The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour

Author(s): Eric French

Source: *The Review of Economic Studies*, Apr., 2005, Vol. 72, No. 2 (Apr., 2005), pp. 395-427

Published by: Oxford University Press

Stable URL: https://www.jstor.org/stable/3700657

Review of Economic Studies (2005) **72**, 395–427 0034-6527/05/00170395$02.00

The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour

ERIC FRENCH
Federal Reserve Bank of Chicago

First version received January 2002; *final version accepted July* 2004 (*Eds.*)

This paper estimates a life cycle model of labour supply, retirement, and savings behaviour in which future health status and wages are uncertain. Individuals face a fixed cost of work and cannot borrow against future labour, pension, or Social Security income. The method of simulated moments is used to match the life cycle profiles of labour force participation, hours worked, and assets that are estimated from the data to those that are generated by the model. The model establishes that the tax structure of the Social Security system and pensions are key determinants of the high observed job exit rates at ages 62 and 65. Removing the tax wedge embedded in the Social Security earnings test for individuals aged 65 and older would delay job exit by almost one year. By contrast, Social Security benefit levels, health, and borrowing constraints are less important determinants of job exit at older ages. For example, reducing Social Security benefits by 20% would cause workers to delay exit from the labour force by only three months.

1. INTRODUCTION

Why do individuals retire when they do? This paper provides an empirical analysis of the effects of the Social Security system and liquidity constraints on life cycle labour supply. It presents the first structural model of labour supply and retirement behaviour where individuals can save to insure themselves against health and wage shocks as well as for retirement, but cannot borrow against future labour, Social Security, and pension income to smooth consumption in the face of an adverse shock. Previous structural analyses of labour supply and retirement behaviour have made diametrically opposed assumptions about a household's ability to borrow and save. At one extreme, Burtless (1986) and Gustman and Steinmeier (1986) assume that households can perfectly smooth consumption by borrowing and lending without limit. At the opposite extreme, Stock and Wise (1990) and Rust and Phelan (1997) assume that households cannot borrow or save, thus allowing no intertemporal consumption smoothing. Clearly, neither of these extreme assumptions is correct.

Understanding the importance of borrowing constraints is critical when considering the effects of the Social Security rules on lifetime labour supply. For example, suppose that Social Security were to become less generous for members of a particular cohort. This would reduce lifetime wealth and diminish the importance of the Social Security work disincentives for members of that cohort. The loss of wealth would cause individuals to work more hours in order to earn more income at some stage of their lifetime. However, it is not clear when they would do so. If individuals are liquidity constrained at the early Social Security retirement age (62), those younger than 62 will not react to the benefit reduction. All consumption and labour supply responses will be after age 62. On the other hand, if individuals are not liquidity constrained, then there may be significant responses by those younger than 62.

The model in this paper allows for a wide range of individual behaviour. The model also captures the fact that the structure of Social Security and pensions causes declining work incentives after age 62. Because individuals in the model can save and decumulate assets, they may leave the labour force and begin dissaving shortly after age 62, as Gustman and Steinmeier (1986) argue. However, they cannot borrow against future Social Security, pension, or labour income. Therefore, individuals may have to remain in the labour-market until they are eligible for Social Security and pension benefits, as Rust and Phelan (1997) argue.[1]

This paper uses the method of simulated moments to match life cycle profiles estimated using data from the Panel Study of Income Dynamics (PSID) to life cycle profiles generated by a dynamic programming model. I match labour force participation, hours worked, and asset profiles. Assuming that preferences are not affected by age (after conditioning on health status and family size), matching these profiles allows me to identify key structural parameters such as the intertemporal elasticity of substitution and the time discount factor. This allows me to consider whether the inability to borrow against future Social Security benefits significantly affects life cycle labour supply. Moreover, I consider the decisions of men aged 30–95, allowing me to consider when in a worker's lifetime we should expect to see labour supply responses to changes in the Social Security rules.

The PSID data are consistent with a low level of labour supply substitutability for young men but a high degree of labour supply substitutability for older men. Consistent with previous research, I find very little life cycle variation in hours worked for men between ages 30 and 55. I also find that work-hours and labour force participation decline sharply after age 55, and especially sharply at ages 62 and 65. These are exactly the ages at which Social Security, pensions, and declining wages provide strong incentives to leave the labour force.

The dynamic programming model produces reasonable preference parameter estimates. It also captures many features of the data, including the sharp decline in labour force participation rates between ages 55 and 70 and the especially large drops at ages 62 and 65. In order to fit both the participation and hours worked profiles, the model estimates a large fixed cost of work. The fixed cost generates a high level of labour supply substitutability at the labour force participation margin. Because of the Social Security and pension incentives to leave the labour force, those in their 60s are near the labour force participation margin. As a result, labour supply elasticities rise from 0·3 at age 40 to 1·1 at age 60.

I use the model to conduct three simulations. First, I consider shifting the early retirement age from 62 to 63. I find that this has almost no effect on labour supply because forward looking agents almost always have sufficient financial resources at age 62 to finance an additional year out of the labour force. Because forward looking agents have positive assets near retirement, liquidity constraints never bind at retirement age. Second, I consider a 20% reduction in Social Security benefits. I find that this would cause individuals to delay job exit from the labour-market by three months in order to develop sufficient financial assets to offset lost retirement income. Because older individuals are the ones most willing to substitute their labour supply, most of the labour supply response would be after age 62. Third, I consider eliminating the tax wedge caused by the Social Security earnings test. I find that this would cause individuals to work an additional one year. Together, these three simulations suggest that the Social Security earnings test is the most significant labour supply incentive of the Social Security system. Interactions between the Social Security system and liquidity constraints are relatively unimportant.

1. Several recent studies, such as Rust, Buchinsky and Benitez-Silva (2003), van der Klaauw and Wolpin (2003) and French and Jones (2004), also account for uncertainty, borrowing constraints, and a savings decision.

The rest of the paper is arranged as follows. Section 2 develops a model of optimal lifetime decision making. Section 3 describes the estimation scheme: the Method of Simulated Moments. Section 4 describes the data. Section 5 presents parameter estimates. Section 6 describes the policy experiments. Section 7 concludes. Technical details of the paper, as well as additional discussion, are in French (2003).

2. THE MODEL

2.1. *The set-up*

This section describes the model of lifetime decision making. Individuals choose consumption, work-hours (including the labour force participation decision), and whether or not to apply for Social Security benefits. They are allowed to save but not borrow. When making these decisions, they are faced with several forms of uncertainty: survival uncertainty, health uncertainty, and wage uncertainty.

Consider a household head seeking to maximize his expected lifetime utility at age (or equivalently, year) t, $t = 1, 2, \ldots, T+1$. Each period that he lives, the individual receives utility, U_t, from consumption, C_t, hours worked, H_t, and health (or medical) status, M_t, so that $U_t = U(C_t, H_t, M_t)$. When he dies, he values bequests of assets, A_t, according to a bequest function $b(A_t)$. Let s_t denote the probability of being alive at age t conditional on being alive at age $t-1$, and let $S(j,t) = (1/s_t)\prod_{k=t}^{j} s_k$ denote the probability of living to age j, conditional on being alive at age t. Since age $T+1$ is the terminal period, $s_{T+1} = 0$. We assume that preferences take the form

$$U(C_t, H_t, M_t) + E_t\left[\sum\nolimits_{j=t+1}^{T+1} \beta^j S(j-1,t)(s_j U(C_j, H_j, M_j) + (1-s_j)b(A_j))\right], \tag{1}$$

where β is the time discount factor. In addition to choosing hours and consumption, eligible individuals can choose whether to apply for Social Security benefits; let the indicator variable $B_t \in \{0,1\}$ equal one if the individual has applied for benefits. The individual maximizes equation (1) by choosing the contingency plans $\{C_j, H_j, B_j\}_{j=t}^{T+1}$, subject to the following equations, described below: a mortality determination equation (4), a health determination equation (5), wage determination equations (6) and (7), a spousal income determination equation (8), and an asset accumulation equation (9).

The within-period utility function is of the form

$$U(C_t, H_t, M_t) = \frac{1}{1-\nu}\left(C_t^{\gamma}(L - H_t - \theta_P P_t - \phi I\{M = \text{bad}\})^{1-\gamma}\right)^{1-\nu}, \tag{2}$$

where the per period time endowment is L and the quantity of leisure consumed is $L - H_t - \theta_P P_t - \phi I\{M = \text{bad}\}$. The 0-1 indicator $I\{M = \text{bad}\}$ is equal to 1 when health is bad and 0 when health is good. Participation in the labour force is denoted by P_t, a 0-1 indicator equal to 0 when hours worked, H_t, equals zero. The fixed cost of work, θ_P, is measured in hours worked per year.[2] Retirement is assumed to be a form of the participation decision. Workers can re-enter the labour force.

2. The distribution of annual hours of work is clustered around both 2000 and 0 hours of work, a regularity in the data that standard utility functions have a difficult time replicating. Fixed costs of work are a common way of explaining this regularity in the data (Cogan, 1981). Fixed costs of work generate a reservation wage for a given marginal utility of wealth. Below the reservation wage, hours worked is zero. Slightly above the reservation wage, hours worked may be large. Individual level labour supply is highly responsive around this reservation wage level although wage increases above the reservation wage result in a smaller labour supply response.

The bequest function is of the form

$$b(A_t) = \theta_B \frac{(A_t + K)^{(1-\nu)\gamma}}{1-\nu}, \tag{3}$$

where K determines the curvature of the bequest function. If $K = 0$ there is infinite disutility of leaving non-positive bequests. If $K > 0$, the utility of a zero bequest is finite.

Given the objective function, individuals face several constraints. Mortality rates depend upon age and previous health status:

$$s_{t+1} = s(M_t, t+1). \tag{4}$$

Next year's health status, $\text{prob}(M_{t+1} \mid M_t, t+1)$, depends on current health status and age. Health status follows a two-state transition matrix at each age with a typical element

$$\pi_{\text{good,bad},t+1} = \text{prob}(M_{t+1} = \text{good} \mid M_t = \text{bad}, t+1). \tag{5}$$

The logarithm of wages at time t, $\ln W_t$, is a function of hours worked, age, and health status, plus an autoregressive component of wages AR_t:

$$\ln W_t = \alpha \ln H_t + W(M_t, t) + AR_t. \tag{6}$$

The function $W(M_t, t)$ is described in detail in Section 3.2. The autoregressive component of wages has a correlation coefficient ρ and a normally distributed innovation η_t:

$$AR_t = \rho AR_{t-1} + \eta_t, \qquad \eta_t \sim N(0, \eta_\zeta^2). \tag{7}$$

By assumption, at time $t-1$ the worker knows the autoregressive component of wages (AR_{t-1}) but only knows the distribution of the innovation in next period's wage (η_t).

Spousal income, described in detail in Section 4.2, depends upon the individual's wage and age:

$$ys_t = ys(W_t, t). \tag{8}$$

The final constraint is the asset accumulation equation:

$$A_{t+1} = A_t + Y(rA_t + W_t H_t + ys_t + pb_t + \varepsilon_t, \tau) + (B_t \times ss_t) - C_t, \qquad A_{t+1} \geq 0, \tag{9}$$

where $Y(rA_t + W_t H_t + ys_t + pb_t + \varepsilon_t, \tau)$ is the level of post tax income, r is the interest rate, τ is the tax structure, pb_t denotes pension benefits (described in Section 2.2), ε_t denotes a pension accrual residual (described in Section 2.3), and ss_t denotes Social Security benefits net of the earnings test (described in Section 2.3). French (2003) describes the computation of taxes, Social Security, and pension benefits in greater detail.

Individuals cannot draw Social Security benefits until age 62. By assumption, the date of pension benefit receipt is 62. Because it is illegal to borrow against Social Security benefits and difficult to borrow against most forms of pension wealth, individuals with low asset levels potentially must wait until age 62 to finance exit from the labour-market.

2.2. *Social security*

There are three major labour supply incentives provided by the Social Security system.[3] All three incentives tend to induce exit from the labour-market by age 65.

3. I use tax and benefit formulae from the *Social Security Handbook Annual Statistical Supplement* for the year 1987 for several reasons. First, 1987 is relatively close to the middle year of the data. Second, there were significant changes to the tax code enacted in 1986 that simplify the dynamic programming problem. Lastly, benefit formulae have not become significantly more or less generous between 1987 and the end of the sample period (although there have been reductions in the Social Security work disincentives).

First, increased labour income leads to increased Social Security benefits, but only for the first 35 years in the labour-market. Social Security benefits depend upon Average Indexed Monthly Earnings, or $AIME_t$, which is average earnings in the 35 highest earnings years. However, after the first 35 years in the labour-market, $AIME_t$ is only recomputed upwards if current earnings are greater than earnings in a previous year of work. French (2003) describes the computation of $AIME_t$.

Second, there are incentives to begin drawing Social Security benefits by age 65. Individuals are ineligible for Social Security benefits before age 62. Upon application for benefits the individual receives them until death. Once the individual has applied for Social Security benefits, the benefits depend on a progressive function of AIME and the year the individual starts drawing benefits. For every year before age 65 the individual applies for benefits, benefits are reduced by 6·7%. This is roughly actuarially fair. For every year between ages 65 and 70 that benefit application is delayed, benefits rise by 3%.[4] This is actuarially unfair and thus generates an incentive to draw benefits by age 65.

Third, the Social Security earnings test taxes labour income for Social Security beneficiaries at a very high rate. If a beneficiary younger than age 70 earns more labour income than a "test" threshold level of \$6000, benefits are taxed at a 50% rate until all benefits have been taxed away. Moreover, the earnings test tax on benefits is in addition to Federal and state income and payroll taxes. Therefore, the marginal tax rate an individual faces is the sum of Federal, state, and payroll marginal tax rates, plus 50%. The incentive to draw benefits by age 65 in combination with the Social Security earnings test for Social Security beneficiaries is a major disincentive for work after age 65.

A common misconception is that the recomputation formulae fully replace benefits lost through the earnings test. Although this is roughly true between ages 62 and 65, a loss of one year's benefits results in only a small upward revision in future benefits after age 65. If a year's worth of benefits are taxed away between 62 and 65, benefits in the future will be raised by 6·7%. If a year's worth of benefits are taxed away between 65 and 70, benefits in the future will be raised by 3%.

The formula for Social Security benefits in the asset accumulation equation (9) captures all of these incentives.

2.3. *Pensions*

Pensions are like Social Security in two important respects. First, pension wealth is illiquid until the early retirement age, which is usually 55, 60, or 62 depending on the pension plan.[5] Therefore, I assume that pension wealth is illiquid until age 62. Second, pension benefits depend on the individual's work history. Because of this, pension benefits are assumed to be a function of $AIME_t$, just like Social Security benefits.

However, pensions are different from Social Security in their age-specific incentives to leave the labour force. Defined benefit pension plans are typically structured in a way that encourages a worker to remain at a firm until the early retirement age[6] and to leave the firm no later than the

4. I reduce AIME for individuals who first receive benefits before age 65. For example, if an individual begins drawing benefits at age 62 we can adjust AIME to account for early retirement. We know that adjusted AIME must result in benefits that are 80% of what they would have been had the individual first received benefits at age 65. Therefore, age at application for Social Security benefits need not be treated as a state variable. However, as pointed out in Section 2.3, pension benefits also depend on AIME, and the date of receipt of Social Security benefits should not affect pension benefits. This adds slightly to the complexity of the AIME adjustment.

5. Although it is often possible to "cash out" of pension plans, there are often penalties for doing so. For example, there are tax penalties for drawing defined contribution wealth before age $59\frac{1}{2}$, except in certain hardship cases.

6. Ippolito (1997) describes these incentives in detail.

normal retirement age (usually 62 or 65). These incentives, as well as my approach to modelling these incentives, are described below.

Pension benefits typically depend on years of service at the firm, the highest annual earnings at the firm (usually the average of the five highest earnings years), and a formula that depends on age and years of service at the firm. The part of the pension formula that depends on age and years of service generates the incentive to leave the firm by the normal retirement age. Up to the normal retirement age, this pension formula component increases with age. After the normal retirement age, it does not. Therefore, delaying exit from the firm after the normal retirement age results in a reduction in the present value of pension benefits. Although delaying benefit receipt causes slightly higher annual benefits (because years of service at the firm have increased), the individual will receive benefits for fewer years.

In order to account for the high pension accrual for those in their 50s and the lower pension accrual at other ages, I take estimates of age-specific accrual rates from Gustman, Mitchell, Samwick and Steinmeier (1998). In order to account for the fact that pension benefits are illiquid, I assume that benefits, pb_t, depend only on $AIME_t$. Unfortunately, the formula for $AIME_t$ does not account for the high accrual rates (measured as the effect of working an additional year increases the present value of pension wealth) for individuals in their 50s or the low accrual rates at other ages. To account for this problem, the variable ε_t represents the difference between using the two different approaches. Thus ε_t is negative at younger ages and is positive at older ages. Nevertheless, the average of ε_t is less than \$1000 at almost every age and is less than 20% of average pension accrual at almost every age. It is treated as labour income in the asset accumulation equation (9).[7]

One final aspect of pensions is worth noting. Accrual rates tend to be higher for high wage workers than for low wage workers. There are two reasons for this. First, the formulae for many defined benefit plans explicitly have higher accrual rates for high wage workers. Second, a higher share of high wage workers tend to have pension plans. I account for this by modelling pb_t as a regressive function of $AIME_t$.

2.4. Heterogeneity and model solution

Optimal decisions depend on the state variables, denoted $X_t = (A_t, W_t, B_t, M_t, AIME_t)$,[8] preferences denoted $\theta = (\gamma, \nu, \theta_P, \theta_B, \phi, L, \beta)$, and the parameters that determine the data generating process for the state variables denoted $\chi = (r, \sigma_\eta^2, \alpha, \rho, W(M_t, t+1), \{\text{prob}(M_{t+1} \mid M_t, t)\}_{t=1}^T, \{S_t\}_{t=1}^T, Y(\cdot,\cdot), \{ys_t\}_{t=1}^T, \{pb_t\}_{t=1}^T, \{ss_t\}_{t=1}^T)$.

The value function is the solution to

$$\begin{aligned} V_t(X_t) = \max_{C_t, H_t, B_t} \Bigg\{ & \frac{1}{1-\nu}\left(C_t^\gamma (L - H_t - \theta_P P_t - \phi 1\{M = \text{bad}\})^{1-\gamma}\right)^{1-\nu} + \beta s_{t+1} \\ & \times \sum\nolimits_{M \in \{\text{good},\text{bad}\}} \int V_{t+1}(X_{t+1}) dF(W_{t+1} \mid M_{t+1}, W_t, t) \text{prob}(M_{t+1} \mid M_t, t) \\ & + \beta(1 - s_{t+1}) b(A_{t+1}) \Bigg\}, \end{aligned} \tag{10}$$

where $F(\cdot|\cdot,\cdot,\cdot)$ is the conditional cdf of next period's wages. The decision rules are solved recursively, starting at time T and working backwards. Since there is no closed form solution

7. Note that this method of accounting for pension accrual only leads to model mis-specification if liquidity constraints and variable marginal tax rates affect behaviour.

8. Pension wealth and spousal income depend on the other state variables and are thus not state variables themselves.

to the problem, the state variables are discretized into a finite number of points on a grid and the value function is evaluated at those points. I use linear interpolation within the grid and extrapolation outside of the grid to evaluate the value function points that were not directly computed. I integrate the value function with respect to the innovation in the wage using quadrature.

Because the fixed cost of work and the benefit application decision mean that the value function need not be globally concave, I cannot use relatively fast hill-climbing algorithms. Therefore, I also discretize the consumption and labour supply decisions and use a grid search technique to find the optimal consumption and hours rules. The grids seem to produce reasonable approximations.

The model solution procedure allows for heterogeneity in the state variables, X_{it}, where i indexes individuals. However, the requirement of computational simplicity does not allow for heterogeneity in preferences θ or in the data generating process for the state variables χ. I assume that individual i responds only to (X_{it}, θ, χ). Different realizations of the stochastic shocks means that wages and health status will differ across individuals, so there may be differences in consumption, labour supply, and benefit application decisions across individuals. However, given the same age, wage, health status, asset level, Social Security application status, and AIME, different individuals will make the same decisions.

3. ESTIMATION

This section describes the method of simulated moments (MSM) estimation strategy. The goal is to estimate the preferences θ given the data generating process for the exogenous state variables χ. Because it would be too computationally burdensome to estimate all parameters simultaneously, I use a two-step strategy. In the first step, I estimate some elements of χ and calibrate others. I assume rational expectations, meaning that individuals know their own state variables X_t at time t, the Markov process that determines their state variables, which is parameterized by χ, and optimize accordingly. In the second step, I use the numerical methods described in Section 2.4 and the estimated data generating process for the state variables to simulate life cycle profiles for a large number of hypothetical individuals. The goal is to find preference parameters that generate simulated profiles that match the profiles estimated from the data.

The next subsection describes the MSM technique in more detail. The following subsections describe construction of the sample profiles that I match to the simulated profiles as well as estimation of some of the elements of χ.

3.1. *Estimation of preferences: the method of simulated moments*

The MSM estimation strategy matches mean assets, hours of work, participation, and also median assets in the PSID to the corresponding moments of the same variables in a simulated sample. The "matching" of moments is done using standard GMM techniques. Because of problems with measurement error, I do not match high order moments.[9] Using means, however, averages out measurement error, as shown below.

The objective is to find a vector of preferences $\theta \in \Theta$ that simulates profiles that "look like" (as measured by a GMM criterion function) the profiles from the data. I assume $\Theta \subset \mathbb{R}^7$

9. See Altonji (1986), Abowd and Card (1989) and French (2004) for attempts to overcome the measurement error problems that plague high frequency analyses of labour supply.

where Θ is a compact set. I assume the PSID data are generated by the model in Section 2, plus measurement error in hours:

$$A_{it} = A_t(X_{it-1}, \theta, \chi), \tag{11}$$

$$\ln H_{it} = \ln H_t(X_{it}, \theta, \chi) + \epsilon_{iHt} \qquad \text{if } P_{it} > 0, \tag{12}$$

$$P_{it} = P_t(X_{it}, \theta, \chi), \tag{13}$$

where A_{it}, $\ln H_{it}$, and P_{it} represent individual i's measured assets, log of hours worked, and participation decision at time t, and ϵ_{iHt} represents measurement error in hours. I assume zero mean measurement error in hours, $E[\epsilon_{iHt} \mid M_{it}, t] = 0$.[10] The computation of $A_t(X_{it-1}, \theta, \chi)$, $\ln H_t(X_{it}, \theta, \chi)$, and $P_t(X_{it}, \theta, \chi)$ is described in Section 2.4.

Assuming that the distribution of the state variables is the same in both the simulations and the data, it is possible to generate moment conditions for median and mean assets as well as mean participation and hours worked conditional upon health status, resulting in the following $6T$ moment conditions:

$$E\big[I\{A_{it} \le \text{median}(A_t(X, \theta, \chi))\} - \tfrac{1}{2} \mid t\big] = 0, \qquad \text{for all } t \in \{1, \dots, T\}, \tag{14}$$

$$E[A_{it} \mid t] - \int A_t(X, \theta, \chi) dF_{t-1}(X \mid t) = 0, \qquad \text{for all } t \in \{1, \dots, T\}, \tag{15}$$

$$E[\ln H_{iMt} \mid M, t] - \int \ln H_t(X, \theta, \chi) dF_{Mt}(X \mid M, t) = 0, \qquad \text{for all } t \in \{1, \dots, T\}, \; M \in \{\text{good}, \text{bad}\}, \tag{16}$$

$$E[P_{iMt} \mid M, t] - \int P_t(X, \theta, \chi) dF_{Mt}(X \mid M, t) = 0, \qquad \text{for all } t \in \{1, \dots, T\}, \; M \in \{\text{good}, \text{bad}\}, \tag{17}$$

where median$(A_t(X, \theta, \chi))$ is the median[11] of the distribution of simulated assets $A_t(X, \theta, \chi)$, $I\{\cdot\}$ is the indicator function, equal to 1 if true, $F_t(X)$ is the cdf of the state variables at time t, and $F_{Mt}(X \mid M)$ is the cdf of the state variables at time t given health status M. The integrals in equations (14)–(17) are computed using Monte Carlo integration. When evaluated at the true preference parameters and the true distribution of the state variables, conditional on age and, in the case of hours and participation, health status, the difference between the data moment and the simulated moment has an expected value of zero.

In summary, the MSM procedure I use can be described as follows. First, I estimate the life cycle profiles for hours worked, labour force participation, and assets from the PSID data. Second, using the same data I use to estimate profiles, I estimate the data generating processes for health status and wages following the estimation techniques described in Sections 3.4 and 4.2. Third, I use the estimated data generating processes to simulate matrices for random health and wage shocks as well as an initial distribution for health, wages, assets, and AIME. These are sequences of lifetime shocks for 5000 simulated individuals, so there is a $5000 \times T$ matrix of health shocks and a $5000 \times T$ matrix of wage shocks. Fourth, I pick an arbitrary vector of preference parameters and compute the decision rules given those parameters and the numerical methods described in Section 2.4. The fifth step is to use the decision rules and the health and wage shocks to simulate hypothetical life cycle profiles for the decision variables. Sixth, the simulated data and the true data are aggregated by age (and in the case of hours and

10. I also allow for zero mean measurement error in participation, conditional on age and health status. The mean asset condition in equation (15) also holds if there is zero mean measurement error in assets. However, the median condition (14) will not typically hold in the presence of measurement error. Nevertheless, dropping the median condition (14) and re-estimating the model does not have a large effect on parameter estimates.

11. See French and Jones (2004) for more on using quantile conditions in a GMM framework.

participation, by health status). Seventh, the difference between the simulated and true profiles is computed and the differences are weighted up to form a distance measure. Finally, a new vector of preference parameters is picked and the whole process is repeated.[12] The preference parameters that minimize the distance between the data moments and the simulated moments described in equations (15)–(17) are the estimated parameters, $\hat{\theta}$. I discuss the distribution of the parameter estimates, the weighting matrix, and the overidentification tests in the Appendix.

3.2. *Estimation of profiles*

This section describes the life cycle profiles for assets, hours, and participation rates to be fed into equations (14)–(17) as well as the life cycle wage profile. When constructing profiles that account for age and health effects, I am concerned about the presence of individual-specific effects, year effects, and family size effects. To generate profiles, I estimate equation (18), where Z_{it} represents an observation for either assets A_{it}, hours $\ln H_{it}$, participation P_{it}, or wages (net of the tied wage-hours effect) $\ln W_{it} - \alpha \ln H_{it}$[13] for individual i at age t:

$$\begin{aligned} Z_{it} = f_i &+ \sum\nolimits_{k=1}^{T} \Pi_{gk} I\{\text{age}_{it} = k\} \times \text{prob}(M_{it} = \text{good} \mid M_{it}^*) \\ &+ \sum\nolimits_{k=1}^{T} \Pi_{bk} I\{\text{age}_{it} = k\} \times \text{prob}(M_{it} = \text{bad} \mid M_{it}^*) \\ &+ \sum\nolimits_{f=1}^{F} \Pi_f \text{famsize}_{it} + \Pi_U U_t + u_{it}, \end{aligned} \tag{18}$$

where f_i is an individual-specific effect, famsize_{it} is family size, U_t is the unemployment rate, $\{\Pi_{gk}\}_{k=1}^{T}\{\Pi_{bk}\}_{k=1}^{T}$, $\{\Pi_f\}_{f=1}^{F}$, and Π_U are parameters, $\text{prob}(M_{it} = \text{bad} \mid M_{it}^*)$ is the probability that health is bad given a noisy health measure M_{it}^* and $\text{prob}(M_{it} = \text{good} \mid M_{it}^*) = 1 - \text{prob}(M_{it} = \text{bad} \mid M_{it}^*)$. French (2001) describes the construction of $\text{prob}(M_{it} = \text{bad} \mid M_{it}^*)$. If M_{it}^* were perfectly measured, then $\text{prob}(M_{it} = \text{bad} \mid M_{it}^*)$ would collapse to a dummy variable. I estimate equation (18) using fixed-effects to control for the individual-specific effect, f_i. I use a full set of age dummy variables when estimating the hours, participation, and asset profiles; however, the wage profile is estimated using a fourth order polynomial in age.[14] For the asset profiles I assume $\Pi_{gk} = \Pi_{bk}$ for all k; that is, I do not condition on health status when generating the asset profile. I use a full set of dummy variables for family size famsize_{it}.

I use the age effects and health effects from equation (18) to generate the data profiles that I will match to the simulated profiles. I set family size equal to three and the unemployment rate to 6·5%, and use the mean individual-specific effect for individuals who were born in 1940, who are age 50, and have the average level of health for 50 year olds (see Appendix). Note that this approach controls for cohort effects. The cohort effect is just the average fixed-effect of all individuals in a single cohort.

12. I use simplex methods to search over Θ. Because the local minimum of the GMM criterion function need not be the global minimum, I try many different starting values. I check to see whether the algorithm will find the global minimum by simulating individuals at assumed parameter values and treating these simulated individuals as data. I then simulate another set of individuals with different wage and health shocks and with different initial utility function parameters. I then use the MSM algorithm to match the second set of simulated individuals to the first set of "data". I find that preference parameters estimated for the second set of individuals come very close to the "true" preference parameters of the first set. Nevertheless, estimated preference parameters usually do not come within two standard errors of the true parameters. This shows that standard errors are underestimated. Footnote 37 provides further evidence that the standard errors are underestimated.

13. Note that this identifies $W(M_{it}, t)$.

14. When creating profiles with the polynomials, I estimate the polynomial using data on individuals five years younger and 10 years older than my sample of interest. This overcomes some of the endpoint problems associated with polynomial smoothing.

3.3. Accounting for selection in the wage profiles

Unfortunately, the fixed-effects estimator does not overcome an important selection problem in the wage equation. The problem is that fixed-effects estimators use wage observations for workers but do not use the potential wages of non-workers. Because the fixed-effects estimator demeans the average level of wages for each individual in the sample, it identifies the growth rate of wages of individuals while working. Because the fixed-effects estimator identifies the individual-level growth rates of workers' wages, composition bias problems—the question of whether high wage or low wage individuals drop out of the labour-market—is not a problem if wage growth rates for workers and non-workers are the same.[15] However, if individuals leave the market because of a sudden wage drop, such as from job loss, then wage growth rates for workers will be greater than wage growth for non-workers. This problem will bias wage growth upward.

In order to account for the selection problem, it is important to distinguish between three separate objects. The first is the unobserved average wage profile for all individuals that is the object of interest.[16] The second is the fixed-effects wage profile estimated using the actual data on workers—this profile is biased for reasons discussed above. The third is the fixed-effects profile using simulated workers.[17] This profile is biased for the same reason that the profile using actual data on workers is biased.

To correct the bias, I assume that the bias in the fixed-effects wage profiles of workers will be the same in both the actual PSID and simulated data.[18] First, I feed the estimated (and biased) fixed-effects wage profile into the model. Second, I solve and simulate the model and estimate the fixed-effects wage profiles for both simulated workers and all simulated individuals. Third, I compute the difference between the profiles for both simulated workers and all simulated individuals so that I can estimate the extent to which growth rates in wages are overestimated by using only simulated workers instead of all simulated individuals. I then use this estimate of the selection bias in the simulated wage profile to infer the extent of selection bias in the PSID data wage profile.

If, for example, the fixed-effects wage profiles overstate average wages at age 60 by 10% in the simulated sample, then it is likely that wages have been overestimated at age 60 by 10% in the PSID data. Therefore, the candidate for the unobserved average wage at age 60 is the fixed-effects estimate from the PSID data, less 10%. This new candidate wage profile is fed into the model and the procedure is repeated. If, for example, the fixed-effects profile using simulated data still indicates a 1% upward bias, the true candidate wage profile is reduced by an additional 1%. This iterative process is continued until a fixed point is found.[19]

Once the process converges, the estimated wage profile for all individuals is fed into the model and preference parameters are estimated using the method of simulated moments. Upon re-estimation of the model parameters, the selection bias is recomputed and the wage profiles are updated. The model parameters are then estimated again.

15. This is an important advantage of panel data over cross-sectional data used by Blundell, Reed and Stoker (2003), for example. Cross-sectional estimators, such as OLS, mix the true wage growth of individuals with spurious wage growth caused by differences in the level of wages between those who enter, exit, and remain in the labour force.

16. This object is $W(M_t, t) = E[\ln W_t - \alpha \ln H_t \mid M, t]$.

17. This object converges to $W(M_t, t) + E[AR_t \mid M, t, H_t > 0]$ as the number of simulations becomes arbitrarily large.

18. This is true if the simulated individuals have the same wage generating process, the same distribution of state variables, and have the same preferences as the individuals in the data.

19. If the value function were concave, it would be possible to prove that this iterative mapping was a contraction. This cannot be proven analytically and in general cannot be proven numerically. However, based upon carefully conducted computations it seems that a unique solution exists.

3.4. Estimation of the health transition matrix

When estimating the health transition matrix in equation (5), I am concerned with both the presence of measurement error in health status and the presence of individual heterogeneity. In order to address both of these concerns I estimate the linear probability model:

$$\begin{aligned}\text{Prob}(M_{it} = \text{good} \mid \alpha_i, M_{it-1}, t) = \alpha_i + \textstyle\sum_{k=0}^{K} \beta_k t^k \times \text{prob}(M_{it-1} = \text{good} \mid M^*_{it-1}) \\ + \textstyle\sum_{k=1}^{K} \gamma_k t^k \times \text{prob}(M_{it-1} = \text{bad} \mid M^*_{it-1}) + \epsilon_{it},\end{aligned} \tag{19}$$

where α_i represents individual heterogeneity in capacity for good health. OLS estimates will be inconsistent for two reasons. First, α_i is correlated with previous health status. In order to circumvent this problem, I first difference equation (19), then use lags of health status and health status interacted with age to instrument for last year's health status change and last year's health status change interacted with age. Second, health status is measured with non-zero mean error. As in equation (18), I take estimates of $\text{prob}(M_{it-1} = \text{good} \mid M^*_{it-1})$ from French (2001). When constructing the health status transition matrix, I set α_i equal to its average level for individuals born in 1940.

4. DATA AND CALIBRATIONS

4.1. Data

I use the Panel Study of Income Dynamics (PSID) for the years 1968–1997. I drop the Survey of Economic Opportunity subsample of poor and minorities to make the data more representative of the U.S. population. Because I model the behaviour of a head of household, I use labour supply variables for the male head of household and household-level asset data.

When estimating the hours worked and labour force participation rate profiles, I use individuals born between 1922 and 1940, resulting in 18,690 person-year observations for labour force participation rates and 15,766 person-year observations for hours worked. For the asset profile, I use individuals born between 1902 and 1965 to increase the sample size, resulting in 8265 person-year observations. For the wage and health profiles, I use the full sample, resulting in 60,714 and 69,347 person-year wage and health observations.

I estimate the asset profile using 1984, 1989, and 1994 PSID wealth surveys. Because I do not wish my estimate of assets to be affected by the extremely wealthy, many of whom inherit their wealth, I exclude observations with over \$1,000,000 in assets. Households in which an entering family member brought assets into the household or an exiting family member took assets out of the household are dropped. The PSID asset measure is fairly comprehensive. It includes real estate, the value of a farm or business, vehicles, stocks, mutual funds, IRAs, Keoghs, liquid assets, bonds, other assets, and investment trusts less mortgages and other debts. It does not include pension or Social Security wealth.

Wages are computed as annual earnings divided by hours and are dropped if wages are less than \$3 per hour or greater than \$100 per hour, 1987 dollars. Hours are counted as zero if measured hours are below 300 hours worked per year.

The PSID has only one measure of health that is asked during all years of the panel. It is the self-reported response to "Do you have any physical or nervous condition that limits the type of work or the amount of work that you can do?" A criticism of self-reported health measures is that respondents often report "bad health" in order to justify being out of the labour force. This will lead me to overestimate the effect of health upon work-hours. Alternatively, the coarse discretization of health status into good and bad when true health status is likely a continuous variable potentially causes measurement error, biasing the effect of health status on different variables to zero effect.

TABLE 1

The variance and persistence of innovations to the wage

Parameter	Variable	Estimate	S.E.
σ_{η}^{2}	Variance of the innovation in wages	0·0141	0·0014
ρ	Autoregressive coefficient of wages	0·977	0·017

The PSID has poor information on mortality statistics. Therefore, I combine PSID data with mortality statistics from the National Center for Health Statistics (NCHS).[20] These statistics use the entire U.S. population as their sample.

4.2. *Remaining calibrations*

In order to estimate preference parameters, I calibrate some of the parameters that determine the data generating process for the state variables χ. These are the parameters that determine the stochastic component of wages $(\sigma_{\eta}^{2}, \rho)$, the effect of work-hours on wages α, the interest rate r, and spousal income.

The parameters from the wage equation $(\sigma_{\eta}^{2}, \rho)$, shown in Table 1, were estimated using equations (6) and (7) and minimum distance techniques. The model of wages allows for a MA(1) measurement error component.[21] The results indicate that $\rho = 0{\cdot}977$; wages are almost a random walk. The estimate of σ_{η}^{2} is 0·0141; one standard deviation of an innovation in the wage is 12% of wages.[22] These estimates imply that long run forecast errors may be large.

The coefficient α, which parameterizes the part-time wage penalty, is set at 0·415 and is similar to the findings of Gustman and Steinmeier (1986) and Aaronson and French (2004). This implies that part-time workers (who work 1000 hours per year) earn 25% less per hour than full-time (2000 hours per year) workers. Controlling for the fact that part-time workers make less per hour than full-time workers eliminates most of the wage declines after age 60 that are shown in Figure 1. The remaining calibrations are as follows. I set the pre-tax interest rate at $r = 0{\cdot}04$ and the age at which individuals receive pension benefits at age 62. Following De Nardi (2004), the object that determines the curvature of the bequest function, K, is set equal to \$500,000. Spousal

20. I compute mortality rates given last year's health status using Bayes' rule:

$$\text{prob}(\text{death}_t \mid M_{t-1} = \text{good}) = \frac{\text{prob}(M_{t-1} = \text{good} \mid \text{death}_t)}{\text{prob}(M_{t-1} = \text{good})} \times \text{prob}(\text{death}_t). \tag{20}$$

I compute $\text{prob}(M_{t-1} = \text{good} \mid \text{death}_t)$ and $\text{prob}(M_{t-1} = \text{good})$ using PSID data, and $\text{prob}(\text{death}_t)$ from the NCHS data. When using PSID data, the estimate of $\text{prob}(\text{death}_t)$ is about 25% lower than when using NCHS data, indicating that the PSID underestimates mortality rates by 25%.

21. In order to obtain these estimates, I use wage residuals from the regression in equation (18) with the fixed effect added back:

$$f_i + u_{it} = AR_{it} + \xi_{it} + \theta\xi_{it-1}, \tag{21}$$

where ξ_{it} is assumed to be measurement error with MA(1) coefficient θ. The AR(1) component, AR_{it}, is potentially non-stationary. French (2004) finds that most of the variance of the MA(1) component of wages is measurement error, so assuming all of the variance of the MA(1) component of wages is measurement error seems reasonable. The estimates in Table 1 are similar to other estimates in the literature, although $\rho = 0{\cdot}977$ is likely at the high end of the range. Card (1994) also finds that wages follow a highly persistent AR(1) process.

22. I scaled up the variance of the innovations to reflect the additional uncertainty due to the aggregate unemployment rate using estimates of the variance of the unemployment rate and Π_U from equation (18). However, the amount of volatility in wages associated with the unemployment rate is tiny, and thus this procedure had only a small effect on σ_{η}^{2}.

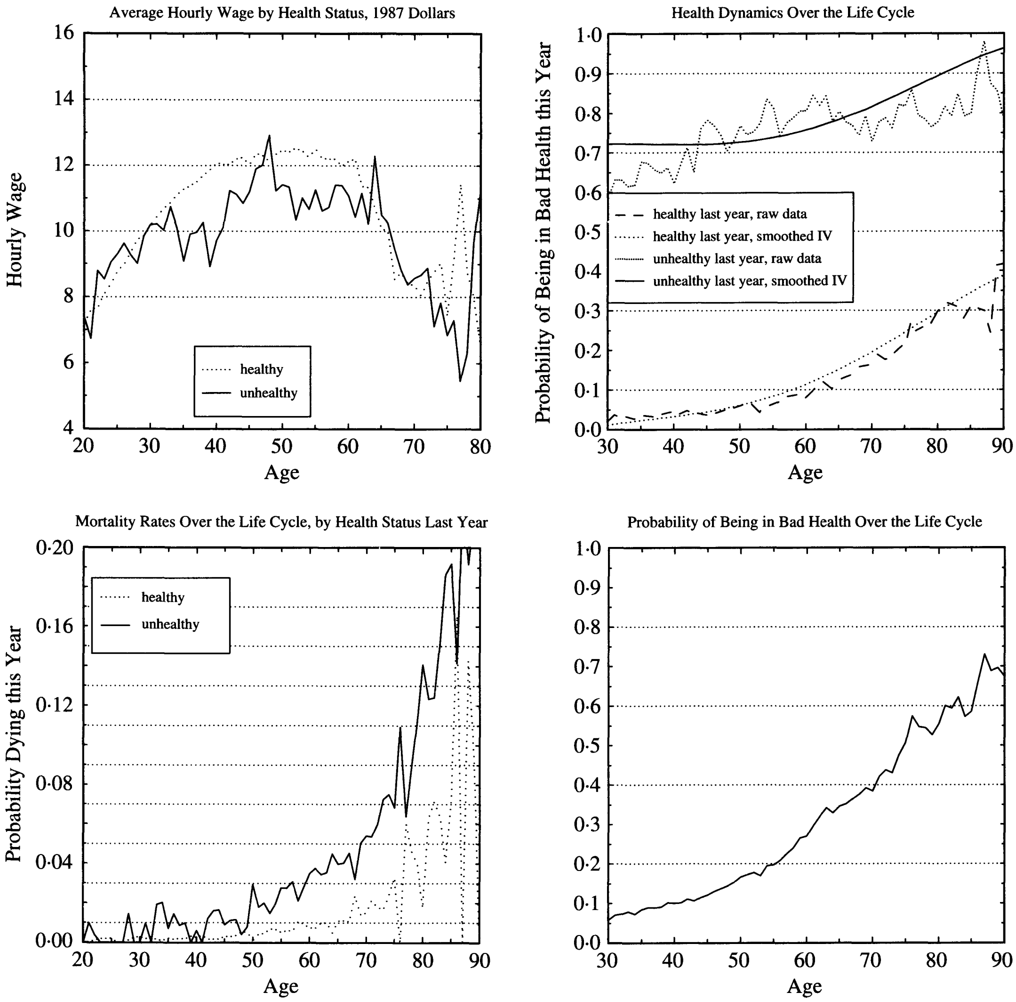

FIGURE 1

Life cycle profiles for exogenous state variables

income is assumed to follow a polynomial in age and the log of the wage.[23] Because the PSID has poor information on pensions and (until the most recent waves) Social Security, I use spousal income when young to predict spousal pension and Social Security benefits when old.

5. RESULTS

The estimated inputs into the MSM algorithm can be divided into data on the exogenous state variables and data on decision variables. The data generating process for the exogenous state variables, parameterized by the vector χ, includes growth rates for wages conditional on health status, health transition matrices, and mortality probabilities. The decision variables are the

23. I regress spouse's income on the husband's log wage (instrumented using education), an age polynomial, and a set of cohort dummy variables. When I construct the spousal income profile, I set the cohort effect equal to those born in 1940.

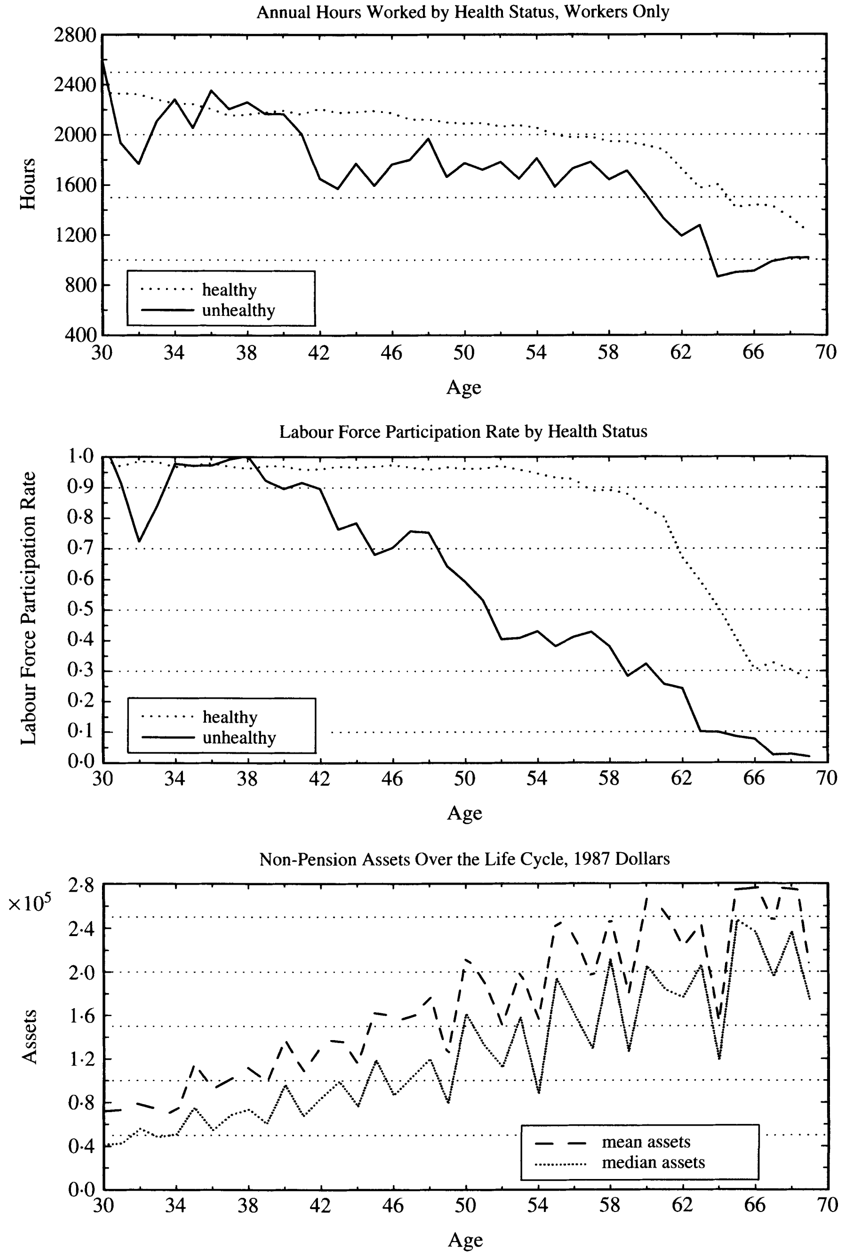

FIGURE 2

Life cycle profiles for decision variables

profiles for hours worked per year (by those who worked), assets, and labour force participation. In order to identify the role of health in explaining the decline of hours near the end of the life cycle, profiles for hours and labour force participation rates are shown for individuals in both good and bad health.

5.1. *Profiles for the exogenous state variables*

This section describes the profiles for wages, health transition matrices, and the survivor probabilities. I use smoothed versions of the profiles when estimating preferences. However, I display unsmoothed profiles to show that the profiles are precisely measured, as displayed by their smooth appearance. On average, profiles for healthy individuals are smoother than for unhealthy individuals. This is because there are more observations on healthy individuals than on unhealthy individuals.

Using the methodology from Section 3.2, the top left panel of Figure 1 displays wage profiles for males by age and health status. Most striking is the hump shape of the wage profiles for both health groups, with wages peaking near age 55. Fixed-effects estimates show a more rapid decline in wages after age 55 than do OLS estimates (see Ghez and Becker (1975), Heckman (1976) and Browning, Deaton and Irish (1985) for profiles constructed using OLS). The reason for this is that high wage individuals tend to remain in the labour force until older ages than do low wage workers. Therefore, OLS estimates suffer from "composition bias" problems, where wage observations will be for all workers at age 55 but only for high wage workers at age 65. Also striking is the small effect of health on wages. Fixed-effects estimates show a smaller role for health than OLS. There are three alternative explanations for the difference between OLS estimates and fixed-effect estimates. First, it may be that some other factor (*e.g.* childhood poverty) causes both poor health and low wages. Second, the Grossman model (Grossman, 1972) predicts that individuals who have higher expected lifetime wages invest more in health human capital when young. Therefore, the Grossman model implies that high wages cause good health, and not vice versa as most interpretations of an OLS regression of wages on health assume. The third explanation for the small estimated effect of health upon wages could be related to a selection problem. It may be only the individuals who get lucky in the labour-market who remain in the labour-market after a bad health shock. Section 5.5 discusses this third point in greater detail.

The two R.H.S. panels of Figure 1 show how health dynamics change over the life cycle. Until age 55, very few individuals experience a change from good health to bad health. This begins to change after age 55, with individuals becoming more and more likely to move from good health to bad health. Note, however, there is no rapid shift in population health that takes place only between ages 55 and 70, the ages at which labour force participation declines most rapidly. Instead, much of the decline in population health takes place after age 70.

Lastly, the lower left panel shows mortality rates over the life cycle. Unsurprisingly, individuals in bad health have higher mortality rates than individuals in good health.

5.2. *Decision profiles*

This section describes the profiles for the decision variables. To recover preference parameters, I make two fundamental identifying assumptions. First, changes in work-hours and consumption affect neither health nor wages (other than through the tied wage-hours effect). Second, preferences depend only upon health and family size. Preferences change with age, but only as a result of changes in health and family size. Therefore, age can be thought of as an "exclusion restriction" which causes changes in the incentives for work and savings but does not change preferences.

The top panel of Figure 2 shows the life cycle profiles for hours worked for men in good and bad health. At any point in the life cycle, the effect of health on hours worked is sizeable, but health only explains a small amount of the variation in work-hours over the life cycle. Hours worked begins to decline rapidly after age 59.[24] This is true even when conditioning on health status, so it appears that health status alone must have a small causal role in the decline in the number of hours worked by workers near retirement.

The middle panel of Figure 2 shows the life cycle profiles for labour force participation. Health appears to affect labour force participation rates more than hours worked. However, the effect is still modest. The fraction of all individuals at age 55 who report bad health is 20% and this rises to only 37% by age 70. Therefore, the change in labour force participation rates attributable to changes in health between ages 55 and 70 is small. The effect can be quantified using the equation

$$\sum_{t=56}^{70} \Delta P_t = \sum_{t=56}^{70} \left(\frac{\Delta P}{\Delta M}\right)_t \Delta M_t, \tag{22}$$

where Δ is the first difference operator, $\left(\frac{\Delta P}{\Delta M}\right)_t$ is the estimated effect of health on participation (estimated by the vertical difference between the upper and lower profiles in the middle panel in Figure 2) at age t, and ΔM_t is the change in population health status between age $t-1$ and t (estimated using the bottom right panel of Figure 1). This technique suggests that declining health between ages 55 and 70 can explain a 7% drop in labour force participation rates. Thus, of the drop in labour force participation rates from 87% to 13% between ages 55 and 70, only 10% can be attributed to the (admittedly crude) health measure. Moreover, the ages at which hours and labour force participation rates decline most rapidly coincides with those ages at which wages decline and at which there are large pension and Social Security work disincentives. For example, labour force participation drops 9 percentage points (or 13%) at age 62 and 7 percentage points (or 18%) at age 65.[25] Therefore, it seems that wages, pensions, and Social Security potentially play a strong role in determining the age of retirement.

The estimated effect of health on wages, hours worked, and labour force participation rates is average for the literature (see Currie and Madrian, 1999). As with most studies, I find a statistically significant effect of health. However, this is the first study to predict what fraction of the life cycle variation in wages, hours worked, and labour force participation rates is explained by health status. The above analysis shows that the amount of explained variation is small.

Finally, the bottom panel of Figure 2 shows both mean and median assets over the life cycle.[26] Note that young people do save. A certainty life cycle model with typical parameters predicts that people dissave when young since wage levels are very low when young. Therefore, the life cycle asset profile is evidence against the standard certainty-equivalent life cycle model. However, it is consistent with a model in which young people save in order to generate a buffer stock of assets for insurance against bad wage shocks when old (Gourinchas and Parker (2002), Cagetti (2003)).

24. This is similar to Browning *et al.* (1985), although Ghez and Becker (1975) find that the drop-off in hours is later in life. Ghez and Becker's result may be different because they use data from 1960, when the average retirement age was later.

25. Blau (1994) finds an even larger decline in labour force participation at age 65 using data from the Retirement History Survey.

26. Note that median assets are not much lower than mean assets. In Cagetti (2003), median assets are much lower than mean assets. His median asset profiles are similar to mine, but his mean asset profile implies a higher level of assets than mine. The difference arises because I topcode assets at $1,000,000, whereas he does not. Topcoding has a much larger effect on mean assets than median assets. Gourinchas and Parker's (2002) simulated asset profile is very different from those of Cagetti or myself. Their simulated asset profile implies that assets are almost zero until age 45. This seems to be the result of difficulties measuring savings using income and consumption data.

TABLE 2

Preference parameter estimates

Parameter and definition	Specification (1)	(2)	(3)	(4)
γ Consumption weight	0·578 (0·003)	0·602 (0·003)	0·533 (0·003)	0· 615 (0·004)
ν Coefficient of relative risk aversion, utility	3·34 (0·07)	3·78 (0·07)	3·19 (0·05)	7·69 (0·15)
β Time discount factor	0·992 (0·002)	0·985 (0·002)	0·981 (0·001)	1·04 (0·004)
L Leisure endowment	4466 (30)	4889 (32)	3900 (24)	3399 (28)
ϕ Hours of leisure lost, bad health	318 (9)	191 (7)	196 (8)	202 (6)
θ_P Fixed cost of work, in hours	1313 (14)	1292 (15)	335 (7)	240 (6)
θ_B Bequest weight	1·69 (0·05)	2·58 (0·07)	1·70 (0·04)	0·037 (0·001)
χ^2 Statistic: (233 degrees of freedom)	856	880	830	1036
$\epsilon_{h,W}(40)$ Labour supply elasticity, age 40	0·37	0·37	0·35	0·19
$\epsilon_{h,W}(60)$ Labour supply elasticity, age 60	1·24	1·33	1·10	1·04
Reservation hours level, age 62	885	916	1072	1051
Coefficient of relative risk aversion	2·35	2·68	2·17	5·11

Standard errors in parentheses

Specifications described below:
(1) Does not account for selection or tied wage-hours offers
(2) Accounts for selection but not tied wage-hours offers
(3) Accounts for tied wage-hours offers but not selection
(4) Accounts for selection and tied wage-hours offers

5.3. *Initial distributions*

To generate the simulated initial joint distribution of assets, wages, AIME,[27] and health status, I take random draws from the empirical joint distribution of assets, wages, AIME, and health status for individuals aged 29–31. I adjust the mean of log wages for healthy and unhealthy individuals to match the estimated life cycle fixed-effects profile for wages. Average assets at age 30 are equal to \$42,100 and are highly correlated with wages.

5.4. *Preference parameter estimates*

Table 2 presents estimates of the parameters in the utility function for males, aged 30–95. Because relatively little is known about the extent to which tied wage-hours offers and selection in the wage equation may affect parameter estimates, Table 2 presents parameter estimates given different assumptions about tied wage-hours offers selection. Because of a lack of data on older individuals, I assume that individuals do not work after age 70 and match moments only up to age 70.

One of the objects of interest in this paper is an individual's willingness to intertemporally substitute his work-hours. At age 40, the elasticity of simulated average hours worked given an anticipated transitory change in the wage is 0·19–0·37, depending upon the specification.[28] This labour supply elasticity increases with age. At age 60, the elasticity of simulated hours worked

27. I assume all individuals enter the labour force at age 25 and work 2000 hours per year at the age 30 wage to impute initial AIME.

28. This calculation was made by changing the wage by 20% for all simulated individuals of a given age, then computing the difference in total hours worked at that age. This causes a wealth effect, making the elasticity calculated herein smaller than the Frisch labour supply elasticity. Assuming certainty and interior conditions, the Frisch elasticity of leisure is $\frac{\gamma(1-\nu)-1}{\nu}$ and the Frisch elasticity of labour supply is $-\frac{L-H_t-\theta_P}{H_t} \times \frac{\gamma(1-\nu)-1}{\nu}$. However, one of the advantages of the dynamic programming approach is that it is not necessary to assume certainty or interior conditions.

given an anticipated transitory change in the wage is 1·04–1·33. This increase is due to the fixed cost of work generating volatility on the participation margin. This fixed cost, θ_P, varies between 240 and 1313 hours, depending on the specification. Wage changes cause relatively small hours changes for workers at both age 40 and age 60. However, the substitutability of labour supply at the participation margin rises with age. By age 60, many workers are close to indifferent between working and not working. Small changes in the wage cause large changes in the participation rate.

The fixed cost of work generates a reservation number of work-hours. Individuals will either work more than this many hours or will not work at all. The reservation number of work-hours depends on assets, wages, health status, and AIME. At age 62, for example, individuals never choose to work fewer than 885 hours per year in the baseline specification. This is similar to Cogan's (1981) estimate of 1000 hours per year. The fixed cost of work is identified by the life cycle profile of hours worked by workers. Note that the hours of work profiles, presented in Figure 2, do not drop below 1000 hours per year (or 20 hours per week) even though labour force participation rates decline to near zero. In the absence of a fixed cost of work, we should expect hours worked to parallel the decline in labour force participation. When estimating the model without fixed costs of work or tied wage-hours offers, hours worked tends to decline to about 500 hours per year for individuals aged 65–70. Moreover, the simulated labour supply elasticity rises very little over the life cycle.[29]

Most of the variation in the wage and labour supply profiles is from individuals aged 55–65. Figure 2 shows hours worked and labour force participation rates declining rapidly after age 60, even after the effect of health on hours has been addressed. This decline in hours coincides closely with the decline in wages, pension accrual, and the Social Security incentives to retire. Therefore, the evidence from older individuals indicates that labour supply is responsive to changes in economic incentives.

Many of the studies that estimate the intertemporal elasticity of substitution (Ghez and Becker (1975), MaCurdy (1981) and Browning *et al.* (1985)) obtain identification from a problematic source: the covariation of work-hours and wages of continuously employed young workers. Young workers work many hours although on average their wage is lower than the wage of older workers. This indicates that the intertemporal elasticity of substitution is small within a certainty-equivalent environment,[30] as hours change very little but wages change a lot over the life cycle. However, as Benitez-Silva (2000) and Low (2002) point out, younger workers may work many hours in order to develop enough assets to buffer themselves against the possibility of low wages when old.[31] Therefore, omission of uncertainty will potentially bias the estimated intertemporal elasticity of substitution downwards (see Domeij and Floden, 2002 for more on this point). Note that this problem is in addition to the problem of omitting the labour force participation margin. For these two reasons, previous studies have understated the substitutability of male labour supply.

The coefficient of relative risk aversion (or the inverse of the intertemporal elasticity) for consumption is 2·2–5·1 (depending on the specification),[32] which is similar to previous estimates

29. For example, in the specification that accounts for neither selection nor tied wage-hours offers, the elasticity rises from 0·45 at age 40 to 0·72 at age 60. In the specification that accounts for tied wage-hours offers but not selection, the elasticity rises from 0·32 to 0·77.

30. For example, Ghez and Becker (1975), MaCurdy (1981) and Browning *et al.* (1985) estimate the Frisch labour supply elasticity to be around 0·3 at all ages.

31. Both of these papers are similar to mine in that they solve dynamic programming models of labour supply and consumption decisions under wage uncertainty. However, neither paper considers the importance of fixed costs of work. As a result, neither paper seems to be totally successful in matching the decline in work-hours after age 60.

32. This is measured using the formula $-\frac{(\partial^2 U_t/\partial C_t^2)C_t}{\partial U_t/\partial C_t} = -(\gamma(1-\nu)-1)$. Note that this variable is measured holding labour supply fixed. The coefficient of relative risk aversion for consumption is poorly defined when the labour supply is flexible.

that rely on different methodologies (see Auerbach and Kotlikoff (1987) and Attanasio and Weber (1995) for reviews of the estimates). Identification of this parameter is similar to Cagetti (2003) who estimates a buffer stock model of consumption over the life cycle using asset data. Within this framework, a small estimate of the coefficient of relative risk aversion means that individuals save little given their level of assets and their level of uncertainty. If they were more risk averse, they would save more in order to buffer themselves against the risk of bad income shocks in the future. I also obtain identification from labour supply, as precautionary motives can explain why wages co-vary little with hours when young but a great deal when old. More risk averse individuals work more hours when young in order to accumulate a buffer stock of assets.

My estimate of the time discount factor, β, is larger than most estimates for three reasons. The first two reasons are clear upon inspection of the Euler equation: $\frac{\partial U_t}{\partial C_t} \geq \beta s_{t+1}(1 + r(1 - \tau_t))E_t \frac{\partial U_{t+1}}{\partial C_{t+1}}$, where τ_t is the marginal tax rate.[33] This equation identifies $\beta s_{t+1}(1 + r(1 - \tau_t))$, although not the elements of this equation separately. Therefore, a lower value of s_{t+1} or $(1 + r(1 - \tau_t))$ results in a higher value of β. The first reason for my high estimate of β is that most studies omit mortality risk. In my model, individuals discount the future not by the discount rate β, but by the discount factor multiplied by the survivor function s_{t+1}. Since the survivor function is necessarily less than one, omitting mortality risk will bias β downwards.[34] Second, the post-tax rate of return is smaller than the pre-tax rate of return. Therefore, omission of taxes should also bias β downwards. Third, the life cycle profile of hours shows that young individuals work many hours even though their wage, on average, is low. This is equivalent to stating that young people buy relatively little leisure, even though the price of leisure (their wage) is low. Between ages 35 and 60, people buy more leisure (*i.e.* work fewer hours) as they age even though their price of leisure (or wage) increases. Therefore, life cycle labour supply profiles provide evidence that individuals are patient. Ghez and Becker (1975), Heckman and MaCurdy (1980) and MaCurdy (1981) also find that $\beta(1 + r) > 1$ when using life cycle labour supply data.[35]

The bequest parameter θ_B varies a great deal across specifications. However, the marginal propensity to consume out of wealth in the final period—which is a non-linear function of θ_B, β, γ, ν, and K—is very stable across specifications. For low income individuals, the marginal propensity to consume is 1. For high income individuals, the marginal propensity to consume is between 0·025 and 0·045, depending on specification.

Figure 3 displays both the PSID profiles and the simulated life cycle profiles of the decision variables. It also displays 95% confidence intervals. Simulated profiles appear generally consistent with the data, although a χ^2 overidentification test rejects the model because the simulated profiles frequently lie outside the confidence intervals. There are some differences between the simulations and data that are worthy of mention. The model substantially overpredicts labour force participation rates for unhealthy individuals. This is potential evidence of one of two things. First, that the coarse discretization of health into good and bad is inadequate. It could be that bad health has only a small effect on the labour supply of some people and makes others completely unable to work. Alternatively, disability insurance provides benefits to those in bad health, but only if those people earn very little over the course of the year.[36] Therefore, disability insurance provides income to those who drop out of the labour-market but not to those who work part-time.

33. Note that this is not exactly correct when individuals value bequests. Also note that the Euler equation holds with equality when assets are positive.

34. Quantitatively, this effect is small. When I estimated parameters assuming that all individuals survive to age 65, the estimated value of β fell by less than 0·3 percentage points.

35. All of these papers ignore taxes and mortality.

36. Individuals are only eligible for disability benefits if their income is very low.

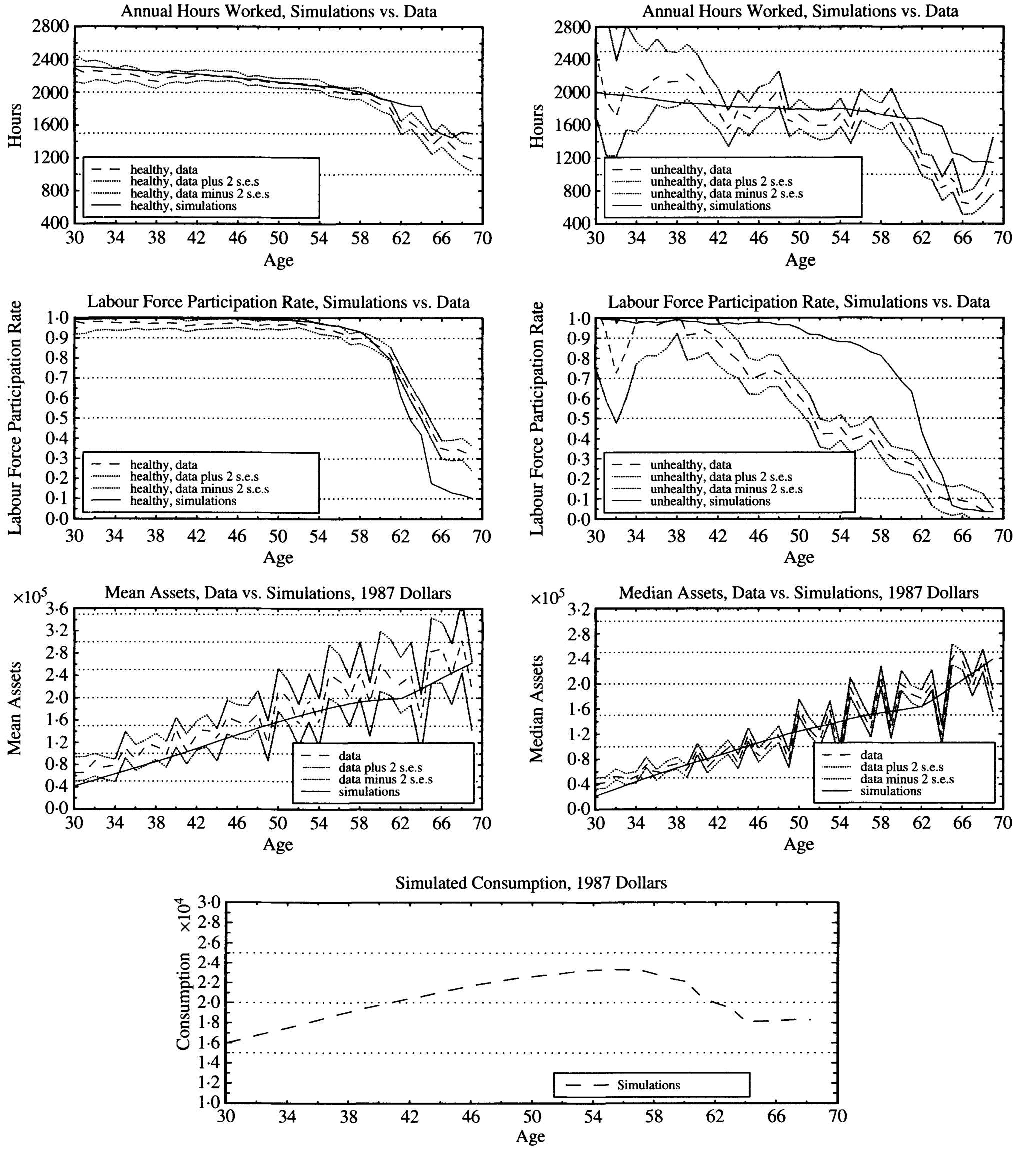

FIGURE 3

Simulated profiles vs. true profiles

There are two reasons for the small standard errors in Table 2. First, the standard error formulae rely on the assumption that the GMM criterion function is quadratic near the minimum of the function. This is true in the case of a linear model but may be a poor approximation in the case of a non-linear model.[37] For example, Gustman and Steinmeier (1986), Palumbo (1999),

37. To address this problem, I tried an alternative technique to obtain a measure of the precision of the estimates. I adjusted ν upwards by 5% and re-estimated the other parameters. A χ^2 test of the unrestricted model vs. the restricted model resulted in only a narrow rejection of the restricted model (the difference in the χ^2 statistics was 5·1, with a critical 5% value of 3·8). Also note the differences in overidentification test statistics between column 1 of Table 2 and column 1

Gourinchas and Parker (2002) and Cagetti (2003) all find extremely small standard errors using smaller data-sets than the one I use. Moreover, those studies have less variation in their data. For example, Gustman and Steinmeier (1986) only match labour supply paths, Cagetti (2003) only matches asset profiles, whereas this study matches both labour supply and asset profiles. Second, small standard errors may result from the assumption that the first stage parameter estimates of χ are measured with no error. Gourinchas and Parker (2002) suggest a method that allows one to incorporate the variance of the first stage parameter estimates into the second stage standard errors.

Note, however, that the profiles for both the decisions and beliefs are precisely estimated as shown by their smooth appearance. This arises from the extremely large sample size used in the analysis. Moreover, using data on labour force participation rates greatly increases the variation in the data. For example, Heckman and MaCurdy (1980) also obtain small standard errors in their fixed-effects Tobit specification for female labour supply. Therefore, it is unsurprising that standard errors are smaller than other analyses using PSID data on continuously employed male workers (*e.g.* MaCurdy, 1981).

5.5. The effects of selection and tied wage-hours offers on wages

Table 2 shows preference parameter estimates with and without controls for the effects of selection and tied wage-hours offers. This section describes how accounting for selection and tied wage-hours offers affects parameter estimates, as well as the interpretation of the estimates. Figure 1 shows that when using fixed-effects, the life cycle profile for wages declines rapidly after age 60. However, there are two reasons to suspect that the fixed-effects estimates do not represent the true productivity decline after age 60. First, as discussed in Section 3.3, I am likely overestimating wage growth because I am using data only on the wage growth of workers. In other words, the average wage decline for all individuals age 60+ is even sharper than the decline shown in Figure 1. Using the methodology described in Section 3.3, I find that true wages are 7% lower at age 62 and are 11% lower at age 65 than what is shown in Figure 1. Moreover, failure to account for selection leads to a 2% underestimate of the effect of health on wages.

In contrast, failure to account for tied wage-hours offers may lead to a downward bias in productivity growth after age 60. Gustman and Steinmeier (1986) and Aaronson and French (2004) present evidence that the drop in wages after age 60 may result from the drop in work-hours after age 60. The estimates presented herein assume that $\alpha = 0{\cdot}415$, which implies that part-time (1000 hours per year) workers are paid 25% less per hour than full-time (2000 hours per year) workers, resulting in a productivity profile displaying almost no decline after age 60.

Table 2 shows that the estimated fixed cost of work is very sensitive to whether the wage depends on hours worked. Both fixed costs of work and tied wage-hours offers are potential explanations for why hours worked by workers do not drop below 1000 hours. Tied wage-hours offers imply that individuals will not work a small number of hours per year, even if the fixed cost of work is small. Because of the low wages paid to part-time workers, a small number of hours worked results in very little labour income, making part-time work undesirable. However, if the wage does not depend on hours worked, a large fixed cost of work is necessary to explain why average hours worked does not decline below 1000 hours per year. When estimating preferences, the model fits the data almost as equally well with and without tied wage-hours offers; that is,

of Table 5. When $\theta_B = 0$, the test statistic rises from 856 to 968. This shows that the hypothesis of $\theta_B = 0$ can be rejected at almost any level. This tends to show that while the standard errors are being underestimated, the model is sharply identified.

the data cannot distinguish whether it is tied wage-hours offers or large fixed costs of work that result in hours worked not declining below 1000 hours per year.[38]

5.6. What causes the high job exit rates at age 62?

Figure 3 shows that the model is able to replicate the high job exit rates at age 62 that are seen in the data. There are several potential reasons for the high job exit rates at age 62: the rapid decline in pension accrual at age 62, actuarial unfairness of the Social Security system, and liquidity constraints. This section discusses the relative importance of these three reasons.

One important modelling decision is when to set the pension eligibility age. Because pension income is taxed and taxation is progressive, there is a jump in an individual's marginal tax rate if he continues to work and begins receiving pensions. This is an important labour supply disincentive. As discussed in Section 2.3, age 62 is the median normal retirement age for pensions. As a result, I assume that all individuals begin drawing pension benefits at age 62. The fact that many people would be pushed into higher tax brackets if they continued to work seems to cause about half the decline in labour supply at age 62. When I either make taxes proportional or change the pension eligibility age, about half the decline in labour supply at age 62 disappears. This result should be taken with a great deal of caution because pension eligibility should be a choice variable and there is a great deal of heterogeneity in the normal pension retirement age.[39] Nevertheless, the correlation between labour supply and the pension eligibility age shows the importance of considering the tax implications of pensions.

The model of pension accrual allows for discontinuous jumps in pension accrual at ages 61, 62, 63, 64, and 65. As Gustman and Steinmeier (1986, 1999) and Stock and Wise (1990) point out, discontinuities in pension accrual are potential explanations for the high job exit rates at ages 62 and 65. When I force pension accrual to be smooth and re-simulate the model, the age 62 downturn in labour force participation rates is 25% smaller. Therefore, together the tax and accrual aspects of pensions explain most of the decline in labour supply at age 62.

Next consider the actuarial unfairness of the Social Security system. Whether or not to apply for benefits at age 62 and face the Social Security earnings test depends largely upon the assumed rate of interest. Given a 4% pre-tax interest rate, Social Security benefit accrual is slightly negative (and is thus actuarially unfair) for individuals who face low marginal taxes at age 62.[40] However, even after using a 3% pre-tax rate of return, making Social Security benefit accrual positive for everyone between ages 62 and 65 results in a small change in labour force participation rates. In other words, actuarial unfairness explains only a small part of the decline in participation rates at age 62.

Liquidity constraints are another potential explanation for the high exit rates at age 62 (Kahn (1988), Rust and Phelan (1997)). Many individuals potentially wish to borrow against pension and Social Security benefits in order to finance retirement before age 62, but because Social Security wealth is illiquid, they are unable to do so. In order for liquidity constraints to affect consumption and labour supply decisions, future illiquid income must be high *vis-à-vis* current income (Deaton (1991), Carroll (1997)), so that consumption (and thus leisure) will rise when income eventually rises. In order to investigate the importance of liquidity constraints, I construct the following measure of the replacement rate: $\frac{pb_{62}+ss_{62}}{2000W_{62}}$. The numerator of this expression is

38. Aaronson and French (2004), who account for endogenous labour supply, show that tied wage-hours offers are necessary to fit the wage data.

39. When I set the pension eligibility age to either 55 or 65 and re-estimate the model, the model underpredicts job exit rates at age 62. Nevertheless, the model still matches the overall decline in job exit rates rather well. Moreover, preference parameter estimates are relatively unchanged.

40. Given the survivor probabilities I am using, the expected present value of benefits should be equal at ages 62 and 63 when the post-tax rate of return is 3·0%.

TABLE 3

The distribution of replacement rates and assets, age 62

	Quintile				
	0%–20%	20%–40%	40%–60%	60%–80%	80%–100%
By replacement rate quintile					
Mean replacement rate (simulated)	45·3%	61·9%	75·1%	92·0%	129·0%
Mean assets (simulated)	\$260,548	\$200,669	\$184,226	\$164,863	\$141,984
Mean hours decline (simulated)*	5·72%	10·3%	28·8%	53·4%	66·1%
By asset quintile					
Mean assets (simulated)	\$31,294	\$88,973	\$154,631	\$246,243	\$431,148
Mean hours decline (simulated)*	14·5%	13·7%	25·4%	37·0%	30·2%
Mean assets (data)	\$15,706	\$61,193	\$121,096	\$242,493	\$521,420

*Mean hours decline is measured as the change in the total number of work-hours between ages 61 and 62 for all individuals in that quintile.

pension and Social Security income after age 62 if the individual applies for benefits at age 62. The denominator of this expression is a measure of labour income when working 2000 hours per year. Therefore, this ratio measures the fraction of current labour income that Social Security and pensions replace. If this ratio is equal to one at age 62, an individual can leave the labour-market at age 62 with no assets and have no decline in consumption upon retirement. Therefore, if this ratio equals one, then an individual may optimally choose to have zero assets at retirement age.

The top panel of Table 3 shows quintiles of the replacement rate. Note that even though all individuals in the model face the same Social Security and pension benefit rules,[41] heterogeneity in wage and labour supply histories creates a large amount of heterogeneity in the replacement rate. Even though the median replacement rate is less than one, the average replacement rate for those in the top quintile of the replacement rate distribution is 129%. Those with high replacement rates tend to have relatively high wages when young (increasing the numerator of the expression) and relatively low wages when old (reducing the denominator of the expression). As a result, many simulated individuals choose to have close to zero wealth at age 62. Note that those with high replacement rates have low assets, on average. The evidence in Table 3 shows that having low asset levels near retirement is an optimal response to having a high replacement rate. The lower panel of Table 3 also shows quintiles of the asset distribution by both the model and by the PSID data. Note that even though I am not matching the distribution of assets, the distribution of assets implied by the model fits the data quite well.

Nevertheless, I find only a very small effect of liquidity constraints on labour supply. Table 3 shows that the mean hours decline at age 62 varies much more by replacement rate quintile than by asset quintile. Moreover, the bottom panel of Table 3 shows that those with low assets have smaller hours declines at age 62 than those with high assets. Therefore, it seems that the individuals who leave the labour force at age 62 are those with high replacement rates and high assets. In other words, those who face the largest jump in marginal tax rates at age 62 (recall that I assume that individuals first start drawing pension benefits at age 62, and pension income often increases the marginal tax rate) are the individuals who drop out of the labour force at age 62. Moreover, almost no individuals have assets equal to zero in either the data or the simulations at age 62. Thus, most individuals would be able to afford one year out of the labour-market at age 61.

41. Gustman and Steinmeier (1999) find a large amount of heterogeneity in pension accrual rates. Therefore, I am most likely underestimating the amount of heterogeneity in my illiquidity measure.

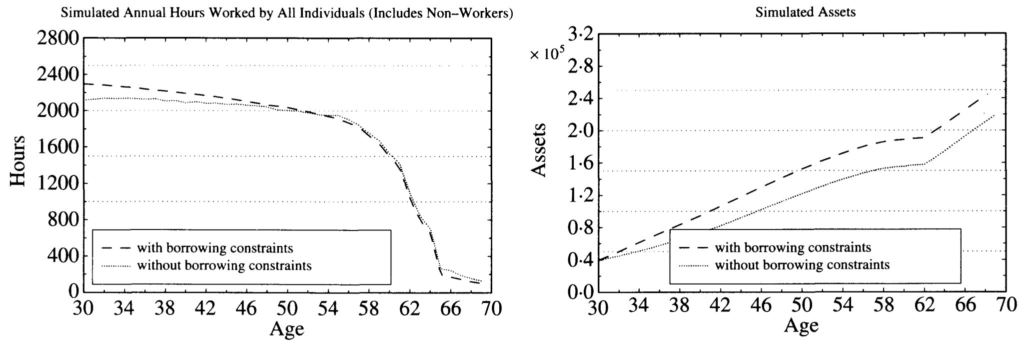

FIGURE 4

The importance of borrowing constraints

Another way of testing the importance of borrowing constraints is to allow people to borrow against their Social Security wealth (*i.e.* the present value of future Social Security income). Figure 4 shows average hours worked and asset levels both when the individual can borrow against future Social Security benefits and when the individual cannot borrow against future Social Security benefits.[42] Note that borrowing constraints affect asset growth a great deal. It also has some effect on hours worked, although most of the effect is when young.

In sum, liquidity constraints and the actuarial unfairness of the Social Security system explain little of the decline in work-hours at age 62. Taxes and pension accrual appear to be the driving factors.

5.7. *Consumption over the life cycle*

Although I do not match the life cycle consumption profile to the data, the model generates an implied life cycle consumption profile. The bottom panel of Figure 3 shows the geometric mean of consumption at each age. The consumption profile displays a pronounced hump shape. There are two reasons for the hump shape to the consumption profile. First, the combination of uncertainty and borrowing constraints implies that young consumers will save more than what they would have done in the absence of borrowing constraints. Because their wages and thus incomes are low, their consumption must be low. Second, non-separabilities between consumption and leisure imply that consumption tracks work-hours over the life cycle.[43]

This profile can be compared to the life cycle consumption profiles of Attanasio, Banks, Meghir and Weber (1999) or Gourinchas and Parker (2002). Both papers account explicitly for cohort effects and family size, attempt to account for time effects, and use Consumer Expenditure Survey data. Unsurprisingly, both papers yield similar life cycle profiles. In both studies, consumption grows approximately 15% between age 30 and age 40, peaks around age 40, then declines 30% by age 65.[44] In contrast, my consumption profile shows consumption rising 44%

42. However, in both situations the individual faces uncertainty. Low (2002) shows that uncertainty may have important effects on life cycle labour supply.

43. Identification of non-separabilities between consumption and leisure is tenuous in this model because consumption data are not used. Recall that leisure and consumption are substitutes if $\nu > 1$. Identification of this parameter has already been discussed. Therefore, the results presented herein should not be taken as strong evidence of non-separabilities between consumption and leisure. Instead, the results should be taken as evidence that this preference specification is consistent with the evidence from other studies.

44. One omission of my model is family size. Although I attempt to account for family size in the first stage estimating equations, my approach is not fully consistent with the formal model. Attanasio and Weber (1995) and

between ages 30 and 55, then declining 22% between 55 and 65. Therefore, my simulated profile peaks at a later age than profiles estimated using Consumer Expenditure Survey data. However, both Attanasio *et al.* (1999) or Gourinchas and Parker (2002) likely underpredict consumption of older individuals because both papers omit data on medical expenses and housing. Both goods tend to be consumed in greater quantities later in life.[45]

Nevertheless, my consumption profile provides some evidence that my baseline specification implies that households are "too patient". This may cause me to find no evidence of liquidity constraints even though they are important in the data.[46] Simulated individuals wish to save for higher consumption in the future, leading to asset levels at age 62 that are higher than those seen in the data. Section 6.1 investigates preference specifications with lower discount factors.

One interesting feature of the model is the sharp decline in consumption near retirement age. Below I show the relative importance of non-separabilities vs. shocks in explaining the consumption decline at retirement. To do this, I estimate the following model using simulated data:

$$\Delta \ln C_{it} = \beta s_t + \phi \Delta P_{it} + u_{it} \tag{23}$$

where u_{it} is a residual. Previous studies using OLS have found that ϕ is positive, indicating that consumption falls upon exit from the labour force. The OLS estimate may be a biased estimate of the consumption response to an anticipated exit from the labour force because a negative wage shock can reduce both labour supply and consumption. In other words, the OLS estimate may just be capturing the fact that the expected present value of future resources tends to decline when people exit the labour force. To check on the size of the bias, I estimate equation (23) using both OLS and IVs. Similar to the approach used by Banks, Blundell and Tanner (1998), I use age-average values of all variables in equation (23), and instrument for age-average values of ΔP_{it} using age-average values of ΔP_{it-2}. Using this approach I obtain an estimate of $\phi = 0{\cdot}34$, much larger than the Banks *et al.* (1998) estimate of 0·26. I also use dummy variables equal to one if the individual is older than 62 and 65 as instruments. This approach produces an estimate of $\phi = 0{\cdot}36$. The OLS estimate of ϕ is 0·37. Therefore, both OLS and instrumental variables produce similar results. This shows that OLS estimates are not severely biased and that shocks do not likely explain the consumption fall at retirement.[47]

6. EXPERIMENTS

Policy makers are interested in how Social Security generosity, the early and normal Social Security retirement age, and the Social Security earnings test affect labour supply. To answer these questions, I conduct four experiments. Table 4 gives accounting statistics for each of these experiments.

The top row of Table 4 displays results from simulations under the 1987 policy environment. The second row displays results where Social Security benefits are reduced by 20%. Reducing

Attanasio *et al.* (1999) emphasize the importance of family size on life cycle consumption. However, previous versions of Gourinchas and Parker (2002) show that accounting for family size only slightly changes the shape of the life cycle consumption profile. For example, their estimates imply that consumption rises 25% between 30 and 45 when not accounting for family size and 18% when accounting for family size. Attanasio *et al.* (1999) seem to find slightly larger effects of family size, but still find a hump shaped profile for life cycle consumption.

45. See Gokhale, Kotlikoff and Sablehaus (1996) for life cycle profiles that include housing and medical consumption. Properly accounting for medical expenses and housing would therefore produce a peak in life cycle consumption later than 40, although perhaps before 55.

46. Table 3 shows that I overpredict assets for those in the bottom quantile of the asset distribution.

47. Moreover, they are similar to what the model would predict assuming certainty and interior conditions. Assuming certainty and interior conditions, a 1% change in leisure brought about by a wage change will result in a $\frac{(1-\gamma)(\nu-1)}{\gamma(1-\nu)-1}$ per cent change in consumption. Given the parameter estimates in column 1 of Table 2, cutting work-hours from 1500 hours per year to 0 hours leads to consumption dropping 29%. This is again similar to the OLS and IV estimates.

TABLE 4

Policy experiments

	Years worked	Hours worked per year	PDV of labour income ($)	PDV of consumption ($)	Assets at age 62 ($)
With borrowing constraints					
1987 policies	32·60	2097	1781	1583	190
Reduce benefits	32·83	2099	1789	1569	200
Reduce benefits, reduce taxes	33·00	2115	1803	1586	203
Shift early retirement age to 63	32·62	2096	1781	1584	190
Eliminate earnings test, age 65+	33·62	2085	1799	1594	188
Without borrowing constraints					
1987 policies	32·39	2067	1764	1603	158
Reduce benefits	32·58	2063	1770	1587	168
Reduce benefits, reduce taxes	32·68	2078	1781	1602	170
Shift early retirement age to 63	32·39	2067	1764	1603	158
Eliminate earnings test, age 65+	33·46	2063	1784	1616	154

PDV stands for present discounted value.
Consumption, labour income, and assets are measured in thousands.

benefits has two labour supply effects, both of which should increase hours worked after age 62. First, the loss of Social Security benefits causes a loss of lifetime wealth. This results in individuals working more hours throughout their lives, as individuals consume less leisure given the loss of wealth. Second, reducing Social Security benefits also effectively reduces the Social Security earnings test and the high marginal tax rates of the earnings test. If an individual receives no Social Security benefits, there are no Social Security benefits to be reduced by the earnings test. Therefore, the substitution effect associated with a benefit cut causes individuals to work more hours when eligible for Social Security benefits and fewer hours at younger ages.

The second row of Table 4 shows that reducing benefits causes individuals to work more hours throughout their lives and thus increase their assets in order to offset reduced benefits. To understand the magnitude of these effects, note that the present value of Social Security benefits at age 62 is equal to about $132,000, on average. Cutting benefits by 20% reduces the present value of Social Security wealth by $26,000. Due to both reduced consumption and increased work-hours when younger than 62, age 62 asset levels are $9800 greater when benefits are reduced. About $\frac{2}{3}$ of this effect is through reduced consumption; the other $\frac{1}{3}$ is from increased labour supply. This highlights the importance of forward-looking behaviour when considering the effects of changing the Social Security rules. Nevertheless, most of the effects are seen after age 62. Increased years in the labour-market after age 62 replace $5500 of the lost income. The reason for this is that most individuals are still working at age 62, and most of the life cycle variability in hours is at the participation margin. Therefore, the substitutability of labour supply is high after age 62. Reduced consumption when old and reduced bequests account for the remaining lost benefits. These labour supply effects are similar to those of Burtless (1986) and Krueger and Pischke (1992) and are fairly average for the literature. A more extreme experiment of eliminating Social Security benefits results in an increase of average years in the labour force between age 30 and 70 to 33·71 years.

In order to check whether substitution effects or wealth effects drive my results, I reduce the present discounted value of taxes paid over the life cycle by an amount equal to the present discounted value of reduced benefits.[48] Note that this is roughly similar to eliminating wealth

48. This was done assuming that population growth is $g_p = 1\%$, each birthyear cohort has annual income that is $g_w = 2\%$ above the previous birthyear cohort, and mortality rates are the same for all cohorts. In this case the net cost of

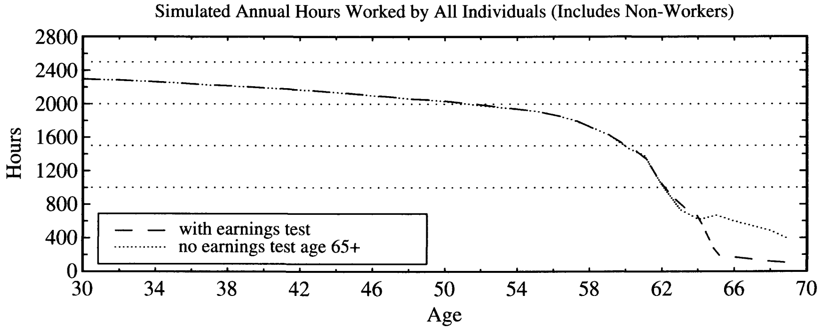

FIGURE 5
The effect of removing the earnings test, age 65+

effects. Nevertheless, the substitution effects of reducing benefits remains. Results are displayed in the third row of Table 4. Upon reducing taxes, hours worked after age 62 are still very high. This experiment highlights the importance of the substitution effect generated by the Social Security work disincentives for individuals age 62 and older.

One potential reform to the Social Security system is to shift the early retirement age from 62 to 63. Recall that benefit recomputation formulae almost fully replace benefits lost through the earnings test at age 62. Therefore, if borrowing constraints do not bind, there should be little if any work disincentive imposed by Social Security at age 62 and thus there should be little if any effect of shifting the Social Security early retirement age to 63. Recall that borrowing constraints bind for very few individuals at age 62. As a result, the fourth row of Table 4 shows that any effect of this policy would be minor. Simulations from the model indicate that shifting the early Social Security retirement age to 63 would leave years in the labour force unchanged.

Finally, I eliminate the Social Security earnings test for individuals aged 65 and older. This has large effects. As shown in the fifth row of Table 4 and also Figure 5, hours worked after age 65 jumps. Years in the labour force rises from 32·60 to 33·62 although average hours worked by workers is largely unchanged. This increase in work-hours is completely a substitution effect, given that eliminating the earnings test will increase lifetime wealth. The wealth effect from increased post-tax wages will lead individuals to consume more of everything, including leisure. Therefore, eliminating the wealth effects from this experiment would lead to an even greater labour supply response.

This final experiment provides the model with a strong out of sample test. The earnings test was in fact abolished for individuals older than 64 in 2000. Therefore, the model predicts that labour force participation rates for individuals should rise sharply over the coming years.[49]

The results in Table 4 highlight the importance of considering labour force participation when conducting policy experiments. Most models (for example, Auerbach and Kotlikoff, 1987)

the Social Security system is

$$\sum_{t=30}^{T}\sum_{n=1}^{N}(1+g_p+g_w)^{(T-t)}S_{30,t}(ss_{n,t}-[(1-\tau_{ss})\times\min\{W_{n,t}H_{n,t},43{,}000\}]), \tag{24}$$

where N is the number of simulations and τ_{ss} is the Social Security tax. The net cost of the Social Security system is zero when $\tau_{ss} = 0{\cdot}065$ and $ss_{n,t}$ comes from the 1987 benefit formulae. This is greater than the employee OASDI tax of 0·052 but less than the employer and employee contribution of 0·104. The net cost is also zero when $\tau_{ss} = 0$ and $ss_{n,t} = 0$. Therefore, I reduce taxes by $0{\cdot}065 \times 20\% = 0{\cdot}013$ up to the OASDI maximum of $43,000 when reducing benefits 20%.

49. Of course, other incentives have been changing over time (Anderson, Gustman and Steinmeier, 1999), so the prediction is somewhat ambiguous. Over the past 15 years there have been important changes to the earnings test for individuals younger than 65 and to benefit recomputation formulae for individuals aged 65 and older.

TABLE 5

Preference parameter estimates

Parameter and definition	Specification (1)	(2)	(3)	(4)
γ Consumption weight	0·589 (0·004)	0·556 (0·005)	0·539 (0·001)	
ν Coefficient of relative risk aversion, utility	5·68 (0·07)	9·98 (0·16)	6·36 (0·08)	
β Time discount factor	1·04 (0·002)	0·95	0·95	0·987 (0·001)
L Leisure endowment	5159·0 (31)	5073 (44)	3937 (27)	5280
ϕ Hours of leisure lost, bad health	559 (8)	429 (7)	175 (12)	153 (4)
θ_P Fixed cost of work, in hours	1378 (15)	772 (9)	273 (3)	553 (7)
θ_B Bequest weight	0	0	0	5·62 (0·26)
χ^2 statistic: (193 degrees of freedom)	968	1093	1158	1107
$\epsilon_{h,W}(40)$ Labour supply elasticity, age 40	0·35	0·41	0·49	0·06
$\epsilon_{h,W}(60)$ Labour supply elasticity, age 60	2·17	0·99	0·50	0·99
Reservation hours level, age 62	885	1226	1059	900
Coefficient of relative risk aversion	3·76	6·00	2·08	0·566

Standard errors in parentheses

Specifications described below:
(1) $\theta_B = 0$
(2) $\theta_B = 0, \beta = 0{\cdot}95$
(3) $\theta_B = 0, \beta = 0{\cdot}95$, intertemporal elasticity of substitution for consumption $= \frac{-1}{\gamma(1-\nu)-1} = 0{\cdot}48$
(4) Separable preferences: $\gamma_C = 0{\cdot}566$ (0·0003), $\gamma_H = 9{\cdot}82$ (0·01), $\phi_H = 3{\cdot}94 \times 10^{32} (2{\cdot}28 \times 10^{30})$, $L = 5280$ by assumption, γ and ν not in specification

focus on hours worked by workers and ignore the labour force participation decision. This model suggests, however, that the dominant margin of labour supply substitutability for men is at the labour force participation decision.

The bottom rows of Table 4 repeat the top rows, but assume that individuals can borrow against future Social Security benefits. Note that relaxing borrowing constraints does have important effects on labour supply and asset accumulation. Asset levels are much lower at age 62 when borrowing constraints are relaxed. Hours worked per year are also lower. However, note from Figure 4 that this reflects changes in labour supply early in life, but not near retirement age. Moreover, note that the effects of changing the Social Security rules on labour supply is similar to the effects when borrowing constraints are enforced. Elimination of the earnings test for those older than 65 has a very large effect on life cycle labour supply, while shifting the early retirement age to 63 has a very small effect. Therefore, the presence of borrowing constraints does not affect the predicted response of labour supply to changes in the Social Security rules.

6.1. *Sensitivity of results to changes in preference parameters and specification*

In this section I present evidence on the sensitivity of results to changes in the preference parameters. Of the estimated parameters, the values of β, θ_B, and ν are potentially the most controversial, as is the source of identification of these parameters. Moreover, it seems worthwhile to test the robustness of the results to alternative preference specifications. In this section, I evaluate whether the results in Table 4 are sensitive to changes in preference parameters and the utility function.[50] Preference parameter estimates are shown in Table 5.

One cause for concern when constructing asset profiles is that the year effects are not well proxied by the unemployment rate. Given that the asset data are from 1984 to 1994, during which

50. However, the importance of unobserved heterogeneity in preferences and pension accrual is not considered.

there was a rapid run-up in the stock market, I may be overstating asset growth. Therefore, I may be overestimating θ_B and β.

In order to understand the importance of this problem, I reduced asset growth 1% during each year of the sample period (*i.e.* I took a 1% growth rate per year out of the data, allowing for a pure time effect) and re-estimated preference parameters. French (2003) shows that this technique likely understates asset growth over the life cycle relative to what we would have anticipated in the absence of a run-up in the stock market. When using this asset profile, the estimate of θ_B falls from 1·69 to 0·85. None of the other parameters changes noticeably. Moreover, the new estimates have only a tiny effect on labour supply and savings responses to changes in the Social Security rules.

However, there are other reasons to suspect that I may be overestimating θ_B and β. The parameter θ_B is identified largely from the shape of the asset profile, but only for individuals younger than 70. However, Hurd (1989) finds significant declines in assets near the end of the life cycle whereas the simulated profiles presented herein do not fall for older individuals. Moreover, I omit medical expense uncertainty. Palumbo (1999) shows that uncertain medical expenses can partly explain why the elderly run down their wealth slowly. In order to address these concerns I tried a more extreme set of experiments. In column 1 of Table 5, I set $\theta_B = 0$. Note that β rises to 1·04 when $\theta_B = 0$. This gives some evidence that a high value of θ_B and a high value of β are alternative explanations for why assets are high near age 70.

A value of $\beta = 1{\cdot}04$ is much higher than most estimates in the literature. In order to assess the sensitivity of my results, I set $\beta = 0{\cdot}95$ and $\theta_B = 0$ in column 2 of Table 5.[51] Note that when $\theta_B = 0$ and $\beta = 0{\cdot}95$, the value of ν rises: because they are less patient, consumers need to be more risk averse to generate the observed asset profile.[52] This high value of ν also generates a low intertemporal elasticity of substitution for labour supply and consumption.[53]

Because the intertemporal elasticity of substitution for consumption in column 2 is lower than most estimates,[54] I also consider an intertemporal elasticity of substitution for consumption of 0·48, estimated by Attanasio *et al.* (1999). The results are in column 3 of Table 5. The estimates in column 3 produce the lowest asset levels and are thus the most likely to produce a large effect of liquidity constraints on life cycle labour supply. Average assets are $10,000 at age 62. In order to assess the importance of liquidity constraints, Table 6 shows the same experiments as in Table 4, but with the preferences in column 3 of Table 5. The largest change between the results in Tables 4 and 6 is the effect of changing generosity. Reducing benefits now has an even larger effect on life cycle labour supply. Because asset levels are low at all ages, reducing benefits has only small effects on savings before age 62 and bequests. Most of the response to the benefit cut is in reduced consumption and leisure after age 62. Although lower patience factors affect the labour supply response to Social Security generosity, they do not affect the labour supply response to shifting the early retirement age to 63.

51. As I pointed out earlier, β is identified partly by the life cycle labour supply profiles. The estimated profiles indicate that individuals cut their work hours (or equivalently, increase their leisure consumption) between ages 50 and 60, even though the wage and pension incentives (and thus the price of leisure) are at their greatest near age 60. These facts can only be reconciled by a high value of β. However, one could argue that individuals reduce work-hours at these ages because of declining health, and that this decline is not fully captured by the one simple health measure that I use. Individuals may be impatient, but the disutility of work rises sharply at these ages.

52. Figure 4 of Cagetti (2003) shows the relationship between risk aversion and impatience more explicitly. Note, however, that ν is also partly identified by the labour supply profiles.

53. Assuming certainty and interior conditions, the consumption Euler equation is $\Delta \ln C_t = \frac{-1}{\gamma(1-\nu)-1} \ln(\beta(1+r)) + \frac{(1-\gamma)(\nu-1)}{\gamma(1-\nu)-1} \Delta \ln(L - H_t - \theta_P P_t - \phi I\{M = \text{bad}\})$.

54. Moreover, a wide range of values of ν all seem to give relatively similar criterion functions if other parameters are re-estimated. From a statistical standpoint, one can easily distinguish between different values of ν. Nevertheless, the profiles they generate look relatively similar.

TABLE 6

Policy experiments, $\theta_B = 0$, $\beta = 0{\cdot}95$, $\frac{-1}{\gamma(1-\nu)-1} = 0{\cdot}48$

	Years worked	Hours worked per year	PDV of labour income (\$)	PDV of consumption (\$)	Assets at age 62 (\$)
With borrowing constraints					
Current policies	36·77	2003	1788	1840	10
Reduce benefits 20%	37·42	2000	1805	1828	13
Reduce benefits 20%, reduce taxes	37·66	2011	1819	1840	14
Shift early retirement age to 63	36·77	2003	1788	1840	11
Eliminate earnings test, age 65+	37·91	2005	1811	1858	8

PDV is present discounted value.
Consumption, labour income, and assets are measured in thousands.

One final specification test is to change the utility function so that it is separable in consumption and leisure. Consider the following utility and bequest functions:

$$U(C_t, H_t, M_t) = \frac{1}{1-\gamma_C} C_t^{1-\gamma_C} + \frac{\phi_H}{1-\gamma_H}(L - H_t - \theta_P P_t - \phi I\{M = \text{bad}\})^{1-\gamma_H}, \quad (25)$$

$$b(A_t) = \theta_B \frac{(A_t + K)^{1-\gamma_C}}{1-\gamma_C}. \quad (26)$$

The results from this specification are in column 4 of Table 5. This utility function does not fit the data as well as the non-separable preference specification, although there are no striking differences between this preference specification and the non-separable one. Nevertheless, when repeating the experiments in Table 4, shifting forward the early retirement age has almost no effect on the lifetime labour supply, whereas eliminating the Social Security earnings test after age 65 increases years in the labour force by 1·4 years.

7. CONCLUSION

In this paper I present estimates from the first estimable dynamic structural model of labour supply, retirement, and savings behaviour where assets must be non-negative in all periods. When augmented to include uncertainty over future wages and health status, the model fits the life cycle profile of assets rather well. It also does a good job of fitting the life cycle profiles of hours worked and labour force participation rates.

This allows me to assess how the Social Security system affects life cycle labour supply. Of central importance is whether Social Security affects labour supply because (i) Social Security wealth is illiquid until age 62 and/or (ii) because of the taxation and actuarial unfairness of the system. I find that allowing individuals to borrow against future Social Security benefits would reduce work-hours when younger than 40. However, the fact that benefits are illiquid until 62 cannot explain the high job exit rates at 62 or 65. Instead, it seems that the taxation and actuarial unfairness of pensions and Social Security explains the sharp decline in labour supply at these ages.

The value of this model lies in its ability to predict how labour supply and retirement patterns of individuals might change in response to changes in the Social Security rules. Simulations suggest that a 20% drop in Social Security benefits results in an increase in labour supply throughout the life cycle. However, the effect is rather small; individuals would spend an additional three months in the labour force. In contrast, simulations suggest that the elimination of the Social Security earnings test for those older than 65 will cause individuals to delay exit from the labour force by one year, showing the important work disincentives of the earnings test.

APPENDIX: MOMENT CONDITIONS

In this Appendix I describe the GMM minimization procedure where I account for the three data problems discussed in the text. The first data problem is that I wish to match profiles that are uncontaminated by cohort and family size effects. The second problem is that person-specific unobservables are potentially correlated with both health and preferences for work.

For concreteness, consider the moment condition for hours worked for individuals in good health. I wish to set the following moment condition to zero:

$$E[[\ln H_{it,M=\text{good}} \mid \text{birthyear} = 1940, M = \text{good}, \text{famsize} = 3] - \ln \tilde{H}_{t,M=\text{good}}] = 0, \tag{A.1}$$

where $\ln \tilde{H}_{M,t}$ is the simulated geometric mean of log hours worked. In order to generate this moment condition, I use parameter estimates from equation (18) and make three modifications to hours worked, shown in equation (A.2):

$$\begin{aligned}\ln H_{it,M=\text{good}} = f_i + E[f_i \mid \text{birthyear} = 1940, \text{prob}(M = \text{good}) = \text{prob}(M = \text{good} \mid t = 50), t = 50] \\ - E[f_i \mid \text{birthyear}_i, \text{prob}(M_{it} = \text{good}), t] + \Pi_g \text{age}_{it} + \Pi_f(\text{famsize} = 3) + u_{it}.\end{aligned} \tag{A.2}$$

Note the three modifications to the data. First, there is no probability of being in good or bad health. Instead, individuals are in either good or bad health for certain. Second, it is not the size of the family that is used but a family size of three. In this way life cycle family size effects will not contaminate profiles. Third, I adjust the person-specific effects f_i. I adjust the person-specific effect so that everyone has the same cohort effect, set to birthyear $= 1940$, which means profiles will be uncontaminated by cohort effects. The second aspect of the person-specific effect that I adjust for is the possible correlation between the person-specific effect and the health status of the individual. Below I point out that the adjusted person-specific effects are uncorrelated with health status by construction.

Next I address the third data problem of having an unbalanced panel and the problem of not knowing an individual's health status with certainty. If there are I separate individuals in the data there will be a total of I possible contributions to both the healthy and unhealthy moment conditions for hours at age t. However, not all individuals are observed working for all possible time periods. Assume instead that there are $I_t \leq I$ individuals observed working at age t. The idea is to treat a moment contribution as equal to zero if it is missing.

This means that the moment condition for individuals of age t and health state $M = \text{good}$ is generated by

$$\frac{1}{I}\sum_{i=1}^{I_t}\{\ln H_{it,M=\text{good}} - \ln \tilde{H}_{t,M=\text{good}}\} \times \text{prob}(M_{it} = \text{good}), \tag{A.3}$$

where $\ln H_{it,M=\text{good}}$ is adjusted work-hours described in equation (A.2). The relative weight of this moment condition rises as I_t, the number of observed workers rises and as the probability that these workers are healthy rises. Note that $\text{prob}(M_{it} = \text{good})$, which determines selection into the moment condition, might be correlated with the person-specific fixed effect f_i but will not be correlated with its adjusted value $f_i + E[f_i \mid \text{birthyear} = 1930, \text{prob}(M = \text{good}) = \text{prob}(M = \text{good} \mid t = 50), t = 50] - E[f_i \mid \text{birthyear}_i, \text{prob}(M_{it} = \text{good}), t]$ by construction.

Define the vector of the $6T$ moment conditions (the distance between the simulated profiles and estimated profiles for assets, hours, and participation) as $\tilde{g}(\theta; \chi)$. Assuming W_T is an optimal weighting matrix, the minimized GMM criterion function

$$\min_{\theta\in\Theta} \frac{I}{1+\tau}\tilde{g}(\theta; \chi)' W_T \tilde{g}(\theta; \chi) \tag{A.4}$$

is distributed asymptotically as chi-squared with $6T - 7$ degrees of freedom if the model is correctly specified. τ is the ratio of the number of observations to the number of simulated observations. My estimate of W_T is the inverse of the $6T \times 6T$ variance covariance matrix of the (adjusted) data. That is, W_T^{-1} has a typical element along the diagonal of a variance $\frac{1}{I}\sum_{i=1}^{I_t}\{[\ln H_{it,M=\text{good}} - E[\ln H_{it,M=\text{good}}]] \times \text{prob}(M_{it} = \text{good})\}^2$ and a typical element of a covariance on the off-diagonal. When computing the chi-square statistic and the standard errors, the estimated value of $E[\ln H_{it,M=\text{good}}]$ is replaced with its simulated counterpart.

Under the regularity conditions stated in Pakes and Pollard (1989) and Duffie and Singleton (1993), the MSM estimator $\hat{\theta}$ is both consistent and asymptotically normally distributed. Denoting θ_0 and $\hat{\theta}$ as the true and estimated parameter vector, $\hat{\theta}$ converges in distribution to

$$\sqrt{I}(\hat{\theta} - \theta_0) \rightsquigarrow N(0, V), \tag{A.5}$$

where V is estimated by $\hat{V} = (1+\tau)(\hat{D}'\hat{W}\hat{D})^{-1}$ and $\hat{D} = \frac{\partial \tilde{g}}{\partial \theta}\mid_{\theta=\hat{\theta}}$.

Acknowledgements. I thank Joe Altonji, John Jones, John Kennan, Rody Manuelli, Jonathan Parker, Dan Sullivan, and Jim Walker for detailed comments and encouragement. Jonathan Parker also provided me with data and computer code. I also thank Dan Aaronson, Peter Arcidiacono, Hugo Benitez-Silva, Marco Cagetti, Glen Cain, Chris Carroll, Meredith Crowley, Morris Davis, Nelson Graff, Ed Green, Alan Gustman, Lars Hansen, Yuichi Kitamura, Grigory

Kosenok, Spencer Krane, Tecla Loup, Derek Neal, Victor Rios-Rull, Marc Rysman, Karl Scholz, three very helpful referees, the editors (Orazio Attanasio and Bernard Salanie), and numerous seminar participants. Jonathan Hao, Kate Godwin, and especially Tina Lam provided excellent research assistance. Financial support provided by the National Institute on Mental Health. The views of the author do not necessarily reflect those of the Federal Reserve System.

REFERENCES

AARONSON, D. and FRENCH, E. (2004), "The Effect of Part-Time Work on Wages: Evidence from the Social Security Rules", *Journal of Labor Economics*, **22** (2), 329.

ABOWD, J. and CARD, D. (1989), "On the Covariance Structure of Earnings and Hours Changes", *Econometrica*, **57** (2), 411–445.

ALTONJI, J. (1986), "Intertemporal Substitution in Labor Supply: Evidence from Microdata", *Journal of Political Economy*, **94** (3), S176–S215.

ANDERSON, P., GUSTMAN, A. and STEINMEIER, T. (1999), "Trends in Male Labor Force Participation and Retirement: Some Evidence on the Role of Pensions and Social Security in the 1970s and 1980s", *Journal of Labor Economics*, **17** (4), 757–783.

ATTANASIO, O., BANKS, J., MEGHIR, C. and WEBER, G. (1999), "Humps and Bumps in Lifetime Consumption", *Journal of Business and Economic Statistics*, **17** (1), 22–35.

ATTANASIO, O. and WEBER, G. (1995), "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey", *Journal of Political Economy*, **103** (6), 1121–1157.

AUERBACH, A. and KOTLIKOFF, L. (1987) *Dynamic Fiscal Policy* (Cambridge: Cambridge University Press).

BANKS, J., BLUNDELL, R. and TANNER, S. (1998), "Is There a Retirement Savings Puzzle?", *American Economic Review*, **88** (4), 769–788.

BENITEZ-SILVA, H. (2000), "A Dynamic Model of Labor Supply, Consumption/Saving, and Annuity Decisions Under Uncertainty" (Manuscript).

BLAU, D. (1994), "Labor Force Dynamics of Older Men", *Econometrica*, **62** (1), 117–156.

BLUNDELL, R., REED, H. and STOKER, T. (2003), "Interpreting Aggregate Wage Growth: the Role of Labor Market Participation", *American Economic Review*, **93** (4), 1114–1131.

BROWNING, M., DEATON, A. and IRISH, M. (1985), "A Profitable Approach to Labor Supply and Commodity Demands Over the Life-Cycle", *Econometrica*, **53** (3), 503–543.

BURTLESS, G. (1986), "Social Security, Unanticipated Benefit Increases, and the Timing of Retirement", *Review of Economic Studies*, **53** (5), 781–805.

CAGETTI, M. (2003), "Wealth Accumulation Over the Life Cycle and Precautionary Savings", *Journal of Business and Economic Statistics*, **21** (3), 339–353.

CARD, D. (1994), "Intertemporal Labor Supply: An Assessment", in C. Sims (ed.) *Advances in Econometrics: Sixth World Congress*, vol. 2, (Cambridge: Cambridge University Press).

CARROLL, C. (1997), "Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis", *Quarterly Journal of Economics*, **102** (1), 1–55.

COGAN, J. (1981), "Fixed Costs and Labor Supply", *Econometrica*, **49** (4), 945–963.

CURRIE, J. and MADRIAN, B. C. (1999), "Health, Health Insurance and the Labor Market", in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, vol. 3C, (Amsterdam: North-Holland).

DEATON, A. (1991), "Saving and Liquidity Constraints", *Econometrica*, **59** (4), 1221–1248.

DE NARDI, C. M. (2004), "Wealth Inequality and Intergenerational Links", *Review of Economic Studies*, **971** (3), 743–762.

DOMEIJ, D. and FLODEN, M. (2002), "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased" (Stockholm University).

DUFFIE, D. and SINGLETON, K. (1993), "Simulated Moments Estimation of Markov Models of Asset Prices", *Econometrica*, **61** (4), 929–952.

FRENCH, E. (2001), "How Severe is Measurement Error in Health Status: Evidence from the PSID" (Mimeo).

FRENCH, E. (2003), "The Effects of Health, Wealth and Wages on Labor Supply and Retirement Behavior" (Chicago Fed Working Paper 2000-12; available at `http://www.chicagofed.org/`).

FRENCH, E. (2004), "The Labor Supply Response to (Mismeasured but) Predictable Wage Changes", *Review of Economics and Statistics*, **86** (2), 602–613.

FRENCH, E. and JONES, J. (2004), "The Effects of Health Insurance and Self-Insurance on Retirement Behavior" (Mimeo).

GHEZ, G. and BECKER, G. (1975), "The Allocation of Time and Goods Over the Life Cycle" (NBER).

GOKHALE, J., KOTLIKOFF, L. and SABLEHAUS, J. (1996), "Understanding the Postwar Decline in U.S. Saving: A Cohort Analysis", *Brookings Papers on Economic Activity*, 315–390.

GOURINCHAS, P. and PARKER, J. (2002), "Consumption Over the Life Cycle", *Econometrica*, **70** (1), 47–89.

GROSSMAN, M. (1972), "On the Concept of Health Capital and the Demand for Health", *Journal of Political Economy*, **80**, 223–255.

GUSTMAN, A., MITCHELL, O., SAMWICK, A. and STEINMEIER, T. (1998), "Evaluating Pension Entitlements" (Mimeo).

GUSTMAN, A. and STEINMEIER, T. (1986), "A Structural Retirement Model", *Econometrica*, **54** (3), 555–584.

GUSTMAN, A. and STEINMEIER, T. (1999), "Effects of Pensions on Savings: Analysis with Data from the Health and Retirement Study", *Carnegie–Rochester Conference Series on Public Policy*, 271–324.

HECKMAN, J. (1976), "A Life-Cycle Model of Earnings, Learning, and Consumption", *Journal of Political Economy*, **84** (4), S11–S44.

HECKMAN, J. and MACURDY, T. (1980), "A Life-Cycle Model of Female Labour Supply", *Review of Economic Studies*, **47**, 47–74.

HURD, M. (1989), "Mortality and Bequests", *Econometrica*, **57** (4), 779–813.

IPPOLITO, R. (1997) *Pension Plans and Employee Performance* (Chicago: University of Chicago Press).

KAHN, J. (1988), "Social Security, Liquidity, and Early Retirement", *Journal of Public Economics*, **35**, 97–117.

KRUEGER, A. and PISCHKE, J. (1992), "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation", *Journal of Labor Economics*, **10**, 412–437.

LOW, H. (2002), "Self-Insurance and Unemployment Benefit in a Life-Cycle Model of Labour Supply and Savings" (IFS Working Paper).

MACURDY, T. (1981), "An Empirical Model of Labor Supply in a Life-Cycle Setting", *Journal of Political Economy*, **89** (6), 1059–1085.

PAKES, A. and POLLARD, D. (1989), "Simulation and the Asymptotics of Optimization Estimators", *Econometrica*, **57**, 1027–1057.

PALUMBO, M. (1999), "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle", *Review of Economic Studies*, **66** (2), 395–421.

RUST, J., BUCHINSKY, M. and BENITEZ-SILVA, H. (2003), "An Empirical Model of Social Insurance at the End of the Life Cycle" (Mimeo).

RUST, J. and PHELAN, C. (1997), "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets", *Econometrica*, **65**, 781–831.

STOCK, J. and WISE, D. (1990), "An Option Value Model of Retirement", *Econometrica*, **58**, 1151–1180.

UNITED STATES SOCIAL SECURITY ADMINISTRATION, *Social Security Bulletin: Annual Statistical Supplement* (United States Government Printing Office) (selected years).

VAN DER KLAAUW, W. and WOLPIN, K. (2003), "Social Security, Pensions and the Savings and Retirement Behavior of Households" (Mimeo).