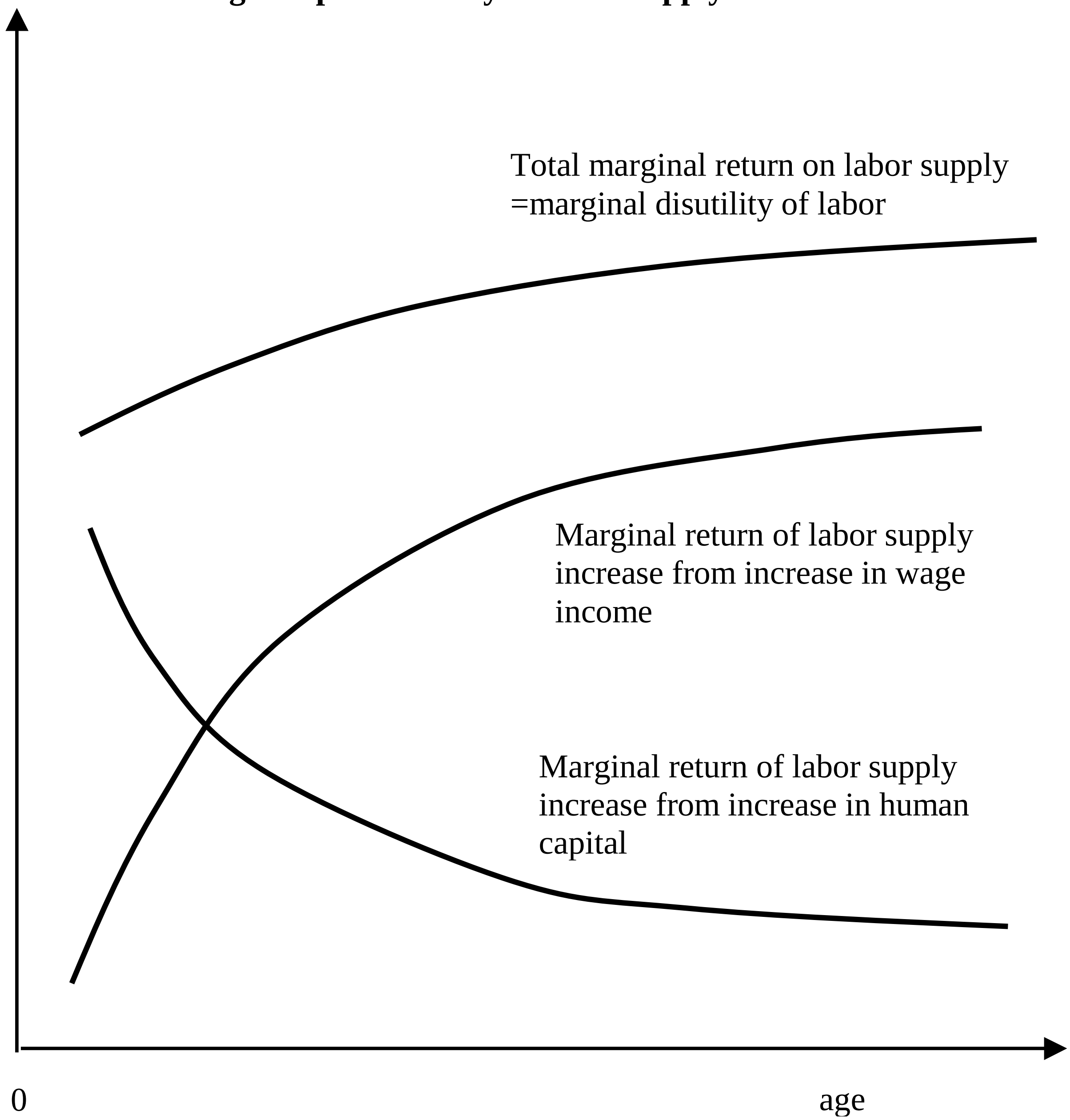

Fig 2: Human Capital Production Function

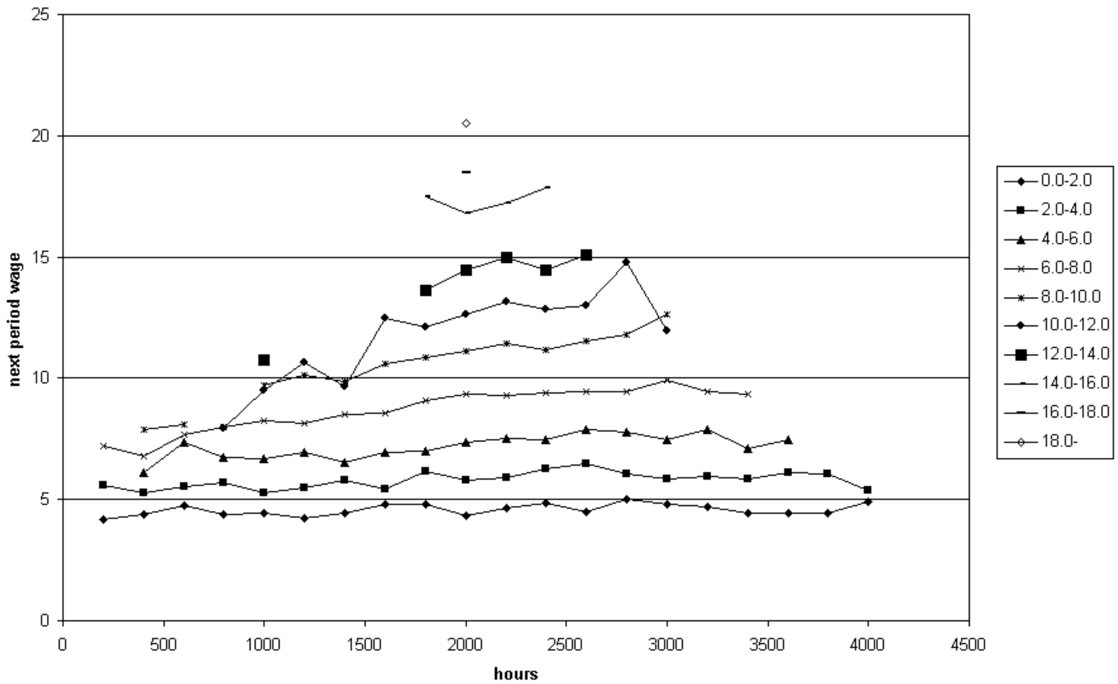

Fig 3: Age Wage Profiles

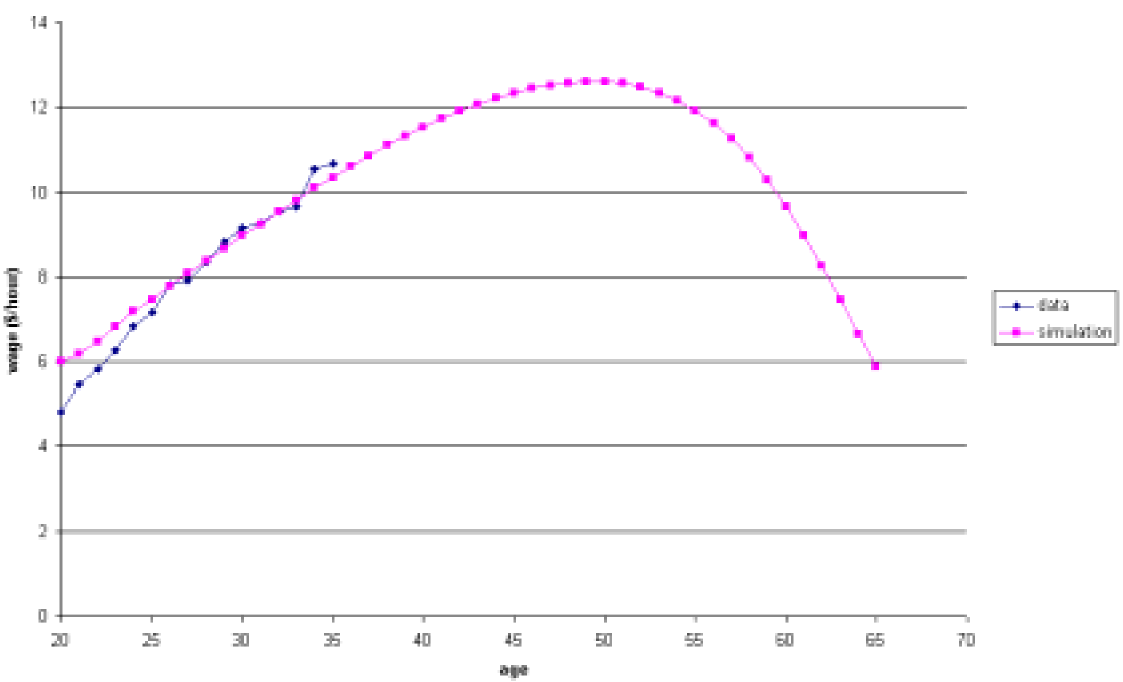

Fig 4: Age Hours Profiles

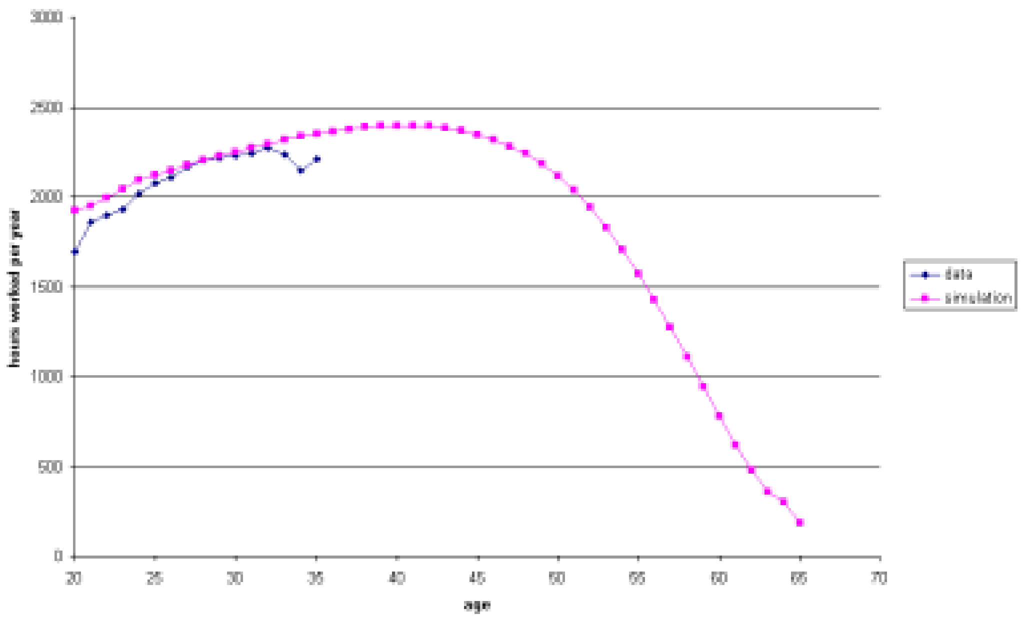

Fig 5: Age Asset Profiles

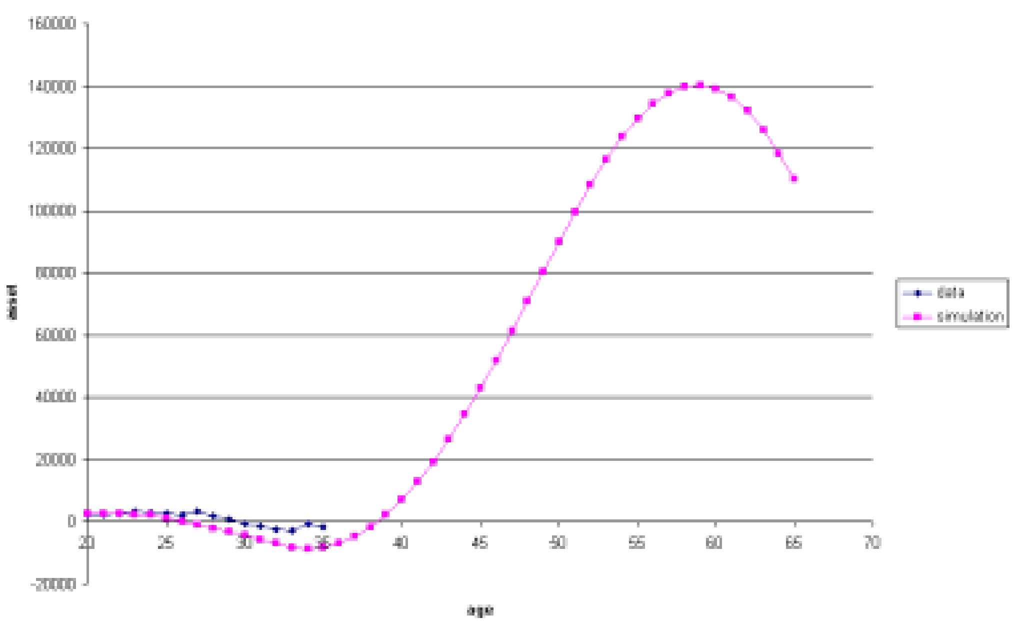

Fig 6: Simulated MRS/Wage profile

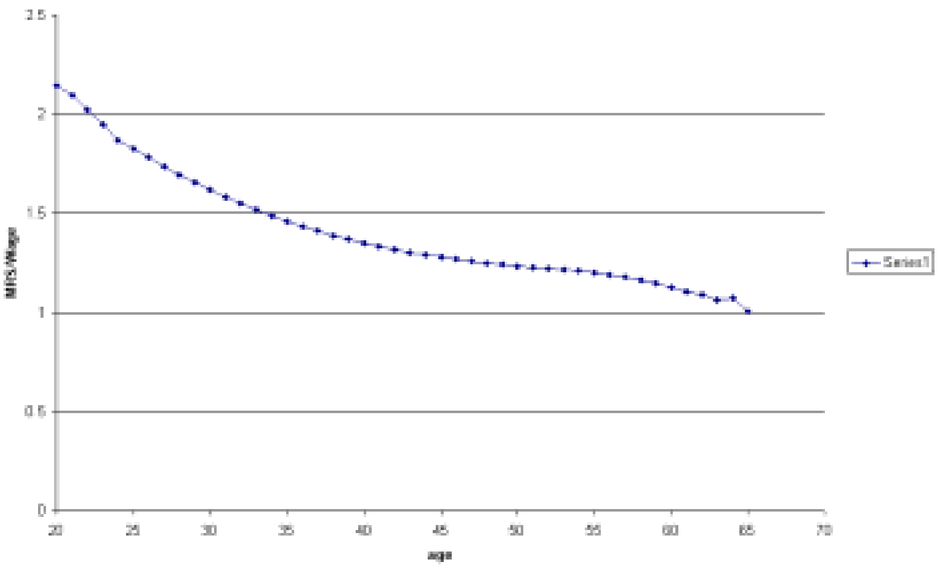

Fig 7: Age Hours Change Profile