Estimation of a Life-Cycle Model with Human Capital, Labor Supply and Retirement

Xiaodong Fan
University of New South Wales
Ananth Seshadri
University of Wisconsin-Madison
Christopher Taber
University of Wisconsin-Madison

March 11, 2015

Abstract

We develop and estimate a life-cycle model in which individuals make decisions about consumption, human capital investment, and labor supply. Retirement arises endogenously as part of the labor supply decision. The model allows for both an endogenous wage process through human capital investment (which is typically assumed exogenous in the retirement literature) and an endogenous retirement decision (which is typically assumed exogenous in the human capital literature). We estimate the model using the Method of Simulated Moments to match the life-cycle profiles of wages and hours from the SIPP data. The model replicates the main features of the data—in particular the large increase in wages and small increase in labor supply at the beginning of the life-cycle as well as the small decrease in wages but large decrease in labor supply at the end of the life cycle. We also estimate versions of the model in which human capital is completely exogenous and in which human capital is exogenous conditional on work (learning-by-doing). The endogenous human capital model fits the data the best.

KEYWORDS: human capital, Ben-Porath, labor supply, retirement

JEL Classification: J22, J24, J26

1 Introduction

The Ben-Porath (1967) model of life-cycle human capital production and the life-cycle labor supply model are two of the most important models in labor economics. The former is the dominant framework used to rationalize wage growth over the life-cycle; the latter has been used to study hours worked over the life-cycle, including retirement. Quite surprisingly, aside from the seminal work in Heckman (1976, 1975), there has been little effort integrating these two important paradigms. This paper attempts to fill this void by estimating a life-cycle model in which workers choose human capital and labor supply jointly. Perhaps the most important aspect of our model is that we do not treat retirement as a separate decision. It occurs endogenously as part of the optimal life-cycle labor supply decision.

The retirement literature typically takes the wage process as given and estimates the incidence of retirement. For individuals who work, raw wages in the cross-section fall substantially before retirement. They decline by over 25% between ages 55 and 65. In the retirement literature, this trend is critical in explaining retirement behavior. By contrast, life-cycle human capital models take the retirement date as given, but model the formation of the wage process. While most work to date on the life-cycle human capital model aims to explain wage growth early in the life-cycle, there has been little work studying the interaction between human capital and labor supply at the end of the working life. We estimate a model wherein the wage, labor supply and retirement choices are rationalized in a unified setting. After endogenizing both labor supply and human capital, this model is rich enough to explain the life-cycle pattern both wages and labor supply, with a focus on wage patterns and retirement at the end of working life.

Specifically, we develop and estimate a Ben-Porath type human capital model in which workers make consumption, human capital investment, and labor supply decisions. We estimate the model using the Method of Simulated Moments (MSM), matching the wage and hours profiles of Male high school graduates from the Survey of Income and Program Participation (SIPP). With a parsimonious life-cycle model in which none of the parameters explicitly depend upon age or experience, we are able to replicate the main features of the data. In particular we match the large increase in wages and very small increase in labor supply at the beginning of the life-cycle as well as the small decrease in wages but very large decrease in labor supply at the end of the life-cycle.

The key to our ability to fit both ends of the life-cycle is human capital depreciation. In a simple model without human capital depreciation, there is no a priori reason for workers to concentrate their leisure towards the end of the life-cycle. However, this is no longer the case with human capital depreciation. When workers take time off in the middle of their career, their human capital depreciates and they earn less when they return to the labor market. On the other hand, if this period of nonworking occurs at the end of the career, the fall in human is much less a concern. As a result of the shorter time horizon, an older worker may choose not to re-enter at a lower wage so they continue to stay out of the labor market. We show that if we restrict our framework to exogenous human capital accumulation over the life-cycle , the model does not fit both the end and beginning of the lifecycle. When tastes for leisure do not vary over the life-cycle, the standard model cannot simultaneously reconcile the small increase in labor supply and large increase in wages at the beginning of the life-cycle and the small decrease in wages and large decrease in labor supply at the end. Of course if one exogenously allowed both wages and labor supply to depend upon age in a completely flexible way one could easily fit the joint pattern. Moreover, it is not clear that this model would have any testable implications so we cannot reject it. The goal of this paper is to try to fit the profiles without resorting to arbitrary age varying taste preferences and exogenous wage variation.

An interesting aspect of our model is that even though the preference for leisure does not vary systematically over the life-cycle, we do find that measured "labor supply elasticities" do vary over the life-cycle. In our dynamic model, the shadow cost of not working is much higher early in the life-cycle (as pointed out by e.g. Imai and Keane, 2004) and it is also lower for older workers as opposed to peak earners. We find that early in the life-cycle the measured labor supply elasticity is low for younger workers, around 0.2. However, around the standard retirement age, workers are more sensitive to wage fluctuations with elasticities between 0.6 and 1.0.

While our baseline model does not incorporate health, we estimate a specification that allows the taste for leisure to depend on health and for this effect to increase with age. Surprisingly, such an "enhanced" model does not significantly improve the fit of the life-cycle patterns of wage and labor supply of the SIPP data. We also show that even within this model that allows a direct and flexible effect of health on labor supply, health plays a relatively minor role in the decline in labor supply late in life.

We also use the estimated model to simulate the impacts of various Social Security policy changes. Much serious work has been developed to quantitatively estimate the economic consequences of an aging population and to evaluate the policies (Gustman and Steinmeier, 1986; Rust and Phelan, 1997; French, 2005; French and Jones, 2011; Haan and Prowse, 2014). They model retirement as a result of declining wages, increasing actuarial unfairness of the Social Security and pension system, and increasing tastes for leisure. However, there is a major difference between our model and the previous retirement literature. Prior work typically takes the wage process as given and focuses on the retirement decision itself. For example, when conducting the counterfactual experiment of reducing Social Security benefits by 20%, the previous literature takes the same age-wage profile as in the baseline model with the earnings test and re-estimates the retirement behavior under the new environment without it. As the wage has already been declining significantly and exogenously approaching the retirement age, under the new policy working is still less likely attractive for many workers. However, as we show in our model, when workers expect a smaller Social Security benefit and therefore higher labor supply later in the life-cycle, they adjust their investment over the life-cycle, which results in a higher human capital level as well as higher labor supply earlier in the life-cycle. On average the observed wage levels are 5% higher between 65 and 80. Over the whole life-cycle, observed average yearly wages, total labor income, and total labor force participation rates increase by 1.5%, 2.17%, and 1.57%, respectively. By contrast, in the model with exogenous human capital, the percentage increases in yearly wages, total labor income and total labor supply are less significant, by 0.2%, 1.26%, and 1.31%, respectively. The differences are more dramatic when we evaluate the effect of removing the Social Security system, with the exogenous model underestimating most results.

2 Relevant Literature

Human capital models have been widely accepted as a mechanism to explain life-cycle wage growth as well as the labor supply and income patterns. In his seminal paper, Ben-Porath (1967) develops the human capital model with the idea that individuals invest in their human capital "up front." In what follows we use the two terms—"human capital model" and "Ben-Porath model" —interchangeably. Heckman (1975) and Heckman (1976) further extend the model and present more general human capital models in which each individual makes decisions on labor supply, investment and consumption. In both papers, each individual lives for finite periods and the retirement age is fixed. In their recent paper, Manuelli et al. (2012) calibrate a Ben-Porath model to include the endogenous retirement decision. All three models are deterministic.

Relative to the success in theory, there hasn't been as much work empirically estimating the Ben-Porath model. Mincer (1958) derives an approximation of the Ben-Porath model and greatly simplifies the estimation with a quadratic in experience, which is used in numerous empirical papers estimating the wage process (Heckman et al. (2006) survey the literature). Early work on explicit estimation of the Ben-Porath model was done by Heckman (1975, 1976); Haley (1976), and Rosen (1976). Heckman et al. (1998a) is a more recent attempt to estimate the Ben-Porath model. They utilize the implication of the standard Ben-Porath model where at old ages investment is almost zero. However, this implication does not hold any more when the retirement is uncertain, where each individual always has an incentive to invest a positive amount in human capital. Other more recent work includes Taber (2002), who incorporates progressive income taxes into the estimation, and Kuruscu (2006), who estimates the model nonparametrically. Browning et al. (1999) survey much of this literature.

Another type of human capital model, the learning-by-doing model, draws relatively more attention in empirical work. In the learning-by-doing model human capital accumulates exogenously, but only when an individual works. Thus workers can only impact their human capital accumulation through the work decision. In these models, the total return from labor supply is not only the direct wage income at current time, but also all the extra wage income from the augmented human capital in all future time. Shaw (1989) is among the first to empirically estimate the learning-by-doing model, using the PSID model and utilizing the Euler equations on consumption and labor supply with translog utility. Keane and Wolpin (1997) and Imai and Keane (2004) are two classic examples of research that directly estimate a dynamic life-cycle with learning-by-doing. These papers assume an exogenously fixed retirement age. Wallenius (2009) points out that such a learning-by-doing model does not fit the pattern of wages and hours well at old ages.[1] Heckman et al. (2003) study the potential effects of wage subsidies on skill formulation by comparing on-the-job training models with learning-by-doing models. They simulate the effects of the 1994 EITC schedule for families with two children and find evidence that EITC lowers the long-term wages of people with low levels of education. They find that the learning-by-doing models predictions of the EITC policy effects fit the actual changes better than does a Ben-Porath style model.

There is a large and growing literature on many aspects of retirement. In these models, typically retirement is induced either by increasing utility toward leisure (Gustman and Steinmeier, 1986) or similarly increasing disutility toward labor supply (Blau, 2008). Haan and Prowse (2014) estimate the extent to which the increase in life expectancy affects retirement. Blau (2008) evaluates the role of uncertain retirement ages in the retirement-consumption puzzle.

Retirement can also be induced by declining wages at old ages and/or fixed costs of working. Rust and Phelan (1997) estimate a dynamic life-cycle labor supply model

[1] However, if one interprets the hourly wages as labor income and hours as labor force participation rates (since there is no participation decision in their model), the fit in Imai and Keane (2004) will be improved at old ages.

with endogenous retirement decisions to study the effect of Social Security and Medicare in retirement behavior. French (2005) estimates a more comprehensive model including savings to study the effect of Social Security and pension as well as health in retirement decisions. French and Jones (2011) evaluate the role of health insurance in shaping retirement behavior. Casanova (2010) studies the joint retirement decision among married couples. Prescott et al. (2009) and Rogerson and Wallenius (2010) present models where retirement could be induced by a convex effective labor function or fixed costs.

In all the retirement literature listed above—theoretical or empirical—the wage process is assumed to be exogenous. That is, even when the environment changes while conducting counterfactual experiments, for example changing the Social Security policies, the wage process is kept the same and only the response in the retirement decision is studied.

3 Model

We present and estimate a Ben-Porath human capital model with endogenous labor supply and retirement in which individuals choose consumption, human capital investment, and labor supply (including retirement as a special case). For simplicity we suppress the individual subscript i for all variables. We allow for heterogeneity in some of the parameters when estimating the model. We delay discussion of this to Section 4.2 for expositional convenience.

3.1 Set-up

Time is discrete. Each individual lives from period $t = 0$ to $t = T$. At the beginning of the initial period, each individual is endowed with an initial asset $A_0 \in \mathbb{R}$ and an initial human capital level $H_0 \in \mathbb{R}^+$.

We model the extensive margin of labor supply, so at each period the individual decides either to work or not. The flow utility at time t is

$$u_t\left(c_t, \ell_t\right) = \frac{c_t^{1-\eta_c}}{1-\eta_c} + \gamma_t \ell_t \tag{1}$$

where c_t is consumption and $\ell_t \in \{0, 1\}$ is leisure. The coefficient γ_t represents taste for leisure. We allow for shocks in γ_t which is assumed to be an i.i.d. random variable for each individual.[2]

[2]A key part of our exercise is that we do not allow γ_t to vary systematically across age. We describe the

If an individual chooses to work, he decides on how much time, I_t, to invest in human capital and spends the rest, $1 - I_t$, at effective (or productive) work from which the wage income is earned. Human capital is produced according to the production function

$$H_{t+1} = (1-\delta) H_t + \xi_t \pi I_t^{\alpha_I} H_t^{\alpha_H} \tag{2}$$

where H_t is the human capital level at period t. The ξ_t is an idiosyncratic shock to the human capital innovation. If an individual chooses not to work, he does not invest in human capital (so $I_t = 0$) and human capital depreciates at rate δ.

The labor market is perfectly competitive. We normalize the rental rate of human capital to one so that the wage for the effective labor supply equals the human capital H_t. Thus pre-tax labor income at any point in time is $H_t (1 - \ell_t) (1 - I_t)$.

While we have tried to keep the basic model as simple as possible, the social security system in the U.S. is such a crucial part of the retirement decision that we incorporate it into the model. We model the social security enrollment decision as a one time decision. Once a person turns 62 they can start claiming social security and once they have started claiming, they continue to collect benefits until their death. We will let ssa_t denote a binary decision variable indicating whether a person starts claiming at time t and let ss_t be a state variable indicating whether a person began claiming prior to period t. Since claiming is irreversible, once $ss_t = 1$ then ssa_t is no longer a relevant choice variable. Thus the law of motion can be written as

$$\begin{aligned} ss_0 &= 0 \\ ss_t &= \max \{ss_{t-1}, ssa_{t-1}\}. \end{aligned}$$

The claiming decision (ssa_t) is made separately from the labor force participation decision (ℓ_t) so that one can receive the social security benefit while working (subject to applicable rules such as the earnings test).

Once they have begun claiming, an individual collects befits ssb_t which are a function of the claiming age and the Average Indexed Monthly Earnings ($AIME_t$) when $ss_t = 1$. In practice we approximate the AIME and use the social security rules as of 2004. Details are in the Appendix. This is incorporated into the budget constraint

$$A_{t+1} = (1+r)A_t + Y_t (H_t (1-\ell_t)(1-I_t), ssb_t) - c_t + \tau_t, \tag{3}$$

where A_t stands for asset and r is the risk free interest rate. $Y_t(\cdot, \cdot)$ is the after-tax income

exact process in the next subsection.

which is a function of wage income, the social security benefit ssb_t, and the tax code. Government transfers, τ_t, provide a consumption floor $\underline{c}$ as in Hubbard et al. (1995) so

$$\tau_t = \max\left\{0,\ \underline{c} - ((1+r)A_t + Y_t)\right\}. \tag{4}$$

Life ends at the end of period T and each individual values the bequest he will leave. It takes the form

$$b(A_{T+1}) = b_1 \frac{(b_2 + A_{T+1})^{1-\eta_c}}{1-\eta_c} \tag{5}$$

where b_1 captures the relative weight of the bequest motive and b_2 determines its curvature as in DeNardi (2004).

3.2 Solving the model

Two shocks affect individuals: the leisure shock, γ_t, and the human capital innovation shock, ξ_t. The timing of the model works as follows: at the beginning of each period γ_t is realized by the agent. He then simultaneously chooses consumption, labor supply, human capital investment, and social security application when relevant. After these decisions are made, the human capital innovation shock ξ_t is realized, which determines the human capital level in the following period. Both γ_t and ξ_t are i.i.d. shocks from the perspective of the agents-so agents have no private information about their value prior to their realizations.

The recursive value function can be written as

$$V_t\left(X_t, \gamma_t\right) = \max_{c_t, \ell_t, I_t, ssa_t} \left\{u_t\left(c_t, \ell_t, \gamma_t\right) + \beta E\left[V_{t+1}\left(X_{t+1}, \gamma_{t+1}\right) \middle| X_t, c_t, \ell_t, I_t, ssa_t\right]\right\} \tag{6}$$

where $X_t = \{A_t, H_t, AIME_t, ss_t\}$ is the vector of state variables. The expectation is over the leisure shock in γ_{t+1} and the human capital innovation ξ_t.

The solution to the agent's problem each period is done in two stages. We first solve for the optimal choices conditional on the labor supply decision and then we determine the labor supply decision.

The optimal consumption $\mathcal{C}_{t,0}\left(X_t\right)$, investment $\mathcal{I}_{t,0}\left(X_t\right)$, and social security $\mathcal{SSA}_{t,0}(X_t)$ claiming decisions conditional on participating in the labor market ($\ell_t = 0$) depend only on X_t and can be obtained from

$$\left\{\mathcal{C}_{t,0}\left(X_t\right), \mathcal{I}_{t,0}\left(X_t\right), \mathcal{SSA}_{t,0}(X_t)\right\} \equiv \operatorname*{argmax}_{c_t, I_t, ssa_t} \left\{\frac{c_t^{1-\eta_c}}{1-\eta_c} + \delta E\left[V_{t+1}\left(X_{t+1}, \gamma_{t+1}\right) \middle| X_t, c_t, \ell_t = 0, I_t, ssa_t\right]\right\} \tag{7}$$

and the conditional value function is

$$\widetilde{V}_{t,0}(X_t) \equiv \frac{(\mathcal{C}_{t,0}(X_t))^{1-\eta_c}}{1-\eta_c} + \beta E\left[V_{t+1}(X_{t+1},\gamma_{t+1})\middle| X_t, \mathcal{C}_{t,0}(X_t), \ell_t = 0, \mathcal{I}_{t,0}(X_t), \mathcal{SSA}_{t,0}(X_t)\right] \quad (8)$$

Similarly, conditional on not working ($\ell_t = 1$), we can calculate the optimal consumption and claiming decision from

$$\{\mathcal{C}_{t,1}(X_t), \mathcal{SSA}_{t,1}(X_t)\} \equiv \underset{c_t, ssa_t}{\operatorname{argmax}} \left\{ \frac{c_t^{1-\eta_c}}{1-\eta_c} + \beta E\left[V_{t+1}(X_{t+1},\gamma_{t+1})\middle| X_t, c_t, \ell_t = 1, I_t = 0, ssa_t\right] \right\} \quad (9)$$

and define the conditional value function apart from γ_t to be

$$\widetilde{V}_{t,1}(X_t) \equiv \frac{(\mathcal{C}_{t,1}(X_t))^{1-\eta_c}}{1-\eta_c} + \beta E\left[V_{t+1}(X_{t+1},\gamma_{t+1})\middle| X_t, \mathcal{C}_{t,1}(X_t), \ell_t = 1, I_t = 0, \mathcal{SSA}_{t,1}(X_t)\right]. \quad (10)$$

Notice that since there is no serial correlation in the stochastic shocks of leisure, γ_t, the conditional policy and value functions defined in equations (7)-(10) do not depend on γ_t.

The individual works if

$$\widetilde{V}_{t,0}(X_t) \geq \widetilde{V}_{t,1}(X_t) + \gamma_t.$$

This means that there exists a threshold value $\gamma_t^*(X_t)$ such that

$$\ell_t = \begin{cases} 1, & \text{if } \gamma_t > \gamma_t^*(X_t) \\ 0, & \text{if } \gamma_t \leq \gamma_t^*(X_t) \end{cases} \quad (11)$$

where the threshold value $\gamma_t^*(X_t)$ is defined as

$$\gamma_t^*(X_t) = \widetilde{V}_{t,0}(X_t) - \widetilde{V}_{t,1}(X_t) \quad (12)$$

We use the the parametric form for γ_t,

$$\gamma_t = exp\left(a_0 + a_\varepsilon \varepsilon_t\right) \quad (13)$$

where ε_t follows an independent and identically-distributed standard normal distribution. Therefore γ_t follows a log-normal distribution, $\ln \gamma_t \sim \mathcal{N}\left(a_0, a_\varepsilon^2\right)$. Then we can calculate the threshold value of ε_t as[3]

$$\varepsilon_t^*(X_t) \equiv \frac{1}{a_\epsilon}\left\{\log\left(\gamma_t^*(X_t)\right) - a_0\right\}. \tag{14}$$

Since γ_t is log-normal,

$$E\left(\gamma_t \mid \gamma_t > \gamma_t^*(X_t)\right) = exp\left(a_0 + \frac{a_\epsilon^2}{2}\right)\frac{\Phi\left(a_\epsilon - \varepsilon_t^*(X_t)\right)}{\Phi\left(-\varepsilon_t^*(X_t)\right)}$$

So

$$\begin{aligned} E\left[V_{t+1}(X_{t+1},\gamma_{t+1}) \mid X_{t+1}\right] = {} & \Phi\left(\varepsilon_{t+1}^*(X_{t+1})\right)\widetilde{V}_{t+1,0}(X_{t+1}) + \left(1 - \Phi\left(\varepsilon_{t+1}^*(X_{t+1})\right)\right)\cdot \\ & \left[\widetilde{V}_{t+1,1}(X_{t+1}) + exp\left(a_0 + \frac{a_\epsilon^2}{2}\right)\frac{\Phi\left(a_\epsilon - \varepsilon_{t+1}^*(X_{t+1})\right)}{\Phi\left(-\varepsilon_{t+1}^*(X_{t+1})\right)}\right] \end{aligned}$$

Finally note that X_{t+1} is a known function of $X_t, c_t, \ell_t, I_t, ssa_t$ and ξ_t, so to solve for

$$E\left[V_{t+1}(X_{t+1},\gamma_{t+1}) \mid X_t, c_t, \ell_t, I_t, ssa_t\right] = E\left[E\left(V_{t+1}(X_{t+1},\gamma_{t+1}) \mid X_{t+1}\right) \mid X_t, c_t, \ell_t, I_t, ssa_t\right]$$

we just need to integrate over the distribution of ξ_t. We assume it is i.i.d and follows a log-normal distribution,

$$\log(\xi_t) \sim \mathcal{N}\left(-\frac{\log\left(\sigma_\xi^2 + 1\right)}{2}, \log\left(\sigma_\xi^2 + 1\right)\right) \tag{15}$$

so that ξ_t has mean of one and variance of σ_ξ^2.

4 Estimation

The estimation of the model is carried out using a two-step strategy. First, we pre-set parameters that either can be cleanly identified without explicitly using our model or are not the focus of this paper. In the second step, we estimate the remaining preference and production parameters of the model using the method of simulated moments (MSM). The model is described by equations 1-6 and we summarize the parameters here. The parameters related to preferences are the intertemporal elasticity of consumption, η_c, the discount rate, β, the parameters determining the taste for leisure, a_0 and a_ε, and the be-

[3] We are implicitly assuming that $\widetilde{V}_{t,0}(X_t) > \widetilde{V}_{t,1}(X_t)$. If this is not the case, the individual works with probability 1.

quest parameters b_1 and b_2. Human capital production is determined by π, α_I, α_H and σ_ξ. Parameters related to the budget constraint are the interest rate (r) and the consumption floor ($\underline{c}$). Finally there are initial values for the state variables, assets, A_0, human capital, H_0, and Averaged Indexed Monthly Earnings, $AIME_0$.

4.1 Pre-set Parameters

The set of parameters pre-set in the first stage includes the interest rate, initial wealth and initial AIME, the time discount rate, CRRA of utility in consumption, consumption floor, and bequest shifter. Specifically, we do not try to match moments related to consumption or assets as those are not the focus of this paper. Of course the marginal utility of consumption plays an important role in determining the optimal human capital accumulation. Given the total wealth level, the consumption allocation across periods are jointly determined by β, η_c, A_0, $\underline{c}$, b_1, and b_2. Separately identifying these parameters require matching moments related to consumption or asset accumulation. For this reason we fix β, η_c, A_0, $\underline{c}$, and b_2, and only estimate b_1. In Section 7.2 we look at the sensitivity of some of our results to these values.

One period is defined as one year.[4] The initial period in our model corresponds to individuals age 18 and ends at age 80.[5] The early retirement age is 62 and the normal retirement age is 65. The risk free real interest rate is set as $r = 0.03$ and the time discount rate is set as $\beta = 0.97$. The coefficient of constant relative risk aversion (CRRA) in the utility from consumption is set as $\eta_c = 4.0$. The consumption floor is set as $\underline{c} = 2.19$, following French and Jones (2011).[6]

The parameter which determines the curvature of the bequest function is set as $b_2 = 300$.[7] This number is close to French (2005) where he sets $b_2 = 250$ or French and Jones (2011) where they estimate $b_2 = 222$.[8] We assume all individuals start off their adult life with no wealth and zero level of AIME at age 18. These normalized or pre-set parameters are summarized in Table 1.

[4]Mid-year retirement might be an issue. However, more than half of workers are never observed working half-time approaching retirement, so it would not be a big issue.

[5]The life expectancy for white males is 74.1 in 2000 and 76.5 in 2010.

[6]$\underline{c} = 4380/2000 = 2.19$ since we normalize the total time endowment for labor supply at one period as one.

[7]It is equivalent to \$600,000 in 2004 U.S. dollar.

[8]They are \$500,000 and \$444,000 in their papers.

Table 1: Normalized or Pre-set Parameters

Parameters		Normalized/Pre-set Values
Interest rate	r	0.03
Discount	β	0.97
Risk Aversion	η_c	4.0
Initial wealth[a]	A_0	0.0
Initial AIME[a]	$AIME_0$	0.0
Consumption floor[b]	$\underline{c}$	2.19
Bequest shifter[c]	b_2	300.0

[a]The initial age is 18.
[b]The consumption floor is equivalent to \$4380 in 2004\$, since we normalize the total time endowment for labor supply at one period—which is 2000 hours—as one.
[c]The bequest shifter is equivalent to $\$600,000$.

4.2 Heterogeneity

This leaves the following parameters: $b_2, a_0, a_\varepsilon, \pi, \alpha_I, \alpha_H$, and H_0. We allow for heterogeneity in three of these: ability to learn (π) , ability to earn (H_0), and tastes for leisure (a_0). For computational reasons we only have nine types determining the joint distribution of (a_0, π). Specifically, we model it as a nine point Gauss-Hermite approximation of a joint normal distribution, which depends on five parameters: the mean and variance of a_0, the mean and variance of π, and the correlation between the two. Respectively we write this as $(\mu_{a_0}, \sigma_{a_0}, \mu_\pi, \sigma_\pi, \rho)$. We emphasize that since we are only using nine points we are not assuming that the Gauss-Hermite is a good approximation of a normal, but rather view this as a parametization itself. That is, we assume that the joint distribution of (a_0, π) is a parametric discrete distribution with 9 points determined by the parameter vector $(\mu_{a_0}, \sigma_{a_0}, \mu_\pi, \sigma_\pi, \rho)$.

Since human capital is already a state variable in our model, we can be more flexible in modeling initial human capital. We also it to be correlated with (a_0, π) though the functional form

$$H_0 = \exp\left(\gamma_0 + \gamma_{a_0} a_0 + \gamma_\pi \pi + \sigma_{H_0} \nu\right) \tag{16}$$

where $\nu \sim \mathcal{N}(0,1)$ is an i.i.d standard normal random variable.

4.3 Estimation Procedure

We apply the method of simulated moments to estimate the parameters of interest, Θ,

$$\Theta = \left\{ \underbrace{\mu_{a_0}, \sigma_{a_0}, a_\varepsilon,}_{\text{leisure}} \underbrace{\delta, \alpha_I, \alpha_H, \sigma_\xi, \mu_\pi, \sigma_\pi,}_{\text{human capital production}} \underbrace{b_1,}_{\text{bequest}} \underbrace{\rho,}_{\text{correlation } (a_0,\pi)} \underbrace{\gamma_0, \gamma_{a_0}, \gamma_\pi, \sigma_{H_0}}_{\text{initial human capital}} \right\}$$

according to the following procedure.

i) Calculate the moments from the data.

ii) Iterate on the following procedure for different sets of parameters of Θ until the minimum distance has been found.

   (a) Given a set of parameters, solve value functions and policy functions for the entire state space grid.

   (b) Generate the life-cycle profile for each simulated individual.

   (c) Calculate the simulated moments.

   (d) Calculate the distance between the simulated moments and the data moments.

4.4 Data and Moments

Our primary data set is the Survey of Income and Program Participation (SIPP). The SIPP is comprised of a number of short panels of respondents and we use all of the panels starting with the 1984 panel and ending with the 2008 panel. To focus on as homogeneous group as possible, the sample only includes white male high school graduates.

Our measure of labor force participation is a dummy variable for whether the individual worked during the survey month.[9] Clearly the aggregation is imperfect. We could use participation in a year, but this would miss much of the extensive labor supply decisions of men. Ideally we would estimate the model at the monthly level, but this is not computationally feasible. We construct the hourly wage as the earnings in the survey month divided by the total number of hours worked in the survey month.

We begin estimation of the model from age 22 rather than 18 for two reasons. First, we have a short panel meaning that many 19 year old high school graduates may return to college after they leave the panel. Second, our model does not include any search or matching behavior, which might be important for the labor force patterns among very

[9] In SIPP an individual is observed in at most three months each year. If an individual is observed working more than 50% of the time then he is categorized as participating in the labor force, otherwise not. If one is sampled twice for the year and is observed working in one month only, the participation status is determined randomly (50% for each possibility).

recent labor force entrants as they transition from school to work as suggested in the literature (Topel and Ward, 1992; Neal, 1999). Our model does over-predict the labor supply for those individuals.

Four sets of moment conditions at each age from 22 to 65 are chosen to represent the life-cycle profiles. We use a total of 230,657 panel observations from 80,519 different respondents.

i) The labor force participation rates (LFPR);
ii) The first moments of the logarithm of observed wages;
iii) The first moments of the logarithm of observed wages after controlling for individual fixed effects.[10]
iv) The second moments (standard deviation) of the logarithm of observed wages.

As is standard in the literature on estimation of Ben-Porath style human capital we assume that wages in the data correspond to

$$W_t = H_t \left(1 - I_t\right) \tag{17}$$

in the model. We use this value to match the data moments.[11] We match both age-wage profiles, with and without controlling for individual fixed effect as the two have quite different patterns.

Figures 1a-1c present these four profiles. Figure 1a plots the labor force participation rates between age 22 and 65. Figure 1b plots two log wage profiles. The first one is the log wage profile from the pooled sample, while the second one is the log wage profile after controlling for individual fixed effects. The original log wage profile has a hump shape, but the one filtering out individual fixed effects does not decline within the examined period which is between age 22 and 65. Figure 1c shows that the extent to which the variance of log wages increases with age.

The most interesting result in Figures 1a-1c is the discrepancy between the age-wage profile with or without controlling for individual fixed effects. This has been documented in various data sets, including the National Longitudinal Survey of Older Men (NLSOM) data (Johnson and Neumark, 1996), the Panel Study of Income Dynamics (PSID) data

[10] To construct these moments we first regress log wage on the age dummies and survey year dummies and obtain the predicted log wage, denoted as z. We pick a base age (age 30) and calculate the average predicted log wage at the base age for each year, denoted as $\bar{z}_{a,j}$, where a is the base age and j is for survey year. We then pick a base year y and calculate the difference of $\bar{z}_{a,j}$ between each year j and the base year y, denoted as $\Delta\bar{z}_{a,j}$. Finally we calculate the difference between the original log wage and $\Delta\bar{z}_{a,j}$ and define the result as $\ln \tilde{W}_t$, which is the log wage after filtering out the time fixed effects.

[11] That is, econometricians cannot distinguish the effective labor supply $\left(1-\ell_t\right)\left(1-I_t\right)$ from the observed labor supply $\left(1-\ell_t\right)$ since the investment time I_t is not observable.

(Rupert and Zanella, 2012), and the Health and Retirement Survey (HRS) data (Casanova, 2013). These papers find that after controlling for individual fixed effects the age-wage profile is flatter than the hump-shaped age-wage profile estimated using pooling observations, and it does not decline until 60s or late 60s. All of these papers argue that this evidence is not consistent with the traditional human capital model since the traditional human capital model would predict a hump-shaped wage. The intuition is that when the human capital depreciation outweighs the investment, wages start to decline which generates a hump-shaped profile. Fitting the wage profile after controlling for fixed effects makes our problem more challenging because we need to explain the decrease in labor supply later in life when there is little evidence that wages decline.

To further verify this result we compare our SIPP results with the Current Population Survey (CPS) data. From the CPS Merged Outgoing Rotation Groups (MORG) data, we match the same respondent in two consecutive surveys[12] using the method proposed in Madrian and Lefgren (2000). We have a short panel with each individual interviewed twice, one year apart. We construct a similar short panel from the CPS March Annual Social and Economic Supplement files (March). The difference is that the wage information is collected from the reference week in the CPS MORG data and from the previous year in the CPS March data.

Figure 2 presents the age-wage profiles with or without controlling for individual fixed effects for male high school graduates from the CPS MORG data and the CPS March data. We find a very similar discrepancy in the age-wage profiles as in the SIPP data presented in Figure 1b. Our model is able to reconcile such a discrepancy in the age-wage profiles, as we show in the next section.

5 Estimation Results

The estimates of the parameters are listed in Table 2. Of particular importance are the depreciation rate, δ, curvature in the human capital production function, α_I, and a_ε which determines the elasticity of labor supply. Before discussing these parameter values we examine the fit of the model in Figures 3a-3d.[13]

The first and central point is that our parsimonious model can reconcile the main facts

[12] For MORG data, they are the fourth and eighth interview.

[13] The overidentification test statistic is reported in the bottom of Table 2. The model is rejected at the 1% level but not at the 0.5% level. The fact that we reject is not surprising given the simplicity of our model and the size of our sample. One could easily add some extra parameters to pass the statistical criterion, but this is not our goal. Our goal is to use a simple model that does a very good job of capturing the lifecycle patterns.

Table 2: Estimates in the Baseline Model[a]

Parameters		Estimates	Standard Errors
Leisure: Standard Deviation of Shock	a_ε	0.433	(1.265×10^{-3})
Human Capital Depreciation	δ	0.101	(9.673×10^{-4})
Human Capital Production Function: *I* factor	α_I	0.076	(3.024×10^{-3})
Human Capital Production Function: *H* factor	α_H	0.151	(1.169×10^{-3})
Standard Deviation of Human Capital Innovation	σ_ξ	0.405	(2.412×10^{-3})
Bequest Weight	b_1	$424,070$	(7.727×10^{-3})
Parameter heterogeneity[b]			
Leisure: Mean of Intercept	μ_{a_0}	-6.525	(8.591×10^{-4})
Leisure: Standard Deviation of Intercept	σ_{a_0}	0.874	(4.289×10^{-4})
Human Capital Productivity, Mean	μ_π	1.758	(1.606×10^{-5})
Human Capital Productivity, Standard Deviation	σ_π	0.583	(1.359×10^{-3})
Correlation between $a_0$and π	ρ	-0.893	(5.837×10^{-3})
Initial Human Capital Level at Age 18			
Intercept	γ_0	1.625	(9.813×10^{-4})
Coefficient on a_0	γ_{a_0}	0.052	(7.290×10^{-5})
Coefficient on π	γ_π	0.531	(2.653×10^{-4})
Standard Deviation of Error Term	σ_{H_0}	0.239	(0.944)
χ^2 Statistic = 212[c]		Degrees of freedom = 161	

[a]Method of simulated moments estimates. Estimates use a diagonal weighting matrix. Standard errors are given in parentheses.
[b]The joint distribution of (a_0, π) is a parametric discrete distribution with nine points determined by these five parameters, using a nine point Gauss-Hermite approximation.
[c]This is the J-statistic. The critical values of the χ^2 distribution are $\chi^2_{(161,0.01)} = 206$, $\chi^2_{(161,0.005)} = 212$, $\chi^2_{(161,0.001)} = 222$.

in the data: a small increase in labor supply/large increase in wages at the beginning of the life-cycle along with the large decrease in labor supply/small decrease in wages at the end of the life-cycle.[14]

The simulated labor force participation rate increases slightly between age 22 and 30 as shown in Figure 3a.[15] Our main result is that this simple model is able to generate a massive decline in labor supply between age 55 and 65, which fits the sharp decline of labor force participation rates within that age period in the data and simultaneously the flat wage profile in the fixed effect model.

Our model generates a similar discrepancy between the log wages with and without controlling for individual fixed effects, as shown in Figures 3b and 3c, and both profiles fit the data well. Log wages after filtering out individual fixed effects increase at a decreasing pace from age 22 to age 55 and then decreases slightly (Figure 3b). On the other hand, Figure 3c shows that the original log wage profile presents a hump shape which almost replicates the data profile very well.

The model also replicates the log wage variation as in the data (Figure 3d). This increasing variation mainly comes from the heterogeneity in the parameters. Without heterogeneity in parameters, the wage variation would decrease with age as human capital would converge due to concavity of the production function. With heterogeneity, the human capital level might diverge, depending on parameter values.

We obtain our fit of the life-cycle profiles of labor supply and log wages despite the lack of any explicit time-dependent preference, production or constraints in our model. Two key features of our model make them possible: human capital depreciation and the separation between the effective labor and observed labor. We discuss each of these in turn.

Human capital depreciation is essential for matching the labor force participation profile; especially in inducing massive retirement at old ages without a very large increase in labor supply at young ages. If human capital did not depreciate, then it would not be optimal for workers to stop working while their human capital was still high. Instead, workers would spread their leisure more evenly across their life-cycle.[16] However, as long as human capital depreciates over time, this is no longer the case. Given an initial human capital level which becomes lower when idle, it is better to utilize it immediately and cluster leisure at the end of the life-cycle where human capital becomes relevant.

[14]One should keep in mind that our parsimonious specification might be a limitation on our policy counterfactuals as other features that we have not explicitly modeled might impact those simulations.

[15]Though the increase in our simulations is smaller than the increase in the data—however we show later (Figure 9a) that with initial assets of $50,000 we can fit this profile perfectly.

[16]Borrowing is allowed in our model.

This discussion implies that our estimate of a depreciation value $\delta = 0.101$ is empirically important. Given this, it is important to place this value into the range of estimates in the literature. This is not easily done is that there is a very large range of estimates-some larger than our 10.1% estimate and some smaller. There are broadly three different literatures that estimate related parameters. The first of these literatures is motivated by family leave for women and tries to estimate the effect of career interruption on wages. It finds estimates ranging from 1.5% per year to 25%.[17] A second literature looks at displacement from the Displaced worker survey and also finds a wide range of estimates-many of which are not directly comparable to ours.[18] A third literature examines the effect of the length of an unemployment spell on the wage at rehire. Schmeider et al. (2014) is a recent and convincingly identified paper of this type. They estimate the effect using a regression discontinuity with German data. In Germany the length of eligibility for unemployment insurance depends on age with jumps at ages 42 and at 44. They see an increase in unemployment duration at these two discontinuity points, so they use the kink points as instruments in order to estimate the effect of the length of unemployment duration on reemployment wages. They find that one extra month of unemployment leads to a decrease in wages of 0.8% which gives an annual rate remarkably close to our estimate of 10.1%.

A second import feature for explaining the lifecycle profiles comes from a point emphasized by Heckman et al. (1998a): observed wages are different than observed human capital. We see in figure 3b that in both the model and the data, once fixed effects are accounted for, wages are close to flat for ages 50-65 despite the fact that there is a large

[17] A classic early paper on this topic is Mincer and Polachek (1974) which estimates a net depreciation rate of around 1.5 percent per year. Mincer and Ofek (1982) go beyond this to discuss the difference between short term and long term losses from interruption. In the long run individuals invest in human capital to offset the initial loss, so Mincer and Ofek (1982)'s definition of short term losses is more closely related to our concept of depreciation. Using panel data methods for the National Longitudinal Survey of Mature Women they find estimates ranging from 5.6% to 8.9%. Light and Ureta (1995) use National Longitudinal Survey of Youth 1979 data and estimate that the immediate effect of a year of non-participation in the labor market for one year leads to a decline in earnings of 25%. Kunze (2002) and Gorlich and de Grip (2009) both use German data (IAB employment sample and German Socio-economic panel respectively). Kunze (2002) finds estimates of about 2-5% wages losses for women from unemployment spells but about 13-18% from parental leave. Gorlich and de Grip (2009) find a variety of results ranging from around 1.5% to 5% depending on the type of spell.

[18] While much of this literature is more focused on earnings than wages, some papers look at weekly earnings. Both Farber (1993) and Ruhm (1991) estimate the effect of a displacement on re-employment wages and obtain a range of estimates with most being around declines of 10% but varying from 6.5% to 16.9%. These numbers are not annualized but are just from the incidence of displacement. Li (2013) uses the same data but produces annualized versions so that the effects can be more easily compared to our estimate of δ. She estimates the effects for many different occupations with a huge range of estimates across occupations. Focusing on the three largest occupations she finds a deprecation of 9.4% for Installation and Repair workers, 7.7% for Production workers, and 17.4% for workers in Transportation.

decrease in labor supply. This aspect of the model can help explain this effect. As shown in Figure 4a, at older ages the actual human capital level has already depreciated to a relatively low level (lower than the initial level at age 18), even though the observed wage level is still quite high. This is due to the decline in investment that happens around that time. This means that measured wages, $H_t(1 - I_t)$ can be flat while H_t is decreasing as long as I_t is decreasing as well. The time investment profile in Figure 4b matches this implication. The solid line is the unconditional investment profile while the dashed line is the average investment profile conditional on working. These two profiles are very close to each other at prime ages, and both decrease over time.

This relatively high value of investment late in the working career is also related to why we find a much smaller level of the human capital curvature parameter, α_I compared to the literature summarized in Browning et al. (1999). The larger is α_I the steeper is the decline in human capital investment with age. At the extreme when $\alpha_I = 1$ one gets a "bang-bang" solution with full investment to a point and then zero investment thereafter. Because depreciation is large, in order to fit the relatively flat wage profile that we see at older ages one needs a lot of investment at this age which requires a small value of α_I. Heckman et al. (1998a) fit the wage data with a much larger value of α_I but this results, in part, from the fact that they set deprecation to zero.

At the early stage of the life-cycle, workers invest a considerable amount of time in the human capital production which drives up both the human capital level and the wage. Once the worker reaches his mid-career (around age 45), he reduces the time investment and human capital starts to decrease. As the worker spends less of his working time investing, wages continue to increase. One can see in Figure 4a that the observed wage keeps increasing after age 45 and peaks around 52. After age 52, however, since the worker has already allocated most of his time in effective working, there is no further room for such adjustment. As a result, the observed wage declines at almost the same rate at which human capital depreciates. This leads to large falls in labor supply at older ages.[19]

Such separation also helps generate the pattern that the working hours profile peaks earlier than the wage profile (Weiss, 1986). Working hours increase slightly with age when the worker is young, with a large portion devoted to human capital investment. The working hours profile peaks around age 40 and starts declining. However, with proportionally less time devoted to human capital investment and more time to effective labor supply (Figure 4b), the observed wage increases from labor market entry to about

[19] This also explains why the estimated depreciation rate of human capital is quite high in our estimation, $\sigma = 10.0\%$, comparing with 2.4% in Manuelli et al. (2012).

age 52.

5.1 Elasticity of Labor Supply

In this subsection, we investigate the model's implications for elasticities of labor supply. Since labor supply is discrete, we examine the elasticity along the extensive margin. At the individual level, the labor supply elasticity is zero unless the worker is exactly indifferent between working or not, in which case it is infinite. Therefore, we can not construct the standard Marshallian and Hicksian labor supply elasticities. Instead we construct counterparts to these by increasing the human capital rental rate at different ages by 10% (from 1 to 1.1), and then simulating the percentage change in the labor force participation rate using the baseline model.[20]

Let h_t^b be the labor force participation rate at age t in the baseline model and h_t^t be the labor force participation rate at age t in the simulation in which we increase the rental rate at age t by 10%. Then our version of the Marshallian is calculated as

$$me = \frac{\log\left(h_t^t\right) - \log\left(h_t^b\right)}{\log(1.1)}. \tag{18}$$

We calculate the Intertemporal Elasticity of Substitution (IES) as

$$ies = \frac{\log\left(h_t^t/h_{t-1}^t\right) - \log\left(h_t^b/h_{t-1}^b\right)}{\log(1.1)}. \tag{19}$$

The whole life-cycle age-wage profile will be different in this model even when the only change is in the rental rate at age t. An alternative and probably more empirically relevant way of calculating these elasticities is to compute the percentage changes in the labor supply responding to the percentage changes in the observed wages,

$$me' = \frac{\log\left(h_t^t\right) - \log\left(h_t^b\right)}{\log(w_t^t) - \log(w_t^b)} \tag{20}$$

$$ies' = \frac{\log\left(h_t^t/h_{t-1}^t\right) - \log\left(h_t^b/h_{t-1}^b\right)}{\log\left(w_t^t/w_{t-1}^t\right) - \log\left(w_t^b/w_{t-1}^b\right)}. \tag{21}$$

The calculated Marshallian elasticity and IES at each age from both methods are plotted in Figure 5a. Table 3 also documents both elasticities at selected ages. One can see that labor supply is much more elastic at older ages than at younger ages in both cal-

[20] In both simulations we assume that the increase in rental rates is anticipated.

Table 3: Elasticities at selected ages

Age	Responding to % changes in H rental rate[a]		Responding to % changes in wages[b]	
	Marshallian (*me*)	IES (*ies*)	Marshallian (*me*′)	IES (*ies*′)
20	0.302	0.274	0.271	0.243
25	0.220	0.197	0.201	0.178
30	0.225	0.210	0.207	0.193
35	0.212	0.191	0.197	0.177
40	0.213	0.193	0.198	0.179
45	0.251	0.227	0.234	0.212
50	0.349	0.322	0.329	0.303
55	0.492	0.430	0.475	0.416
60	0.754	0.626	0.724	0.592
65	1.211	1.015	1.154	0.951
70	1.578	1.283	1.719	1.424

[a]The Marshallian is $me = \frac{\log(h_t^t) - \log(h_t^b)}{\log(1.1)}$; the IES is $ies = \frac{\log(h_t^t/h_{t-1}^t) - \log(h_t^b/h_{t-1}^b)}{\log(1.1)}$.

[b]The Marshallian is $me' = \frac{\log(h_t^t) - \log(h_t^b)}{\log(w_t^t) - \log(w_t^b)}$; the IES is $ies' = \frac{\log(h_t^t/h_{t-1}^t) - \log(h_t^b/h_{t-1}^b)}{\log(w_t^t/w_{t-1}^t) - \log(w_t^b/w_{t-1}^b)}$

culations. This is due to the fact that labor supply has dual roles as one can not invest in human capital without working. Thus young workers will respond to a temporary drop in wages in part by increasing human capital investment rather than just increasing leisure. However, this margin is not important for older workers. As a result, the labor supply of young workers is less responsive to temporary wage shocks than is the labor supply of older workers. Note that the second measure of the Marshallian elasticity or IES is universally smaller than the first. The reason is that at age t the percentage change in the wage is larger than that in the human capital rental rate. As a result of workers' responses to the anticipated rental rate increase, they adjust their investment strategy to take advantage of the higher rental rate at age t.

Figure 5b provides some sense of how these temporary effects impact lifetime labor supply. The left panel presents the effect of LFPR profiles for cases where the 10% increase in the human capital rental rate occurs at different ages, specifically at ages 25, 40, and 60. This shows the response in LFPR at different ages for the positive shock at one specific age.

The right panel of Figure 5b plots the total change in LFPR for such positive shocks at different ages. Assume that the human capital rental rate only increases at age t and the timing of this shock is represented by the X-axis of this figure. For this case, the "Overall" represents the overall change in LFPR over the entire life-cycle (from age 18 to 80); the "Before t" represents the total change in LFPR before age t; the "After t" is the total change after age t and the “At t” is the spot change at age t. If the human capital rental rate increases at age t, the spot LFPR increases responding to this positive shock (represented by “At t”). Furthermore, before age t, the expected return to working and investing also increases (represented by “Before t”). This leads to the increase in LFPR before age t. This shows that a rational individual responds to the predicted shock before it occurs in this dynamic model. On the other hand, if the positive shock occurs during the early career, the wealth effect causes a decline in the LFPR at later ages. However, a positive shock at older ages would encourage higher LFPR afterwards. This is because one individual allocates more time in effective working at old ages than at young ages. Thus the substitution effect is more prominent at older ages, when the wage is around the peak.

For individuals under age 50 these estimates are very close to the estimates of labor supply elasticities found in the literature. For example, the early literature estimates the Frisch elasticity being 0.09 (Browning et al., 1985), 0.15 (MaCurdy, 1981), and 0.31 (Altonji, 1986). Chetty (2012) reports extensive (Hicksian) labor supply elasticities around 0.25 combining estimates from many different studies and approaches.

Focusing on the extensive margin, Rogerson and Wallenius (2013) suggest that the IES is 0.75 or greater given empirically reasonable level of nonconvexities or fixed costs. The average of our estimates between ages 60 and 65 is remarkably close to theirs.

5.2 The Role of Health

We have intentionally kept our model simple to show that human capital can explain the dramatic fall in labor supply at the end of the life-cycle. However, there are many alternative reasons why labor supply might decline. Aside from Social Security rules, which we have already incorporated, the most important is health (e.g. Currie and Madrian 1999, French and Jones 2011). If the primary reason for retirement is health, its omission might seriously distort our results. In this section we incorporate health into our model in a very flexible way. We show that while it is an important factor, it is not the primary driver of retirement.

We allow that at each period t there is an additional state variable—health status, $S_t \in \{0, 1\}$, with 0 being in good health and 1 in bad health. Each individual is assumed to have good health at the beginning of the first period, $S_0 = 0$. The health status evolves exogenously according to a time-dependent probability transition matrix,[21] and is realized at the beginning of each period before any choice is made.

[21] The health transition matrix is estimated from the Panel Study of Income Dynamics (PSID) data.

We allow the taste for leisure in the utility function (1) to depend on the health status and change with age,

$$\gamma_t = exp\left(a_0 + S_t\left(a_{s0} + a_{st}t\right) + a_\varepsilon \varepsilon_t\right). \tag{22}$$

That is, individuals with bad health have a different taste for leisure than those with good health and this difference changes as they age.[22] We refer to this model as the baseline health model.

To estimate these two new parameters, a_{s0} and a_{st}, we include the difference in labor force participation rates between workers with good health and and workers with bad health, from age 30 to 65 in our moment conditions. The data moments are derived from the 1963-2008 Current Population Survey data.

We then re-estimate the model. The fit of the model is presented in Figure 6a. Including health (and the additional moments) into the model does not improve its performance on the original moments in any significant way. Our simple model is rich enough to explain the life-cycle patterns of labor supply and log wages.

However, just because the fit does not improve much does not imply that health does not play an important role. It may just be that either health or human capital could explain retirement.[23] To explore the implications of health we use the model estimated with health, but then simulate a counterfactual in which there was no health change. Specifically, we eliminate the importance of health for individuals over 50 in two different ways- we do not allow their health to worsen and we eliminate the interaction between health and preferences for work. We simulate an experiment in which the health status that an individual had at age 50 remains the same for the rest of his life. Secondly, for individuals with bad health status on and after age 50, we assume their taste for leisure does not increase with age. That is, we assume the taste for leisure now is

$$\gamma_t = exp\left(a_0 + S_t\left(a_{s0} + a_{st} \cdot \min\left\{t, 50\right\}\right) + a_\varepsilon \varepsilon_t\right) \tag{23}$$

We then re-solve the modified model and simulate the life-cycle profile for each individual using the same estimates from the aforementioned baseline health model.[24] The pro-

[22] A key aspect of the thought experiment behind this paper is to not allow preferences to vary systematically with age in our baseline model. In practice we can only fit the interaction of health and labor supply in the data by allowing for an interaction between health and tastes for leisure. The main point of this section is that health is not essential to explain the profiles, so even though we are favoring the model with health by allowing this extra flexibility, the fit improves very little.

[23] Note that this is not to say they are not separately identified. The extra moments we use for the health model identify the importance of health.

[24] We are assuming that agents have rational expectations and are aware that their health status will not change. We have also simulated models in which they are not aware that their health status will remain fixed-it does not change the basic message.

files of labor supply and human capital are plotted in Figure 6b. The difference between the counterfactual and the baseline health model is very small in both the labor force participation rate and the human capital level. This implies that at least in our model health is not a major factor driving retirement. This result confirms findings in the previous literature. French (2005) estimates that the changes in health attribute to roughly 10% of the drop in the labor force participation rates between ages 55 and 70, and the contribution to hours worked by workers near retirement is much smaller. Blau and Shvydko (2011) also report that health deterioration is an important but not major cause of retirement.

6 Alternative Human Capital Models

We compare our baseline human capital accumulation model with two variants. All other aspects of the model remain the same. The first variation assumes the innovation part in the human capital production function is completely exogenous. The second variation assumes the innovation only occurs if individuals work, but is exogenous conditional on work. This is essentially a learning-by-doing model as in, for example, Imai and Keane (2004). To keep this comparable, we alter our baseline model as little as possible. We also restrict the number of total parameters to remain the same so that we are comparing models with similar levels of flexibility.

First we consider the model with exogenous human capital. In this case human capital evolves according to the function

$$H_{t+1} = (1-\delta)\, H_t + \xi_t \pi \left(1 + \alpha_1 t + \alpha_2 t^2\right)$$

where t is potential experience. Notice that this is very close to our standard model from equation (2). We have exactly the same parameter names, except that (α_I, α_H) are replaced with (α_1, α_2) since their roles have changed considerably. In this case human capital evolves completely exogenously in the sense that individuals can do nothing to change their human capital.

The parameterization of the second model is analogous. Here we alter the exogenous model so that human capital only grows for workers:

$$H_{t+1} = (1-\delta)\, H_t + (1-\ell_t)\, \xi_t \pi \left(1 + \alpha_1 t + \alpha_2 t^2\right).$$

We refer to this as the "learning-by-doing" model. Even though it looks quite similar to the exogenous model, as a practical matter it is very different as workers can control their human capital through their labor force participation decision. When individuals do not work, their human capital depreciates at rate δ.

In section 5 above we discuss two different reasons why our model can fit the life-cycle profiles of wages and labor supply and in particular the large increase in wages but small increase in labor supply at the beginning of the life-cycle and the large decrease in labor supply but small decrease in wages at the end. The first is human capital depreciation- when workers stop working their earnings fall. The second was the distinction between observed wages and human capital. These two models allow for us to see the relative importance for these two different explanations because the exogenous human capital model lacks both of these features while the learning by doing allows for the former but not the latter.

The estimates of these models are presented in Table 4 and the fits of the two models are presented in Figure 7a. We first discuss the completely exogenous model. As expected, it is difficult for this model to fit both the labor force participation and the fixed effect profile at the same time. The fit of the other two moments is also quite off. The problem is that to fit the decrease in labor supply at the end requires a very large labor supply elasticity (as well as a lot of sample selection bias to give an estimated flat wage). However, the large elasticity to explain labor supply at the end leads to a huge increase in labor supply at the beginning that we do not see in the data. To see the size of the elasticity, we estimate our version of the Intertemporal Elasticity of Substitution as above and present it in figure 7b as well as in table 5 at selected ages. The exogenous model requires a substantially larger elasticity.

By contrast the learning-by-doing model fits the data well—though not quite as well as our baseline model. The elasticity of labor supply is larger—as one can see from figure 7b or from the fact that a_ε takes on the smaller value 0.328 as opposed to 0.433. However the elasticity is much closer to the baseline model than it is to the exogenous model. In comparing the fit, all three models explain the fixed effect wage profile fairly well, but both the exogenous model and the learning-by-doing model are a bit off in the labor force participation rate, especially during the early career. They also perform considerably worse in the log wage profile and the standard deviation profile. It is important to note here that we did not try a wide range of learning-by-doing models, we just did a comparison between our baseline model and a learning-by-doing model chosen to be close to our baseline model. Presumably alternative and more flexible models could fit the data better—though this is true of our baseline model as well.

This comparison between the fit of the three models suggests that the human capital depreciation rate seems to be relatively more important for fitting the data than the

Table 4: Estimates of Alternative Models[a]

Parameters		Exogenous[b] Estimates	Exogenous[b] Standard Errors	Learning-by-Doing[c] Estimates	Learning-by-Doing[c] Standard Errors
Leisure: St. Dev. of Shock	a_ε	0.071	$\left(7.093 \times 10^{-4}\right)$	0.328	$\left(1.848 \times 10^{-4}\right)$
HC Depreciation[d]	δ	0.100	$\left(1.170 \times 10^{-4}\right)$	0.100	$\left(9.208 \times 10^{-4}\right)$
HC Prod.: Coef on Age	α_1	6.140×10^{-3}	$\left(2.440 \times 10^{-7}\right)$	4.356×10^{-3}	$\left(5.863 \times 10^{-7}\right)$
HC Prod.: Coef on Age Squared	α_2	-2.376×10^{-4}	$\left(2.456 \times 10^{-3}\right)$	-1.320×10^{-7}	$\left(5.669 \times 10^{-2}\right)$
St. Dev. of HC Innovation	σ_ξ	1.666	$\left(2.687 \times 10^{-4}\right)$	0.271	$\left(7.190 \times 10^{-3}\right)$
Bequest weight	b_1	$432,924$	$\left(3.345 \times 10^{-3}\right)$	298	$\left(6.181 \times 10^{-2}\right)$
Parameter heterogeneity					
Leisure: Mean of Intercept	μ_{a_0}	-6.800	$\left(5.148 \times 10^{-5}\right)$	-6.975	$\left(1.151 \times 10^{-3}\right)$
Leisure: St. Dev. of Intercept	σ_{a_0}	0.966	$\left(8.196 \times 10^{-6}\right)$	0.999	$\left(2.389 \times 10^{-6}\right)$
HC Prod., Mean	μ_π	1.779	$\left(3.269 \times 10^{-6}\right)$	1.737	$\left(1.702 \times 10^{-5}\right)$
HC Prod., Stand. Dev.	σ_π	0.599	$\left(1.915 \times 10^{-4}\right)$	0.673	$\left(4.433 \times 10^{-4}\right)$
Correlation between $a_0$and π	ρ	-1.000	$\left(5.717 \times 10^{-4}\right)$	-0.949	$\left(3.194 \times 10^{-3}\right)$
Initial HC level at age 18					
Intercept	γ_0	2.320	$\left(2.244 \times 10^{-4}\right)$	1.905	$\left(8.147 \times 10^{-4}\right)$
Coefficient on a_0	γ_{a_0}	0.128	$\left(2.444 \times 10^{-5}\right)$	0.022	$\left(2.705 \times 10^{-5}\right)$
Coefficient on π	γ_π	0.349	$\left(5.253 \times 10^{-3}\right)$	0.181	$\left(2.478 \times 10^{-4}\right)$
St. Dev. of Error Term	$\sigma_{H_{18}}$	0.027	(2.643)	0.557	$\left(5.331 \times 10^{-3}\right)$
χ^2 Statistic		977		662	
Degrees of freedom		161		161	

[a]Method of simulated moments estimates. Estimates use a diagonal weighting matrix. Standard errors are given in parentheses.

[b]In the exogenous model, the human capital production function is $H_{t+1} = (1-\delta) H_t + \xi_t \pi \left(1 + \alpha_1 t + \alpha_2 t^2\right)$.

[c]In the learning-by-doing model, the human capital production function is $H_{t+1} = (1-\delta) H_t + (1-\ell_t) \xi_t \pi \left(1 + \alpha_1 t + \alpha_2 t^2\right)$.

[d]"HC" stands for "Human Capital."

Table 5: Elasticities at selected ages, responding to % changes in wages

	Baseline Model		Exogenous Model		Learning-by-doing Model	
Age	Marshallian[a]	IES[b]	Marshallian[a]	IES[b]	Marshallian[a]	IES[b]
20	0.271	0.243	12.200	13.109	0.306	0.287
25	0.201	0.178	1.435	1.731	0.239	0.230
30	0.207	0.193	1.032	1.155	0.208	0.187
35	0.197	0.177	0.950	1.044	0.231	0.212
40	0.198	0.179	0.954	1.010	0.275	0.249
45	0.234	0.212	1.093	1.159	0.293	0.258
50	0.329	0.303	1.374	1.475	0.351	0.312
55	0.475	0.416	1.853	1.964	0.570	0.494
60	0.724	0.592	2.366	2.484	0.995	0.855
65	1.154	0.951	3.244	3.432	1.676	1.501
70	1.719	1.424	4.049	4.229	2.262	1.880

[a]The Marshallian is $me' = \frac{\log(h_t^t)-\log(h_t^b)}{\log(w_t^t)-\log(w_t^b)}$.

[b]The IES is $ies' = \frac{\log(h_t^t/h_{t-1}^t)-\log(h_t^b/h_{t-1}^b)}{\log(w_t^t/w_{t-1}^t)-\log(w_t^b/w_{t-1}^b)}$

difference between human capital and observed wages.

7 Changes in Tax and Social Security

The preceding sections show that the model fits the life-cycle profiles of labor supply and log wages in the data well. In this section, we use the model to predict how changes in the Tax or Social Security rules would affect behavior in labor supply, human capital investment and the resulting log wage profile. We conduct seven counterfactual policy experiments which reflect various changes in the tax codes and Social Security rules. The results of these experiments are summarized in columns 2-8 in Table 6, where the first column is the baseline model. All numbers are summations throughout the life-cycle (from age 18 to 80).

7.1 The Baseline Model

The first experiment increases the income tax proportionally by 50%. Column 2 shows that after the tax increase, an average individual works additional 1.25 years over the life-cycle, equivalent to 3.1% of the total labor supply. Most of the increase in the labor supply is allocated to effective labor, which increases by 1.19 years. Investment also increases by 0.06 years or 2.6%, which leads to a 2.8% increase in the human capital level and 0.47%

Table 6: Effects of changing taxes or Social Security rules

	1	2		3		4		5		6		7		8	
	Baseline	Tax Increase 50%		No Earnings Test		NRA = 67		Reduce SSB 20%		No SS Taxes		No SS Benefit		No SS System	
	Level[a]	ΔLevel[b]	%Δ[c]	ΔLevel	%Δ	ΔLevel	%Δ	ΔLevel	%Δ	ΔLevel	%Δ	ΔLevel	%Δ	ΔLevel	%Δ
Panel A: Baseline Model															
LFPR	40.356	1.249	**3.096**	0.387	**0.959**	0.409	**1.012**	0.633	**1.570**	-1.947	**-4.826**	5.220	**12.934**	2.647	**6.560**
Effective Labor	37.997	1.187	**3.125**	0.364	**0.957**	0.382	**1.006**	0.594	**1.563**	-1.820	**-4.790**	4.911	**12.926**	2.502	**6.585**
Pre-tax Income	637.759	28.831	**4.521**	5.842	**0.916**	8.484	**1.330**	13.805	**2.165**	-45.555	**-7.143**	97.452	**15.280**	37.153	**5.826**
Average lnw	2.613	0.012	**0.468**	0.008	**0.322**	0.007	**0.268**	0.015	**0.569**	-0.020	**-0.777**	**0.050**	**1.903**	0.021	**0.792**
Human Capital	917.382	25.901	**2.823**	6.465	**0.705**	7.665	**0.836**	12.160	**1.326**	-39.947	**-4.354**	82.071	**8.946**	33.769	**3.681**
Investment	2.359	0.062	**2.631**	0.023	**0.988**	0.026	**1.122**	0.040	**1.684**	-0.127	**-5.395**	0.308	**13.059**	0.145	**6.147**
Panel B: Exogenous Model															
LFPR	40.412	1.524	**3.771**	0.084	**0.208**	0.338	**0.835**	0.530	**1.311**	-2.213	**-5.475**	3.456	**8.552**	0.907	**2.245**
Effective Labor	40.412	1.524	**3.771**	0.084	**0.208**	0.338	**0.835**	0.530	**1.311**	-2.213	**-5.475**	3.456	**8.552**	0.907	**2.245**
Pre-tax Income	655.258	25.849	**3.945**	1.794	**0.274**	5.235	**0.799**	8.285	**1.264**	-33.535	**-5.118**	50.490	**7.705**	13.300	**2.030**
Average lnw	2.625	0.006	**0.248**	0.000	**0.015**	0.001	**0.024**	0.002	**0.076**	-0.003	**-0.103**	0.000	**-0.009**	-0.002	**-0.069**
Panel C: Learning-by-doing Model															
LFPR	40.232	1.287	**3.198**	0.257	**0.638**	0.821	**2.040**	1.186	**2.948**	-3.474	**-8.635**	6.368	**15.827**	3.482	**8.655**
Effective Labor	40.232	1.287	**3.198**	0.257	**0.638**	0.821	**2.040**	1.186	**2.948**	-3.474	**-8.635**	6.368	**15.827**	3.482	**8.655**
Labor Income	629.286	28.301	**4.497**	3.825	**0.608**	13.293	**2.112**	19.829	**3.151**	-60.550	**-9.622**	112.483	**17.875**	49.189	**7.817**
Average lnw	2.572	0.028	**1.075**	0.014	**0.551**	0.016	**0.620**	0.026	**1.013**	-0.181	**-7.026**	0.093	**3.602**	0.064	**2.489**
Human Capital	844.058	23.218	**2.751**	4.107	**0.487**	12.131	**1.437**	16.924	**2.005**	-48.275	**-5.719**	84.723	**10.038**	41.654	**4.935**
Panel D: Model with Health															
LFPR	41.926	1.065	**2.541**	0.191	**0.456**	0.379	**0.904**	0.607	**1.448**	-1.727	**-4.118**	4.275	**10.198**	2.162	**5.156**
Effective Labor	39.625	1.026	**2.589**	0.180	**0.454**	0.355	**0.897**	0.571	**1.440**	-1.621	**-4.090**	4.046	**10.212**	2.060	**5.199**
Pre-tax Income	653.149	25.507	**3.905**	3.179	**0.487**	7.924	**1.213**	12.872	**1.971**	-41.183	**-6.305**	81.956	**12.548**	30.333	**4.644**
Average lnw	2.603	0.014	**0.531**	0.004	**0.143**	0.006	**0.215**	0.009	**0.348**	-0.023	**-0.867**	0.034	**1.296**	0.006	**0.213**
Human Capital	914.487	22.221	**2.430**	3.606	**0.394**	7.127	**0.779**	11.218	**1.227**	-34.736	**-3.798**	67.618	**7.394**	27.030	**2.956**
Investment	2.301	0.039	**1.715**	0.011	**0.487**	0.024	**1.031**	0.036	**1.581**	-0.106	**-4.599**	0.229	**9.958**	0.101	**4.408**

[a]The "Level" column refers to the total value aggregated over the whole life-cycle, except the "Average lnw" which is the average yearly log wages. For example, in the baseline model, the total LFPR is 40.356 years from 18 to 80.
[b]The "ΔLevel" column refers to the difference of the total value between the current experiment and the baseline model. For example, in the "No Earnings Test" case, the LFPR is 0.387 years higher than that in the baseline model across the whole life-cycle from 18 to 80.
[c]The "%Δ" column refers to the percentage of the difference in the "ΔLevel" column relative to the level in the baseline model. For example, in the "No Earnings Test" case, the LFPR increases by 0.387 years which is equivalent to 0.959% of the LFPR in the baseline model.

increase in the observed log wages.[25] A tax hike has both substitution and income effects. The substitution effect discourages labor supply while the income effect encourages labor supply. Our first experiment indicates that in our model the income effect dominates the substitution effect and this is the case with most of our experiments.[26] We also see that human capital investment increases in this experiment. The direct effect of taxes discourages human capital investment, but the increase in labor supply (and in particular delayed retirement) increases human capital investment.

The manner in which Social Security rules affect labor supply and wages is of central interest to policy makers. The six experiments in columns 3-8 are devoted to answering these questions. In the first three we manipulate the current Social Security rules (columns 3-5) while in the last three we decompose the distortionary effects of the current Social Security system (columns 6-8).

First we remove the Social Security earnings test, which is effective between age 62 and 70 in the baseline model. In the second experiment, we delay Normal Retirement Age (NRA) by two years: the new NRA is age 67 in this counterfactual experiments while it is age 65 in the baseline model. In the third one, we reduce the Social Security benefit proportionally by 20%. The results are presented in columns 3-5 in Table 6. Removing the Social Security earnings test between ages 62 and 70 has a smallest effect on all variables; delaying the normal retirement age by two years, has a slightly larger impact; reducing the generosity of the social security benefit has the largest effect among these three.[27] For instance, they increase the labor force participation by four-and-a-half, five, or seven-and-a-half months, respectively. One important feature is that the change in the labor supply does not only happen later in the life-cycle when the policy change is directly effective, it takes place over the whole life-cycle, as indicated in Figure 8a. When the NRA is delayed two years or the Social Security benefit is reduced, workers also invest more and therefore have higher human capital levels, which leads to higher wages at old ages (Figure 8a). The wage difference is negligible before age 60 but increases substantially after that, reaching 3% or 5% around age 67. Ignoring such wage response in experiments involving retirement policy will most likely introduce bias.

[25]Other papers have looked at the effects of taxes and human capital with this type of model. Examples are Heckman et al. (1998b), Heckman et al. (1999), and Taber (2002). These experiments are quite different as labor supply makes a large difference here so the results are not directly comparable.

[26]We also experimented by simulating with a lower coefficient of risk aversion of $\eta_c = 2$ rather than four. In this case the income effect still is larger than the substitution effect, but it is closer. For example in the 50% tax increase labor supply increases by 0.7% rather than 3.1%.

[27]The benefit withdrawn by the earnings test is paid back later in the form of Delayed Retirement Credit (DRC). Therefore the net effect of Social Security earnings test is not clear and depends on the life expectancy which affects the actuarial fairness.

In the last three experiments, we decompose the effect of the current U.S. Social Security system into the individual effects of the Social Security taxes and the Social Security benefit. In Column 6 we keep the Social Security benefit but eliminate the Social Security taxes (the payroll taxes);[28] in Column 7 we remove the Social Security benefit completely but keep the Social Security taxes; in Column 8 we remove the entire Social Security system, that is, both Social Security taxes and the benefit. Removing Social Security taxes induces an average individual to supply 1.95 years less labor. This is not surprising because removing the Social Security taxes is essentially a universal cut in the tax rate. In our tax hike counterfactual, the income effect dominates the substitution effect as is true for the cut in social security taxes as well. Analogously, removing the Social Security benefit induces more labor supply. However, the increase in the labor supply is 5.22 years, which is much higher than 1.95 years reduction of labor supply in the case of removing Social Security Taxes. The combination of these two effects leads to the results in the last experiment where both the Social Security taxes and benefit are removed. Column 8 indicates that eliminating the current Social Security system increases average labor supply by 2.65 years over the life-cycle. Such an observation is also mentioned qualitatively in Gustman and Steinmeier (1986) and Rust and Phelan (1997). Figures 8a and 8b show that the increases in the labor supply and log wages are most pronounced at old ages in the experiment without Social Security system.

Another point worth emphasizing is that, in almost every policy counterfactual,[29] the increase in the endogenously determined wage levels are substantial, especially at old ages: 6% when removing earnings test or reducing Social Security benefit, 3% when delaying NRA by two years, and over 10% when removing Social Security benefit or the entire system. These are caused by increases in the human capital levels as a result of higher investment. For this reason, it is likely that ignoring human capital investment channel will generate a bias in terms of predicting LFPR at old ages in similar experiments.

7.2 Sensitivity to Alternative Models

Table 6 also presents the results of experiments from the alternative models, specifically, Panel B from the exogenous model and Panel C from the learning-by-doing model.

Compared with the baseline model, the labor supply response to the policy changes are smaller in most experiments when the human capital is exogenous (labeled as exogenous model in Panel B), but are larger in most experiments in the learning-by-doing

[28] The income taxes are still effective.

[29] The only exception is the experiment of removing Social Security taxes.

model (Panel C). [30] This result comes from several different features of these three models. Consider the experiment that reduces the Social Security benefit by 20% (Column 5) as an example. The change in labor supply is essentially purely due to the income effect. We see the largest effect on labor supply in the learning by doing model, the second highest in the baseline model, and the lowest in the exogenous model. This is quite surprising because the labor supply elasticity is highest in the exogenous model. The key to understanding this effect is human capital. When the Social Security benefit is reduced, the reduction in the expected wealth induces higher labor force participation particularly for older workers. In the two human capital models this "delayed retirement" increases the expected return to human capital investment, which in turn induces higher participation at earlier ages. This "adjacent complementarity" channel does not operate for the exogenous model, so the change in labor supply leads to a larger response in the learning-by-doing model than in the exogenous model. The response is lower in the baseline model than in the learning by doing model because the baseline model gives workers have an extra channel for adjustment—the allocation of time between investment and working. This enables workers to react to the increased return to human capital more efficiently. In the learning by doing model the only adjustment is through the extensive labor supply channel.

In Panel D, we present the results for the model with health as described in Subsection 5.2. In most cases, the results are reasonably close to those in the baseline model.

8 Robustness Check

Recall that some of parameters are set to certain values taken from the previous literature. In this section we vary those pre-set parameters to see how they affect our estimation results. In particular, we check following variants: (1) increase the consumption floor $\underline{c}$ from 2.19 to 2.5; (2) decrease the consumption floor $\underline{c}$ from 2.19 to 1.8; (3) decrease the time discount rate δ from 0.97 to 0.96 but increase the interest r from 0.03 to 0.04; (4) decrease the time discount rate δ from 0.97 to 0.95; (5) increase the initial asset A_0 from 0.0 to $50,000$; (6) decrease the CRRA coefficient from 4.0 to 2.0. In each case, all other pre-set parameters are kept the same as the baseline model, and then we re-run the whole estimation to obtain the new estimates of parameters of interest. The estimation results are

[30] Two exceptions are the experiment with tax increase (Column 2) and the one without Social Security taxes (Column 6). In these two experiments, the responses in the exogenous model are larger in magnitude than those in the baseline model. This is because taxes changes are essentially equivalent to changes in the human capital rental rate. As we show in Figure 7b, the IES is universally higher in the exogenous model, which implies larger labor supply responses.

listed in Table 7, and the moments are plotted in Figures 9a-9d.

In most cases the simulated moments fit the data moments quite well. The one exception is that when we decrease CRRA the simulated observed log wages are higher than the data before age 45 and lower after age 45. However in that case, the simulated log wages after filtering out the individual fixed effects almost replicate the data profile. In this case, the simulation does not replicate the increase in labor supply at young ages either.

Interestingly the model with larger initial assets fits the data better than our baseline model. In particular it fits the increase in labor supply early in the life-cycle considerably better. For ages 22-25 one can see in Figure 9a that the fit is almost perfect while in Figure 3a we understate the increase in labor supply. This is perhaps not surprising—in our model workers are borrowing constrained so they work more at young ages as a result. Increasing the initial asset level essentially relaxes this borrowing constraint.

In sum, varying pre-set parameters does change the estimated values of some parameters, but in all variants our model generates simulated moments which match data moments quite well.

9 Conclusion

This paper develops and estimates a rich life-cycle model that merges a Ben-Porath style human capital framework with a Neoclassical style framework with endogenous labor supply and retirement. In the model, each individual makes decisions on consumption, human capital investment, labor supply and retirement. Investment in human capital generates wage growth over the life-cycle, while depreciation of human capital is the main force generating retirement. We show that the parsimonious model is able to fit the main features of life-cycle labor supply, wages (with and without fixed effects) as well as retirement. In particular we can fit both the large increase in wages and small changes in labor supply at the beginning of the life-cycle along with the small changes in wages but large changes in labor supply at the end. We incorporate health into the model and show that while this is an important factor, human capital remains the main explanation for the decline in labor supply for older workers.

Despite the fact that our framework does not rely on age and time varying preference or production function parameters, our model implies a rather small and empirically plausible Marshallian elasticity which rises with age. We also estimate the same basic framework using two different approaches to human capital accumulation—exogenous human capital as well as learning by doing. We find that the baseline model is better

Table 7: Estimates in the baseline model and variants[a]

MODEL SPECIFICATIONS		Baseline	1 Larger $\underline{c}$	2 Lower $\underline{c}$	3 Change δ, r	4 Smaller δ	5 Larger A_0	6 Smaller η_c
Interest rate	r	0.03			0.04			
Discount	δ	0.97			0.96	0.95		
CRRA	η_c	4.0						2.0
Initial wealth	A_0	0.0					50,000	
Consumption floor	$\underline{c}$	2.19	2.5	1.8				
Leisure: St. Dev. of Shock	a_ε	0.433	0.431	0.432	0.396	0.209	0.427	0.161
HC Depreciation[b]	σ	0.101	0.101	0.101	0.101	0.101	0.102	0.101
HC Prod. : I factor	α_I	0.076	0.079	0.075	0.057	0.048	0.076	0.108
HC Prod. :H factor	α_H	0.151	0.153	0.151	0.135	0.158	0.157	0.140
St. Dev. of HC Innovation	σ_ξ	0.405	0.415	0.405	0.555	0.972	0.286	0.290
Bequest weight	b_1	424,070	471,626	453,308	873,311	51	431,857	1,075
Parameter heterogeneity								
Leisure: Mean of Intercept	μ_{a_0}	-6.525	-6.529	-6.519	-6.535	-6.869	-6.856	-2.110
Leisure: St. Dev. of Intercept	σ_{a_0}	0.874	0.864	0.869	0.905	0.717	0.590	1.019
HC Prod., Mean	μ_π	1.758	1.756	1.758	1.759	1.753	1.754	1.750
HC Prod., Stand. Dev.	σ_π	0.583	0.585	0.584	0.576	0.601	0.611	0.623
Correlation between $a_0$and π	ρ	-0.893	-0.894	-0.889	-0.892	-0.255	-0.944	-0.041
Initial HC level at age 18								
Intercept	γ_0	1.625	1.633	1.627	1.583	1.904	1.851	1.708
Coefficient on a_0	γ_{a_0}	0.052	0.053	0.052	0.070	0.044	0.043	0.099
Coefficient on π	γ_π	0.531	0.531	0.529	0.535	0.473	0.415	0.393
St. Dev. of Error Term	$\sigma_{H_{18}}$	0.239	0.240	0.240	0.160	0.260	0.324	0.092

[a]Method of simulated moments estimates. Estimates use a diagonal weighting matrix.
[b]HC stands for "Human Capital."

at replicating the main features of the data. In our baseline model, the level of human capital falls for people working with their wages being flat due to investment on the job. This mechanism which is intrinsic to the Ben-Porath framework is not in play in either the learning by doing framework or the exogenous human capital model and this plays a central role in generating a better fit. Finally, policy simulations reveal that changes in the Normal Retirement Age can have sizable effects on wages and more muted effects on labor force participation rates.

We conduct several robustness checks and one of them is worth highlighting. The baseline estimation starts individuals off with no assets. Re-estimation of the model with an initial asset level of $50,000 results in a near perfect fit especially in regards to labor supply early in the life-cycle. This suggests that ignoring the role played by borrowing constraints improves the ability of the model to fit the data.

We have clearly abstracted from many features of labor markets and many can be added to the framework. Two important features of the model are endogenous labor supply and endogenous human capital and our analysis demonstrates that they are inextricably intertwined. Endogenous labor supply is essential for understanding lifecycle human capital investment while lifecycle human capital investment is essential for understanding lifecycle labor supply.

Figure 1a: Labor force participation rate—SIPP data

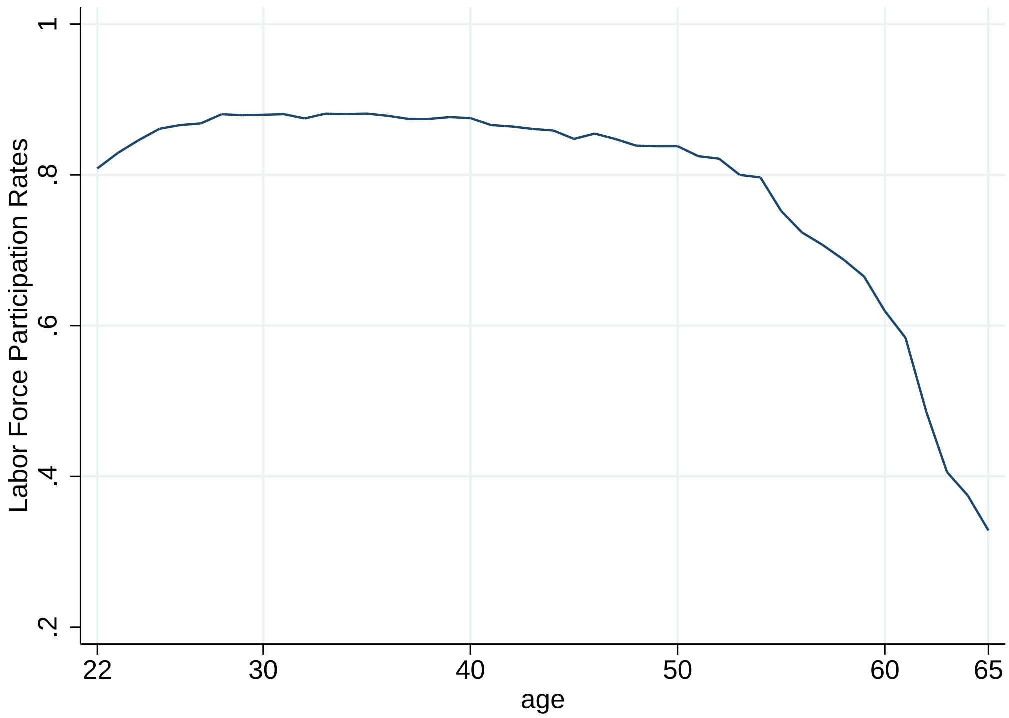


Figure 1b: Log wages—SIPP data

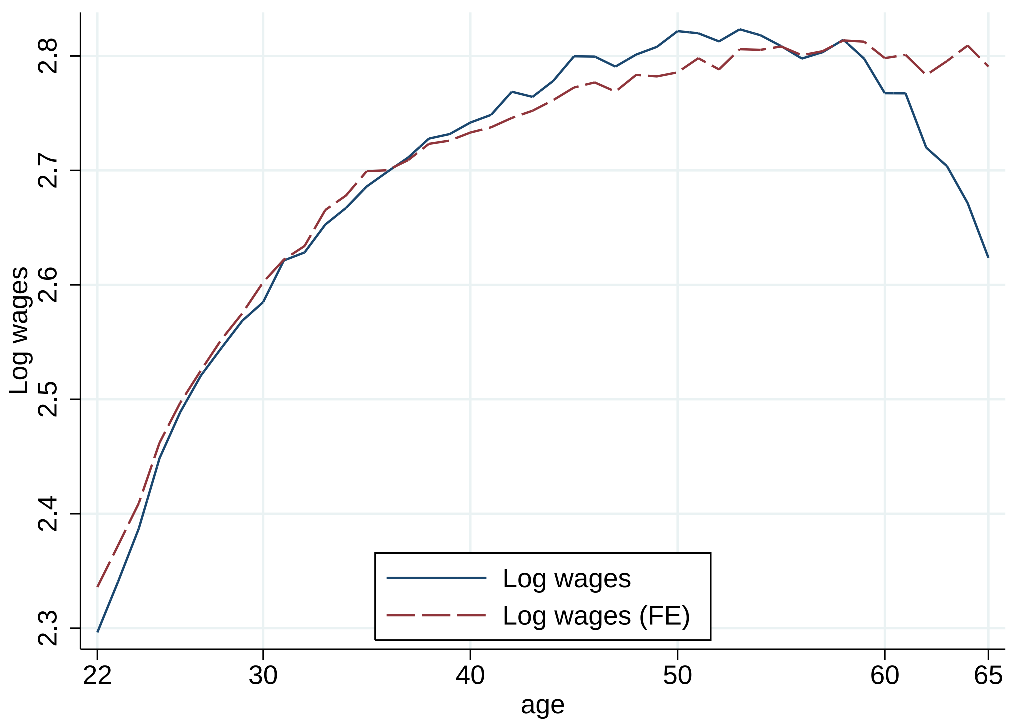


Figure 1c: Standard deviation of log wages—SIPP data

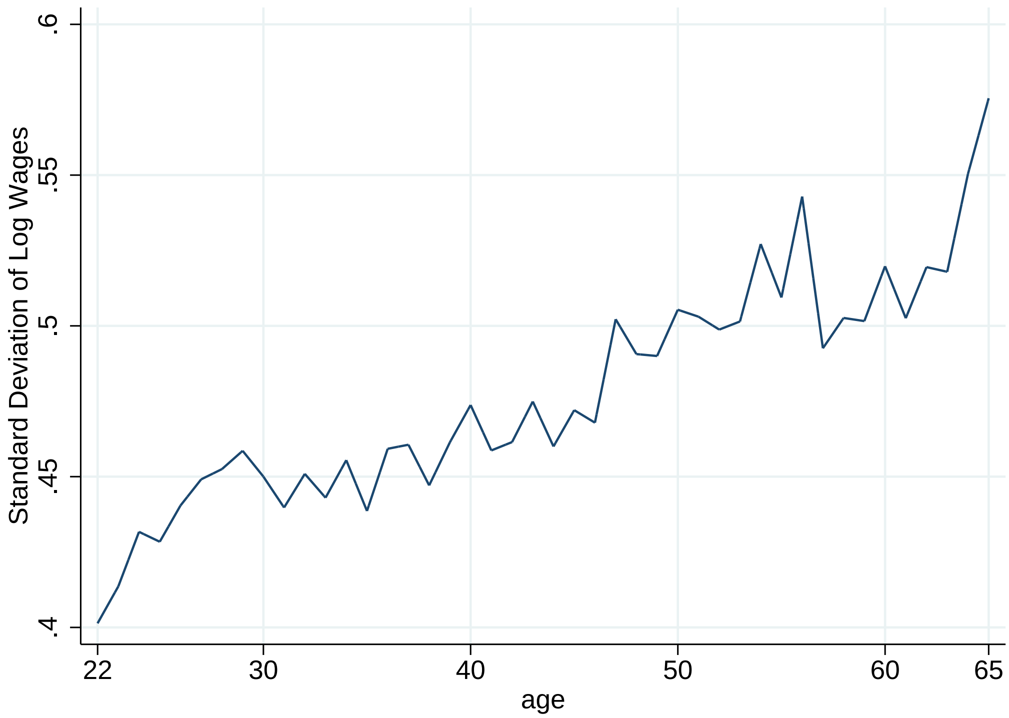


Figure 2: Log wage profiles of male high school graduates, CPS MORG and March data.

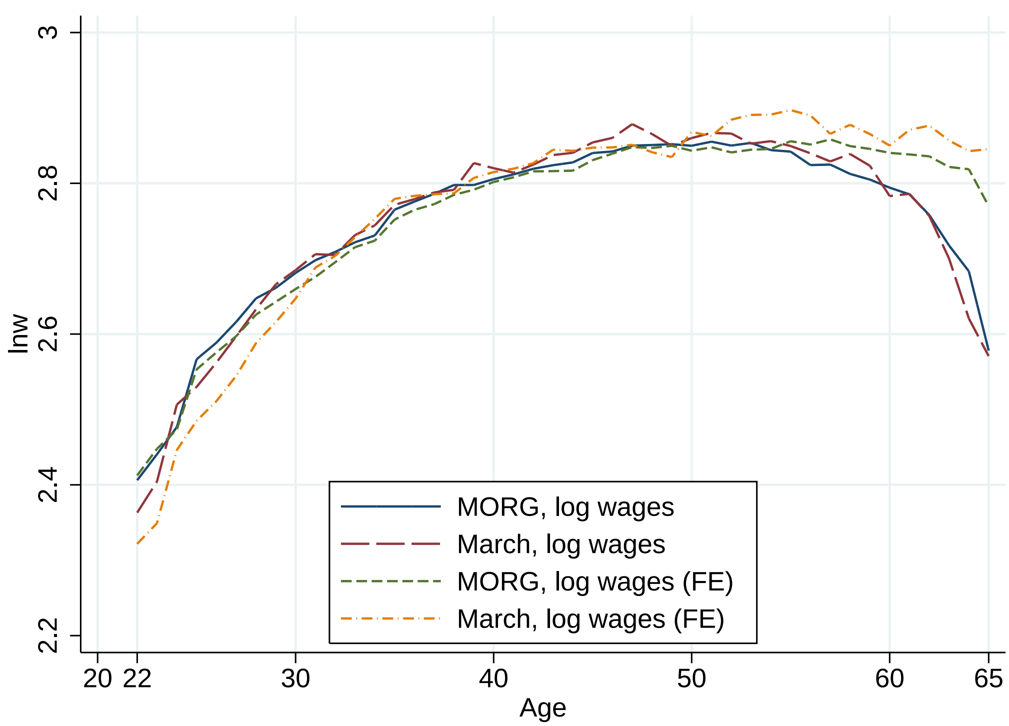


Figure 3a: Fit of model: labor force participation rate

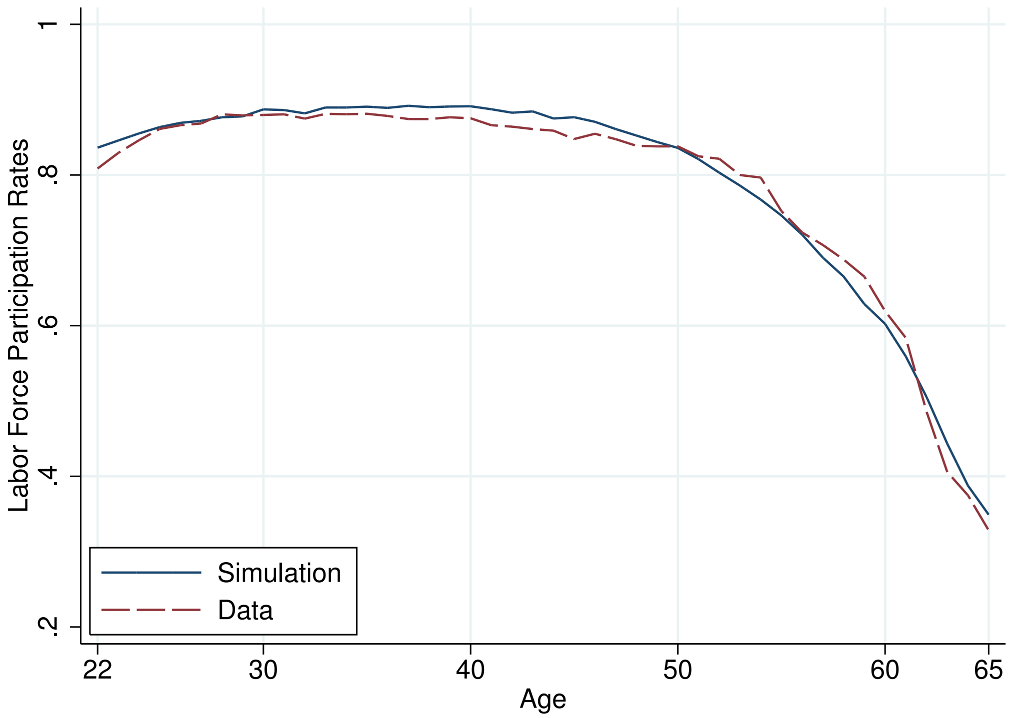


Figure 3b: Fit of model: log wages after controlling for individual fixed effects

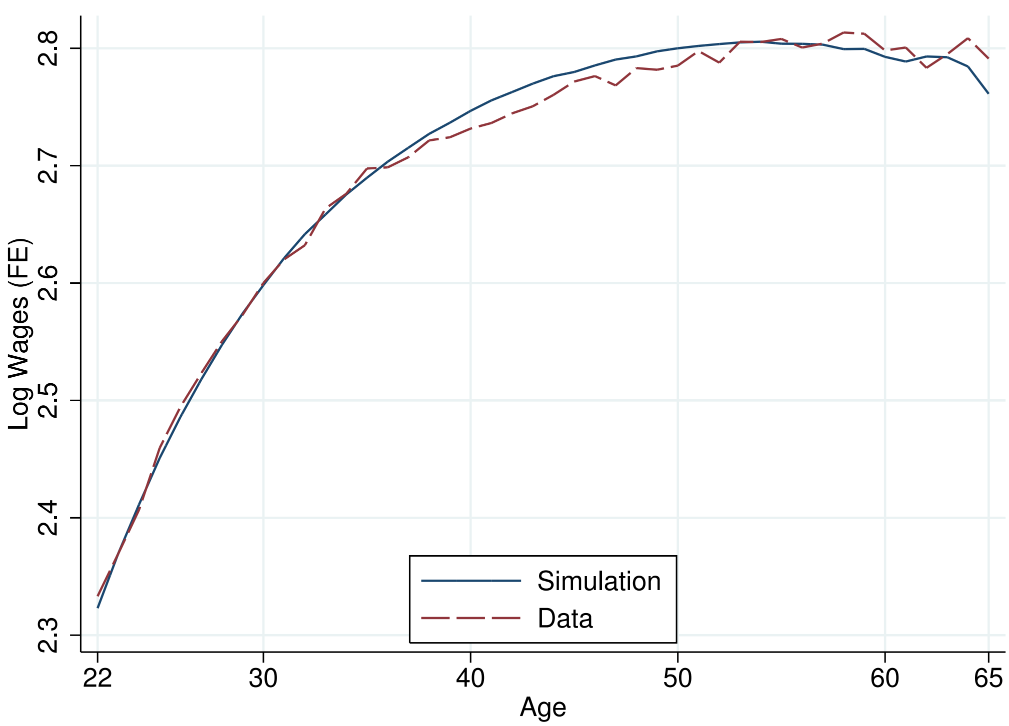


Figure 3c: Fit of model: log wages

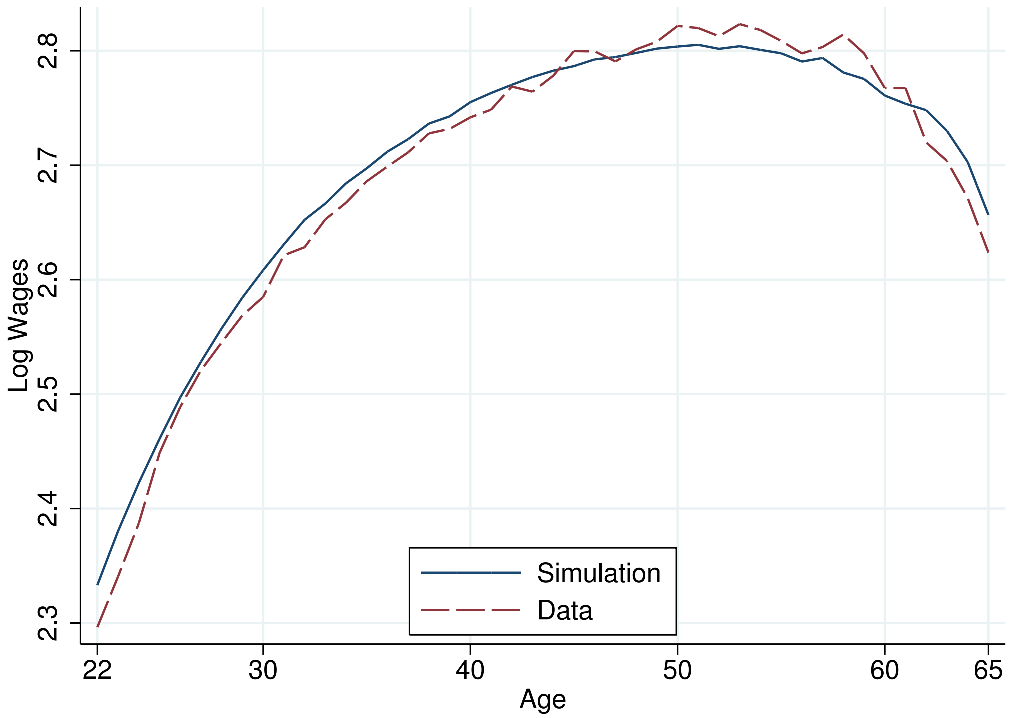


Figure 3d: Fit of model: standard deviations of log wages

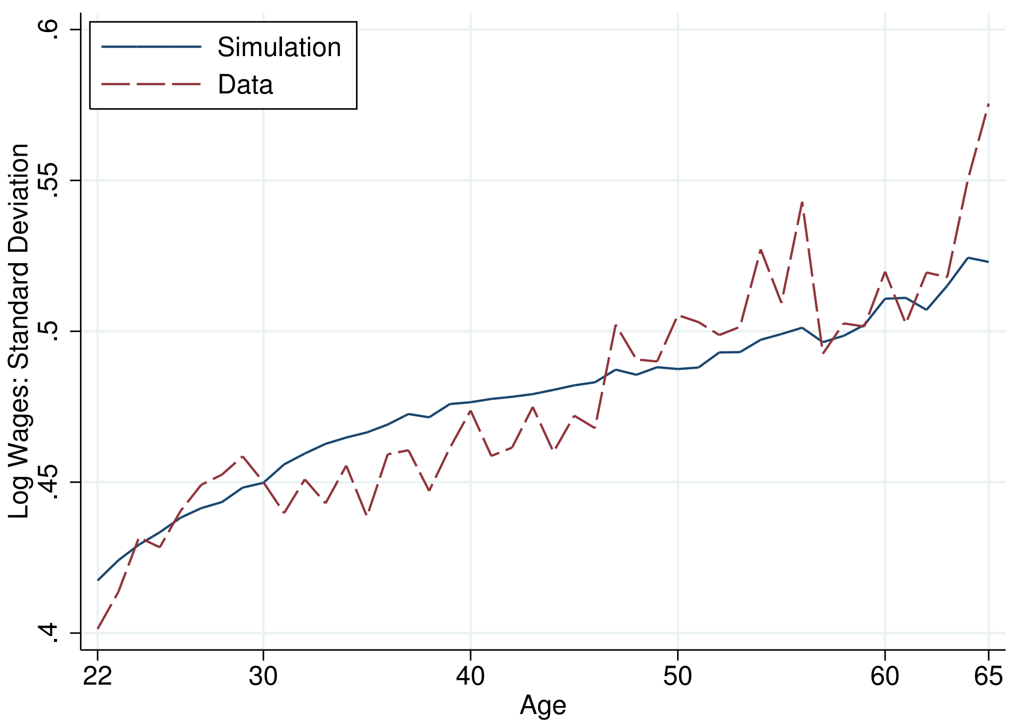


Figure 4a: Log wages and human capital

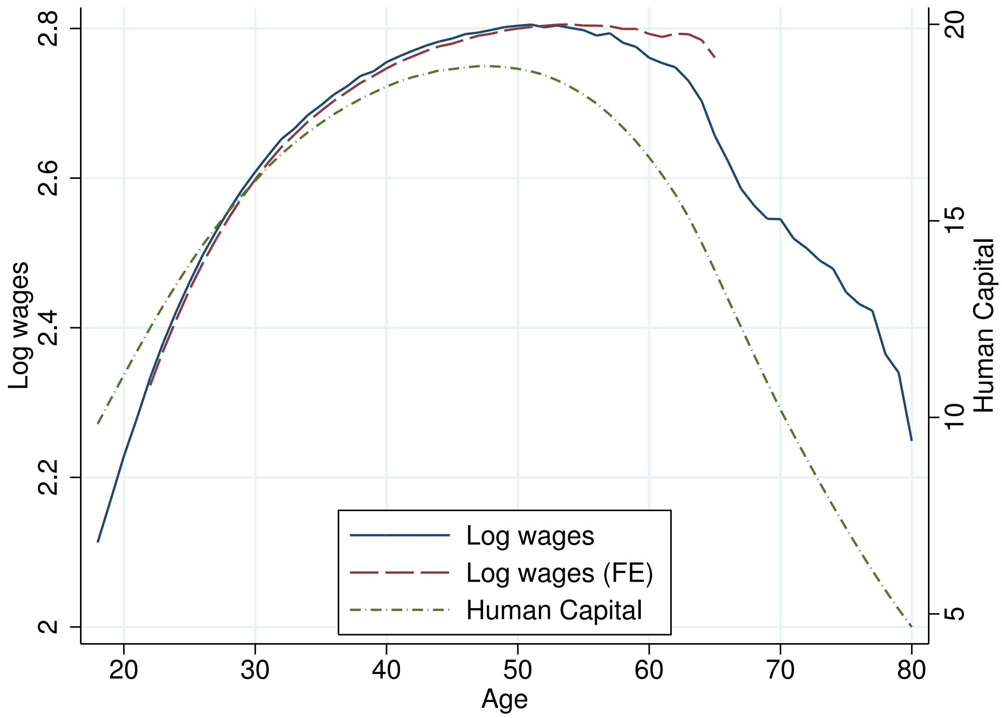


Figure 4b: Investment, and human capital

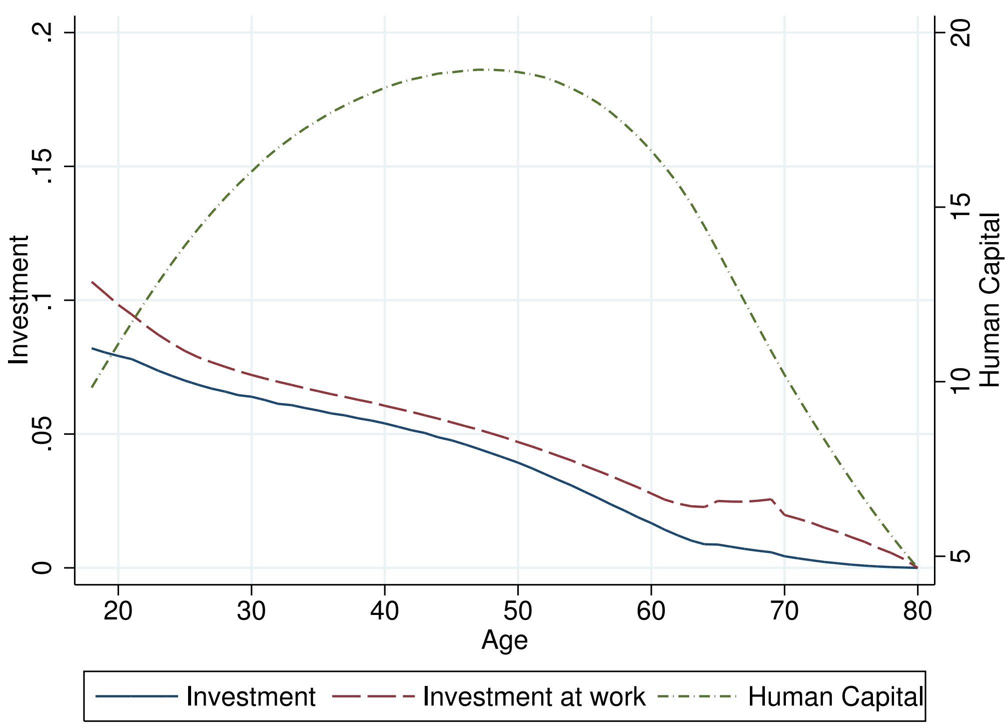


Figure 5a: Calculated elasticities

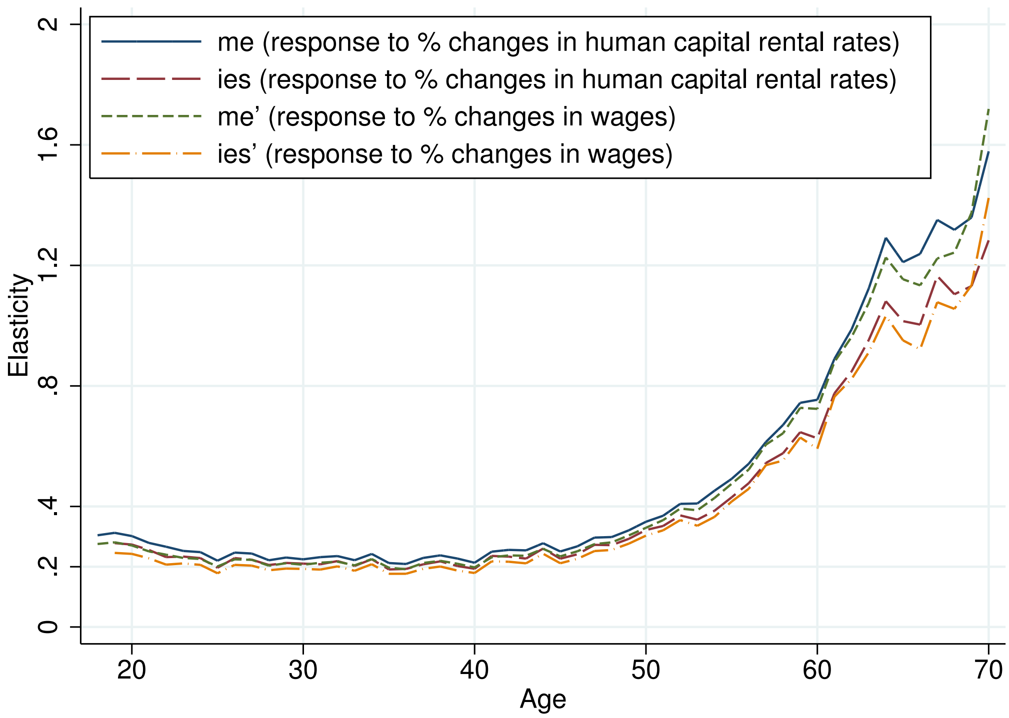
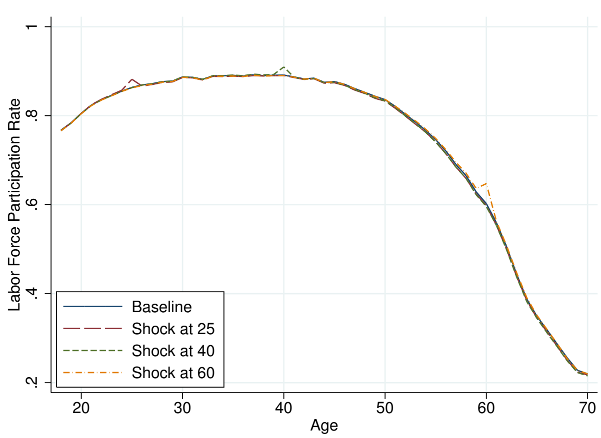


Figure 5b: Labor force participation rates (LFPR) for positive shocks

(i) Shocks at Ages 25, 40 and 60.



(ii) Decomposition of LFPR Profiles.

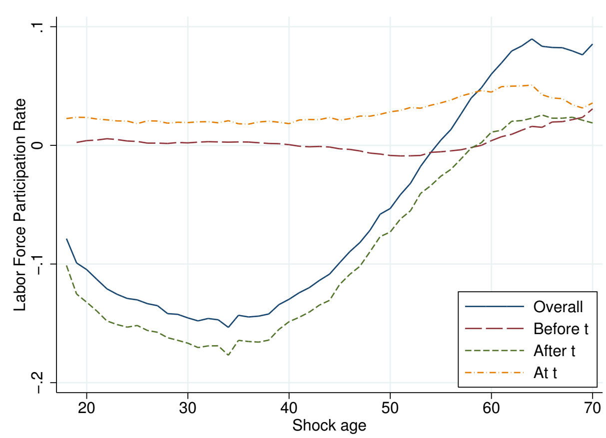


Figure 6a: Fit of model with health

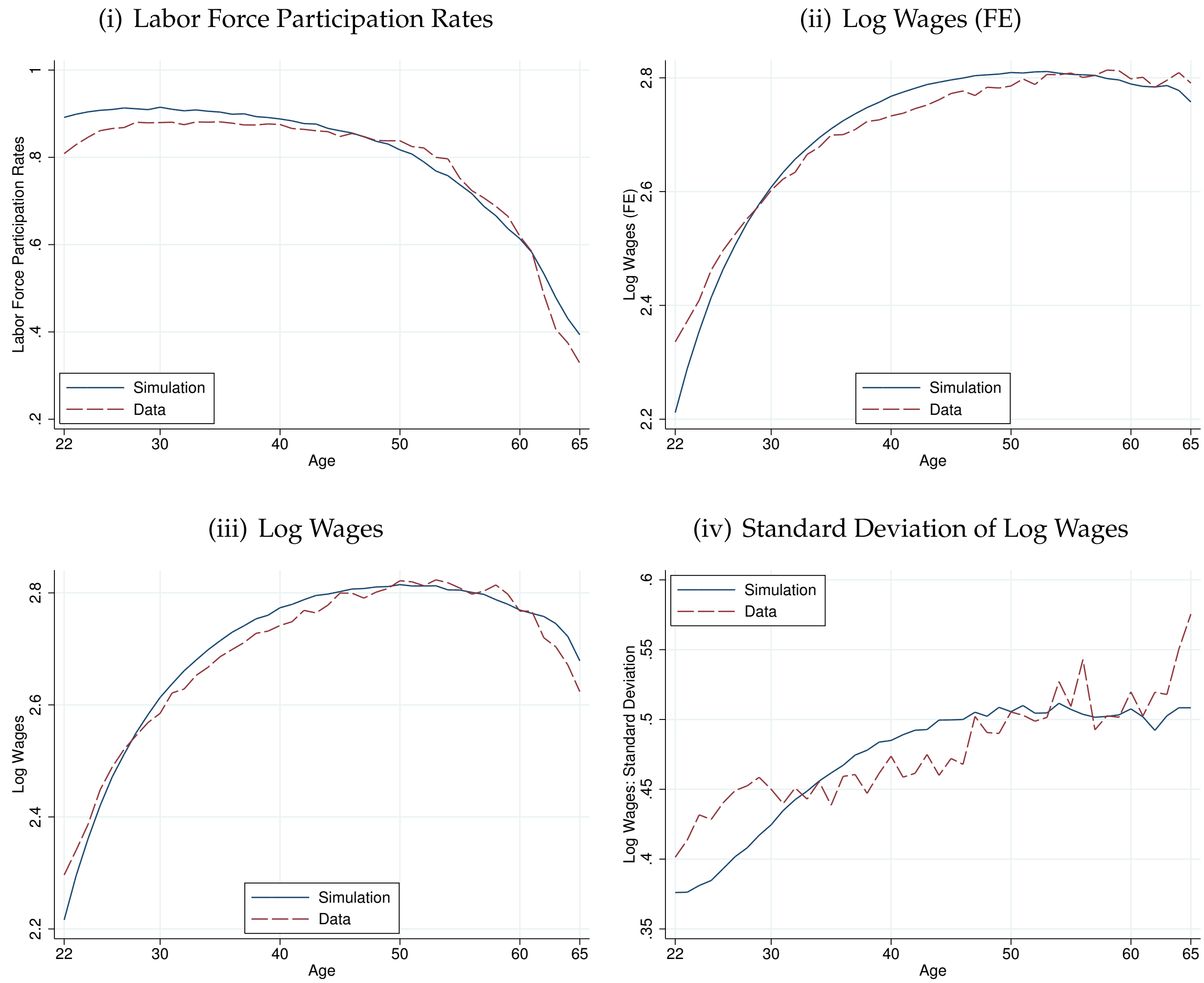


Figure 6b: Sensitivity to heath preferences: health status fixed and taste for leisure unchanged after age 50

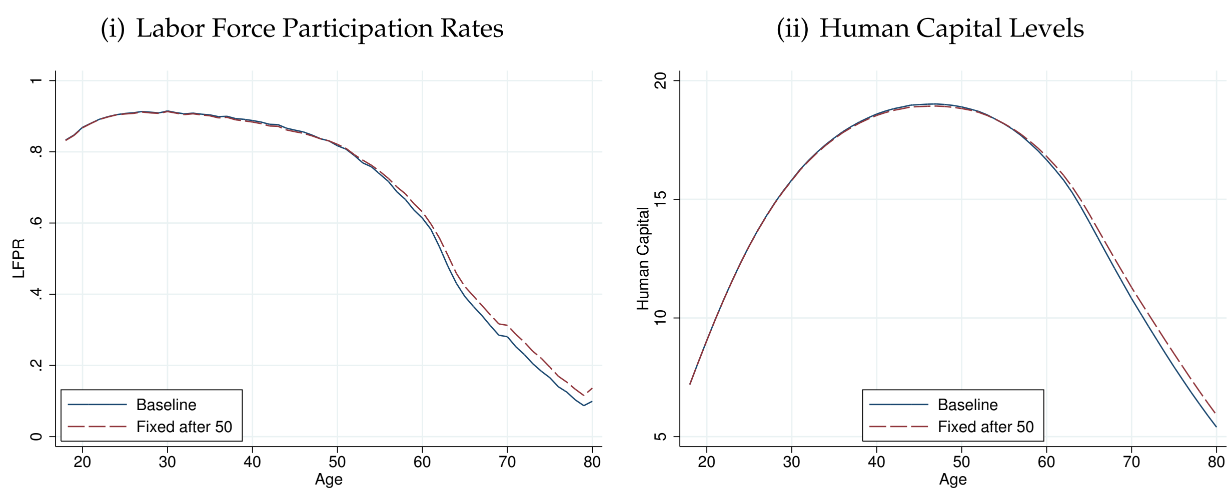
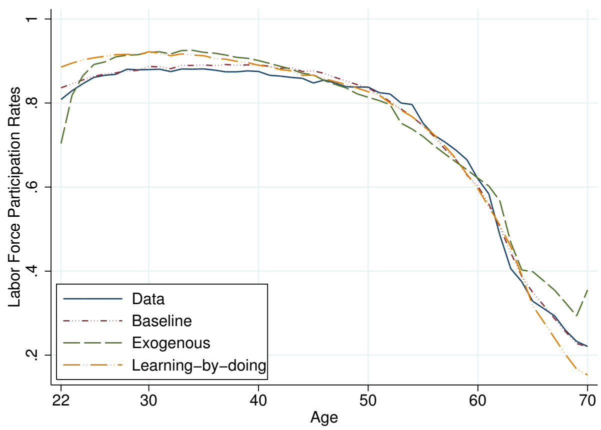
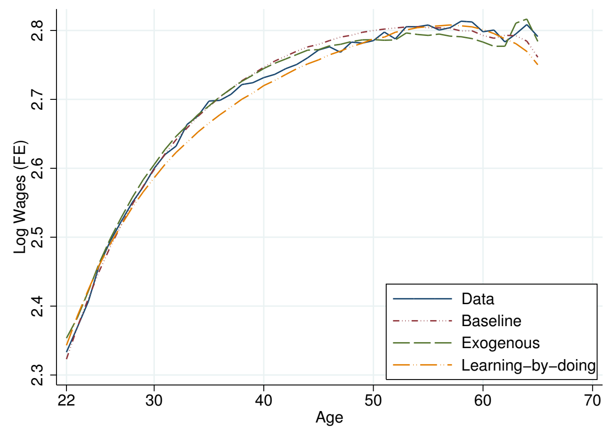


Figure 7a: Exogenous and learning-by-doing models moments

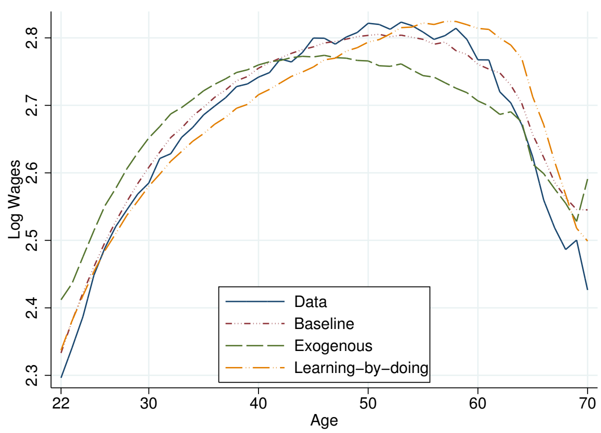
(i) Labor Force Participation Rates



(ii) Log Wages (FE)



(iii) Log Wages



(iv) Standard Deviation of Log Wages

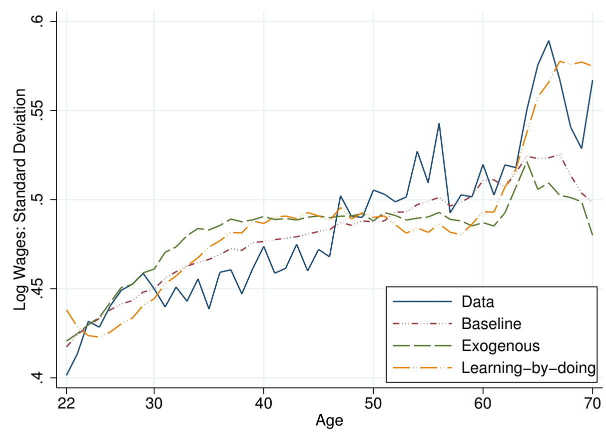
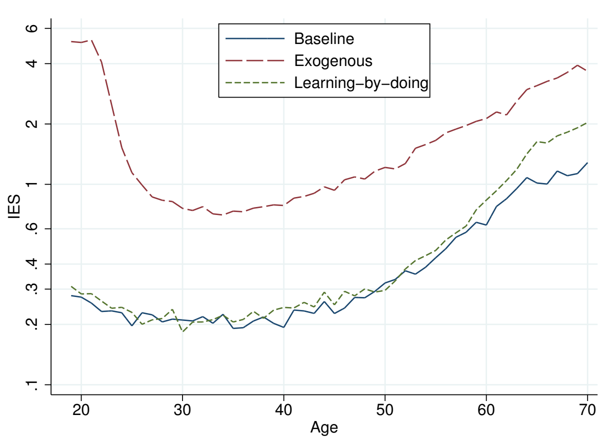


Figure 7b: Comparison of the Intertemporal Elasticities of Substitution (IES).

(i) LFPR Responses to % Changes in H Rental Rate



(ii) LFPR Responses to % Changes in Wages

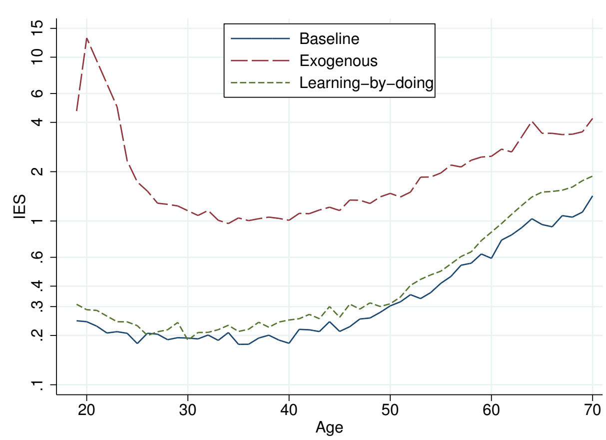
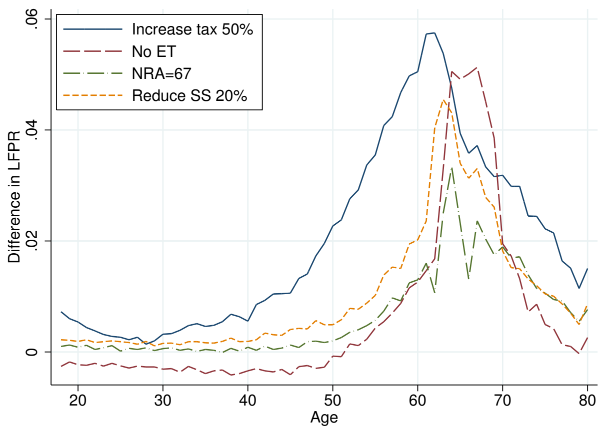
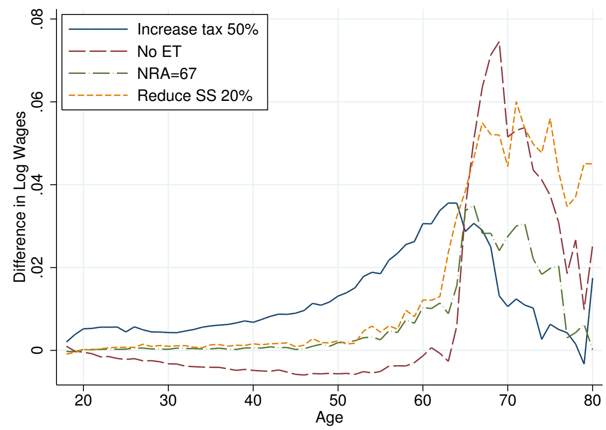


Figure 8a: [Baseline model] Policy experiments: reduce Social Security benefits

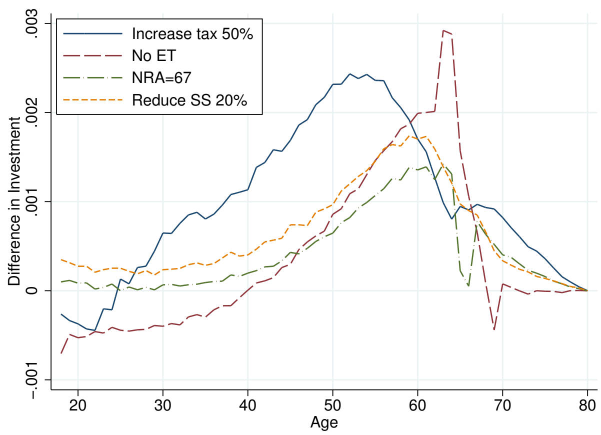
(i) Difference in Labor Force Participation Rates



(ii) Difference in Log Wages



(iii) Difference in Investment



(iv) Difference in Human Capital

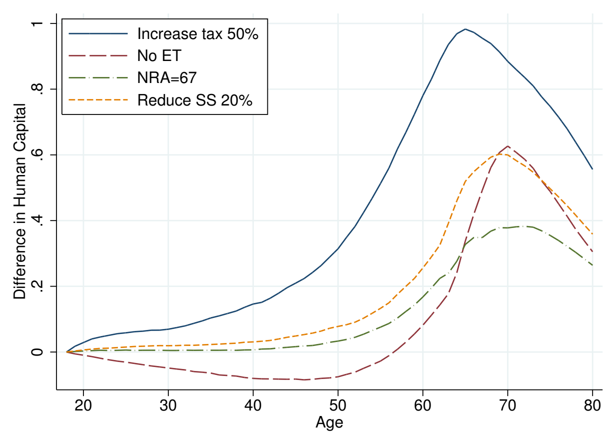
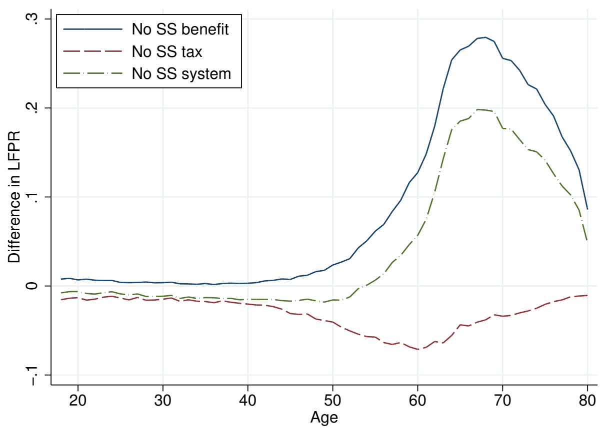
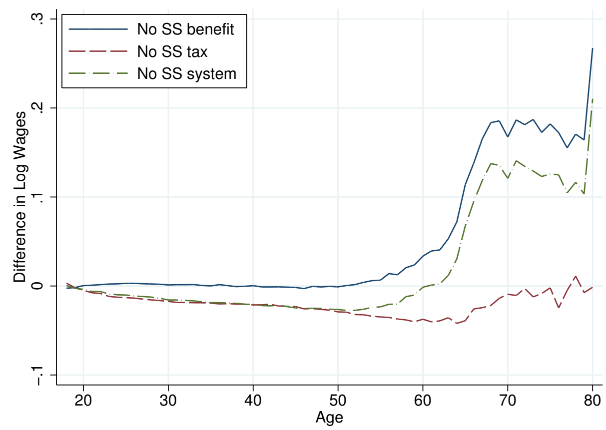


Figure 8b: [Baseline model] Policy experiments: remove Social Security taxes or benefits

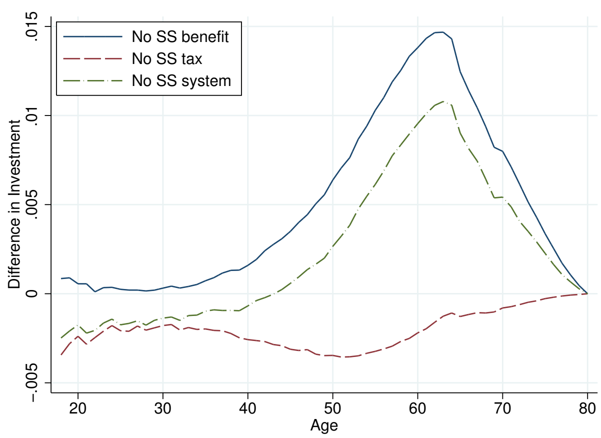
(i) Difference in Labor Force Participation Rates



(ii) Difference in Log Wages



(iii) Difference in Investment



(iv) Difference in Human Capital

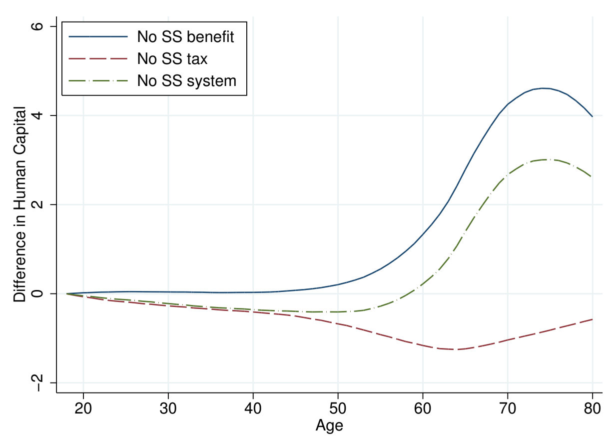


Figure 9a: Fit of alternatives: labor force participation rates

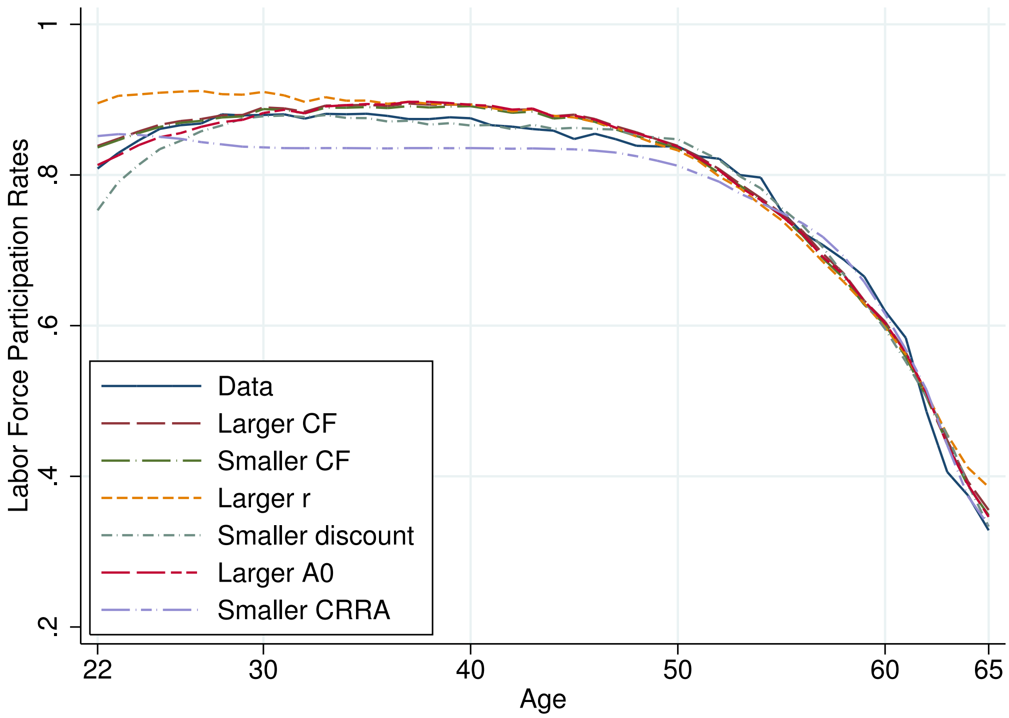


Figure 9b: Fit of alternatives: log wages (FE)

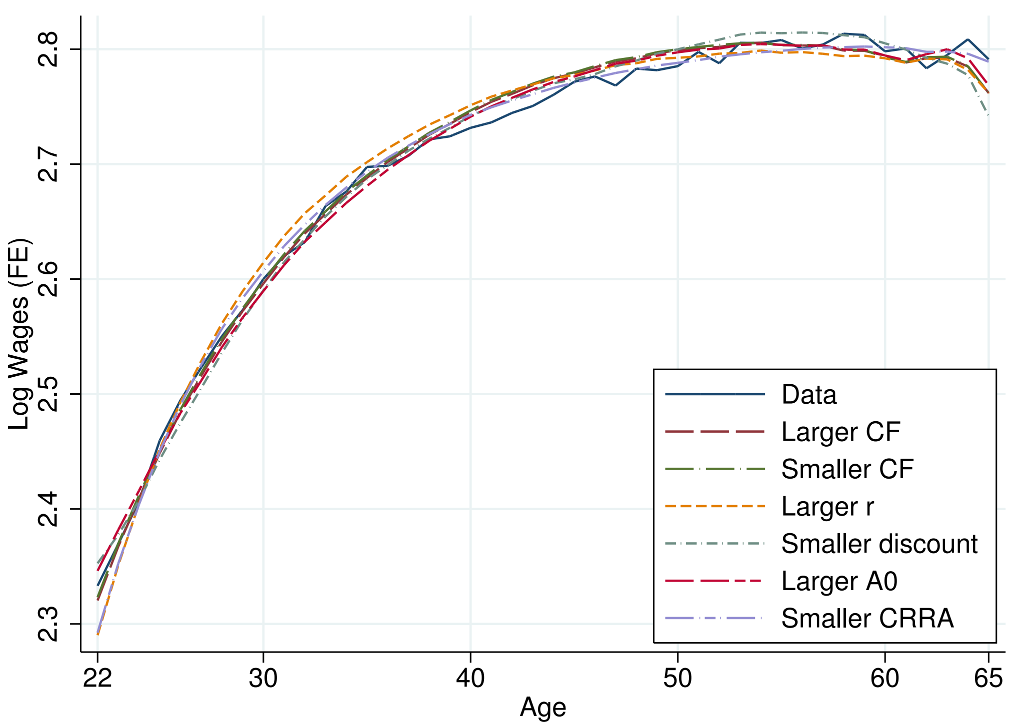


Figure 9c: Fit of alternatives: log wages

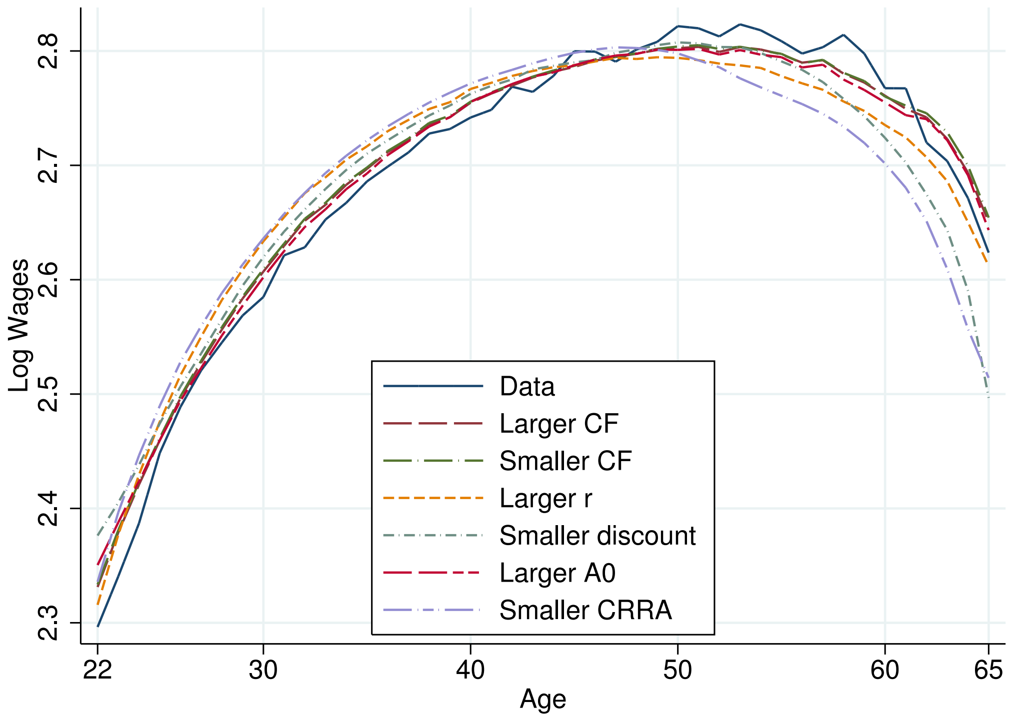
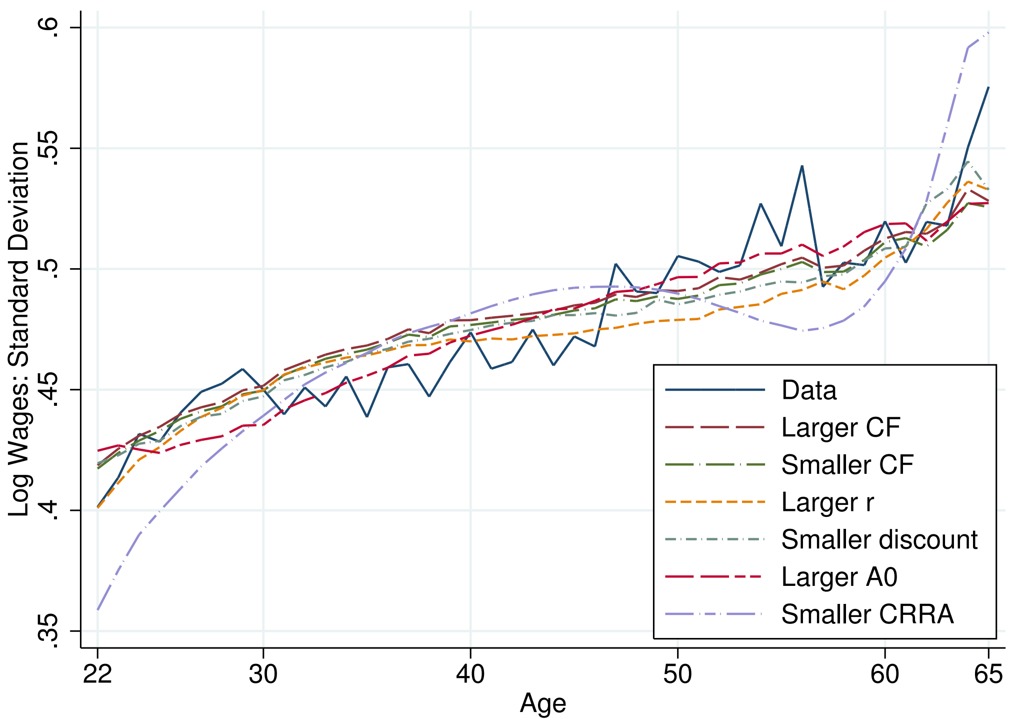


Figure 9d: Fit of alternatives: standard deviations of log wages



References

Altonji, Joseph G, "Intertemporal substitution in labor supply: Evidence from micro data," *The Journal of Political Economy*, 1986, pp. S176–S215.

Ben-Porath, Yoram, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, August 1967, *75* (4), 352–365.

Blau, David, "Retirement and Consumption in a Life Cycle Model," *Journal of Labor Economics*, January 2008, *26* (1), 35–71.

_ **and Tetyana Shvydko**, "Labor Market Rigidities and the Employment Behavior of Older Workers," *Industrial and Labor Relations Review*, April 2011, *64* (3), 464–484.

Browning, Martin, Angus Deaton, and Margaret Irish, "A Profitable Approach to Labor Supply And Commodity Demands over the Life-cycle," *Econometrica*, 1985, pp. 503–543.

_ **, Lars Peter Hansen, and James J. Heckman**, "Micro Data and General Equilibrium Models," *Handbook of Macro Economics*, 1999, *1*, 525–602.

Casanova, Maria, "Happy Together: A Structural Model of Couples' Joint Retirement Choices," November 2010. Working paper.

_ , "Revisiting the Hump-Shaped Wage Profile," August 2013. Working paper.

Chetty, Raj, "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 2012, *80* (3), 969–1018.

Currie, Janet and Brigitte C Madrian, "Health, health insurance and the labor market," *Handbook of labor economics*, 1999, *3*, 3309–3416.

DeNardi, Mariacristina, "Wealth Inequality and Intergenerational Links," *The Review of Economic Studies*, July 2004, *71* (3), 743–768.

Farber, Henry, "The Incidence and Costs of Job Loss: 1982-91," *Brookings Papers: Microeconomics*, 1993, pp. 73–132.

French, Eric, "The Effects of Health, Wealth and Wages on Labor Supply and Retirement Behavior," *Review of Economic Studies*, April 2005, *72* (2), 395–427.

_ **and John Bailey Jones**, "The effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica*, May 2011, *79* (3), 693–732.

Gorlich, Dennis and Andries de Grip, "Human Capital Depreciation During Hometime," *Oxford Economic Papers*, 2009, *61*, i98–i121.

Gustman, Alan L. and Thomas L. Steinmeier, "A Structural Retirement Model," *Econometrica*, May 1986, *54* (3), 555–584.

Haan, Peter and Victoria Prowse, "Longevity, life-cycle behavior and pension reform," *Journal of Econometrics*, 2014, *178*, 582–601.

Haley, William, "Estimation of the Earnings Profile from Optimal Human Capital Accumulation," *Econometrica*, 1976, *44* (6), 1223–1288.

Heckman, James J., *Estimates of a Human Capital Production Function Embedded in a Life-Cycle Model of Labor Supply*, Columbia University Press,

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

_ , Lance J. Lochner, and Petra E. Todd, *Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond*, Vol. 1, Amsterdam: North Holland,

_ , Lance Lochner, and Christopher Taber, "Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents," *Review of Economic Dynamics*, 1998, *1* (1), 1–58.

_ , _ , and _ , "Tax Policy and Human-Capital Formation," *American Economic Review: Papers and Proceedings*, 1998, *88* (2), 293–297.

_ , _ , and _ , "Human Capital Formation and General Equilibrium Treatment Effects: A Study of Tax and Tuition Policy," *Fiscal Studies*, 1999, *20* (1), 25–40.

_ , _ , and Ricardo Cossa, "Learning-by-doing vs. on-the-job training: Using variation induced by the EITC to distinguish between models of skill formation," in Edmund S. Phelps, ed., *Designing inclusion: Tools to raise low-end pay and employment in private enterprise,*, Cambridge University Press 2003, pp. 74–130.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes, "Precautionary Saving and Social Insurance," *The Journal of Political Economy*, April 1995, *103* (2), 360–399.

Imai, Susumu and Michael P. Keane, "Intertemporal Labor Supply and Human Capital Accumlation," *International Economic Review*, May 2004, *45* (2), 601–641.

Johnson, Richard W and David Neumark, "Wage declines among older men," *The Review of Economics and Statistics*, 1996, pp. 740–748.

Keane, Michael P. and Kenneth I. Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, June 1997, *105* (3), 473–522.

Kunze, Astrid, "The Timings of Careers and Human Captial Depreciation," June 2002. IZA Working Paper No. 509.

Kuruscu, Burhanettin, "Training and Lifetime Income," *American Economic Review*, June 2006, *96* (3), 832–846.

Li, Hsueh-Hsiang, "The Effects of Human Capital Depreciation on Occupational Gender Segregation," May 2013.

Light, Audrey and Manuelita Ureta, "Early-Career Work Experience and Gender Wage Differentials," *Journal of Labor Economics*, 1995, *13* (1), 121–154.

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

Madrian, Brigitte C and Lars John Lefgren, "An approach to longitudinally matching Current Population Survey (CPS) respondents," *Journal of Economic and Social Measurement*, 2000, *26* (1), 31–62.

Manuelli, Rodolfo E., Ananth Seshadri, and Yongseok Shin, "Lifetime Labor Supply and Human Capital Investment," January 2012. Working paper.

Mincer, Jacob, "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, August 1958, *66* (4), 281–302.

_ **and Haim Ofek**, "Interuppted Work Careers: Depreciation and Restoration of Human Capital," *The Journal of Human Resources*, 1982, *17* (1), 3–24.

_ **and Soloman Polachek**, "Family Investments in Human Capital: Earnings of Women," *Journal of Political Economy*, 1974, *82* (2), s76–s108.

Neal, Derek, "The Complexity of Job Mobility among Young Men," *Journal of Labor Economics*, April 1999, *17* (2), 237–261.

Prescott, Edward C., Richard Rogerson, and Johanna Wallenius, "Lifetime aggregate labor supply with endogenous workweek length," *Review of Economic Dynamics*, 2009, *12*, 23–36.

Rogerson, Richard and Johanna Wallenius, "Fixed Costs, Retirement and the Elasticity of Labor Supply," October 2010. mimeo, Arizona State University.

_ **and** _ , "Nonconvexities, Retirement and the Elasticity of Labor Supply," *American Economic Review*, June 2013, *103* (4), 1445–1462.

Rosen, Sherwin, "A Theory of Life Earnings," *Journal of Political Economy*, 1976, *84* (4), S45–S67.

Ruhm, Christopher, "Are Workers Permanently Scarred by Job Displacement," *American Economic Review*, 1991, *81* (1), 319–324.

Rupert, Peter and Giulio Zanella, "Revisiting wage, earnings, and hours profiles," June 2012. Working paper.

Rust, John and Christopher Phelan, "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," *Econometrica*, July 1997, *65* (4), 781–831.

Schmeider, Johannes, Till von Wachter, and Stefan Bender, "The Causal Effect of Unemploment Duratoin on Wages: Evidence from Unemployment Insurance Extensions," November 2014.

Shaw, Kathryn L., "Life-Cycle Labor Supply with Human Capital Accumulation," *International Economic Review*, May 1989, *30* (2), 431–456.

Taber, Christopher, "Tax Reform and Human Capital Accumulation: Evidence from an Empirical General Equilibrium Model of Skill Formation," *Advances in Economic Analysis and Policy*, 2002, *2* (1).

Topel, Robert H. and Michael P. Ward, "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, May 1992, *107* (2), 439–479.

Wallenius, Johanna, "Human Capital Accumulation and the Intertemporal Elasticity of Substitution of Labor," September 2009. Working paper.

Weiss, Yoram, "The determination of lifecycle earnings: A survey," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 1, Amsterdam: North-Holland, 1986, pp. 603–640.

Table A1: Wage income tax codes (in 2004$).

Marginal Tax Rate	Pre-tax (Y)	Post-tax Income
0.0765	$\leq 10,250$	$0.9235Y$
0.1765	$10,251 - 20,450$	$9,465.88 + 0.8235\,(Y - 10,250)$
0.2265	$20,451 - 49,150$	$17,865.58 + 0.7735\,(Y - 20,450)$
0.3265	$49,151 - 87,900$	$40,065.03 + 0.6735\,(Y - 49,150)$
0.2645	$87,901 - 110,750$	$66,163.15 + 0.7355\,(Y - 87,900)$
0.2945	$110,751 - 172,950$	$82,969.33 + 0.7055\,(Y - 110,750)$
0.3445	$172,951 - 329,350$	$126,851.43 + 0.6555\,(Y - 172,950)$
0.3645	$\geq 329,351$	$229,371.63 + 0.6355\,(Y - 329,350)$

Appendix

A Taxes

We use taxes codes in the year of 2004. There are two different kinds of taxes that the worker's wage income is subject to, namely the the payroll taxes and the federal income taxes. We ignore the state income taxes. The payroll taxes include the Social Security portion, 6.2% capped at $87,900, and the Medicare tax which is 1.45% and uncapped. The federal income taxes are progressive and we use the tax rules under head of household. The personal exemption for each person is $3,100 and the standard deductions for head of household is $7,150. These all together generate the following tax code used in the paper in Table A1.

B Social Security

We use most Social Security rules in the year of 2004.[31]

B.1 The Social Security Benefits

The normal retirement age (NRA) is 65. The worker receives full Social Security benefits if he applies for the benefits at the NRA. The full retirement benefits equal to the Primary Insurance Amount (PIA), which is a function of Average Indexed Monthly Earn-

[31] Most of information about Social Security benefits in this section are extracted from http://www.ssa.gov.

ings (AIME),

$$\begin{aligned} PIA &= 0.9 * min\left\{bp_1, AIME\right\} + 0.32 * min\left\{bp_2 - bp_1, max\left\{0, AIME - bp_1\right\}\right\} \\ &\quad + 0.15 * max\left\{0, AIME - bp_2\right\}, \end{aligned} \tag{24}$$

where $(bp_1,\ bp_2) = (612,\ 3689)$.

The AIME is computed as the monthly average earning of the 35 years with highest inflation-adjusted earnings. Only earnings subject to the Social Security tax are used in the calculation and therefore AIME is capped. The included earning in a specific year is adjusted for wage inflation by multiplying the wage growth rate relative to the base year, which is at age 60. The wage growth rate is calculated by dividing the average wage in the base year by the average wage in that specific year. Earnings after the base year are not adjusted. Interestingly, the wage growth rate of the national average wage index is very similar to the growth rate of CPI-U after Year 1969, as shown in Figure B1, so we ignore the small difference between these two and use the real wages to update AIME without adjustment.

Computing exact AIME requires keeping tracking of the worker's earning history, which is computationally infeasible. Instead we apply an approximating method, taking into account the wage growth pattern over the life-cycle

$$AIME_{t+1} = AIME_t + max\left\{0,\ \frac{sse_t}{35 \times 12} - share_{min}\left(t\right) \cdot AIME_t\right\} \tag{25}$$

where $sse_t = min\left\{H_t\left(1 - \ell_t\right)\left(1 - I_t\right),\ s\bar{s}e\right\}$ is included earning, capped at $s\bar{s}e = \$87,900$. The $share_{min}$ is the share of minimum wage in AIME. Figure B2 lists the estimated $share_{min}\left(t\right)$ from CPS data for age 52 to 76, assuming the starting working age of 16, and $share_{min}\left(t < 52\right) = 0$.

The early retirement age (ERA) is 62. Starting from ERA, the worker is eligible to receiving the Social Security benefits at a reduced level. In this case, the benefit is reduced 5/9 of one percent for each month before NRA, or 6.67% per year, up to three years. Beyond three years, the benefit is reduced 5/12 of one percent per month or 5% per year.

On the other hand, delayed receiving Social Security benefits after the NRA increases benefits. In this case, the delayed retirement credit (DRC) of 6% is given to the applicant for each delayed year up to age 69.[32] No DRC is given for applicants at age 70 or older.

[32] The 6% DRC is for cohorts born between 1935 and 1936 (inclusive). The DRC varies from 3% for cohorts born in 1924 or earlier to 8% for cohorts born in 1943 or later. In between, it increases by 0.5% every two years.

B.2 The Social Security Earnings Test

We use the Social Security earnings test rules in 1999.[33] The Social Security benefits could be withheld partly or totally if the worker is earning income while taking the Social Security benefits at ages before 70.

For beneficiary under age 65, $1 of benefits for every $2 of earnings in excess of the exempt amount ($10,885 in 2004 dollars) is withheld. The benefit withholding rate for those aged 65-69 is $1 of benefits for every $3 of earnings in excess of the exempt amount ($17,575 in 2004 dollars).

If a whole year's worth of benefits is withheld between ages 62 to 64, benefits in the future will be raised by 6.7% each year. If the benefit is withheld between age 65 to 69, the future benefits will be raised by 6.0%. Given our terminal age at 80, it is favorable for individuals aged 62 to 64 and it is not actuarially fair for individuals aged 65 or older.

B.3 Taxable Social Security Benefits

The Social Security benefits are not taxable if it is the only income. If there is other income, compute "total income" as the sum of half of the benefits and all other income. If total income is no more than the base amount ($25,000 for head of household) then no benefits are taxable. If total income is higher than $34,000 then up to 85% of the benefits could be taxable.

Assume the Social Security benefits are y_{ss} and all the other income is y_o, the taxable part of Social Security benefits is calculated as

$$y_{ss,taxable} = \begin{cases} 0, & \text{if } y_o = 0 \text{ or } y_o + \frac{y_{ss}}{2} \leq 25000 \\ \min\left\{0.85 y_{ss}, \frac{1}{2}\min\left\{y_{ss}, y_o + \frac{y_{ss}}{2} - 25000, 9000\right\}, \right. & \\ \quad \left. + 0.85\max\left\{0, y_o + \frac{y_{ss}}{2} - 34000\right\}\right\} & \text{otherwise.} \end{cases} \tag{26}$$

[33]Before 2000, the earnings test applies to ages before 70. Since 2000, the earnings test is eliminated after reaching NRA.

Figure B1: Relative (to Year 2004) indices of National Average Wage Index and CPI-U.

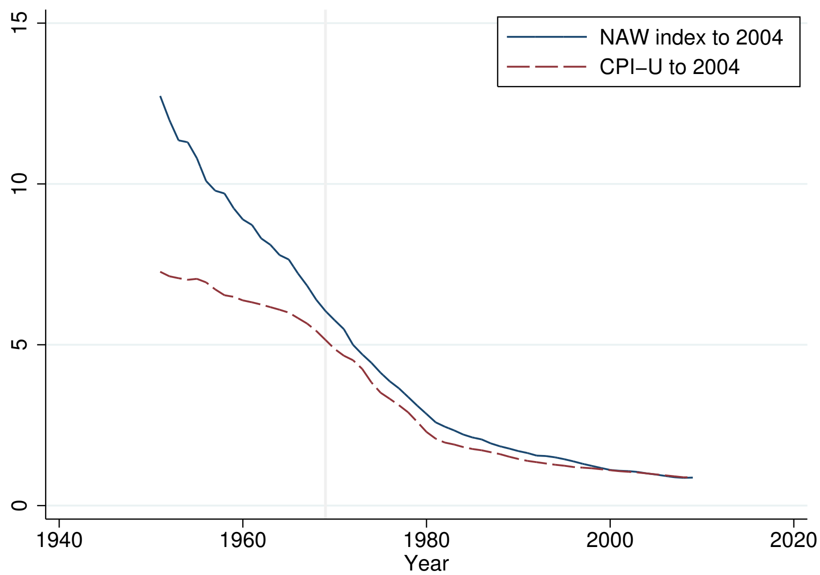


Figure B2: Share of minimum wage on AIME, assuming starting working from age 16. CPS data.