NBER WORKING PAPER SERIES

ESTIMATION OF A LIFE-CYCLE MODEL WITH HUMAN CAPITAL, LABOR SUPPLY AND RETIREMENT

Xiaodong Fan
Ananth Seshadri
Christopher R. Taber

Working Paper 29905
http://www.nber.org/papers/w29905

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2022

We would like to thank the editor (James Heckman), four anonymous referees, seminar participants at the SED, Tokyo, Yale, SOLE, CEPAR, Reading, Royal Holloway, Catholic University of Milan, Minnesota, Toulouse, Cornell, and Johns Hopkins for helpful comments and suggestions. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government. All remaining errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Estimation of a Life-Cycle Model with Human Capital, Labor Supply and Retirement
Xiaodong Fan, Ananth Seshadri, and Christopher R. Taber
NBER Working Paper No. 29905
April 2022
JEL No. J22,J24,J26

ABSTRACT

We develop and estimate a life-cycle model in which individuals make decisions about consumption, human capital investment, and labor supply and use it to analyze changes in Social Security rules. The most important aspect of our paper is human capital towards the end of the life cycle which responds to changes in the rules. 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), an endogenous retirement decision (which is typically assumed exogenous in the human capital literature), and accounts for the Social Security system. We estimate the model using indirect inference to match the life-cycle profiles of employment and measured wages from the SIPP data. The model replicates the main features of the data—in particular the large increase in measured wages and small increase in labor supply at the beginning of the life cycle as well as the small decrease in measured wages but large decrease in labor supply at the end of the life cycle. We use the model to estimate the effects of various changes to tax and Social Security policies and show that allowing for human capital accumulation is critical.

Xiaodong Fan
Department of Economics
Monash University
Clayton, VIC 3800
Australia
xiaodong.fan@monash.edu

Ananth Seshadri
Department of Economics
University of Wisconsin-Madison
1180 Observatory Drive
Madison, WI 53706
and NBER
aseshadr@ssc.wisc.edu

Christopher R. Taber
Department of Economics
University of Wisconsin-Madison
1180 Observatory Dr
Social Sciences Building #6448
Madison, WI 53706-1320
and NBER
ctaber@ssc.wisc.edu

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 (1975, 1976), 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. An important aspect of our model is that we do not treat retirement as a separate decision from labor supply in the model nor do we treat it differently in the data. We use it as a loose term that refers to low levels of labor supply late in life. In our model this declining labor supply over the life cycle occurs endogenously as part of the optimal life-cycle labor supply decision.

The most novel aspect of our paper is examining policy effects on human capital towards the end of working life. This is important as the retirement literature typically takes the wage process as given and estimates the incidence of retirement (e.g., Gustman and Steinmeier, 1986; Rust and Phelan, 1997; French, 2005; French and Jones, 2011). Cross-section raw wages for people who work fall substantially before retirement. They decline by over 25% between ages 55 and 65 (French, 2005). In much of the retirement literature, this trend is critical to understanding retirement behavior. By contrast, life-cycle human capital models take the retirement date as given, but model the formation of the wage process (e.g., Ben-Porath, 1967; Heckman, 1975, 1976; Heckman et al., 1998a; Manuelli et al., 2012). There has been work examining models of learning-by-doing and labor supply, most notably Imai and Keane (2004) and more recently Keane and Wasi (2016). However, these papers do not evaluate the effects of Social Security rules on human capital accumulation.[1] We estimate a model wherein the wage and labor supply choices are rationalized in one unified setting accounting for the Social Security system. After endogenizing both labor supply and human capital, our model is rich enough to explain the life-cycle patterns of both wages and labor supply, with a focus on wage patterns and declining labor supply (i.e. 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 indirect inference, matching the measured wage and labor sup-

[1] Iskhakov and Keane (2020) does look at retirement rules—but in the Australian context which is a very different system.

ply profiles of male high school graduates from the Survey of Income and Program Participation (SIPP). With a parsimonious life-cycle model in which the taste for leisure does not depend upon age or experience, we are able to replicate the main features of the data. In particular, we match the large increase in measured wages and very small increase in labor supply at the beginning of the life cycle as well as the small decrease in measured wages but very large decrease in labor supply at the end of the life cycle.

An important component of our model facilitating the fit in both ends of the life cycle is human capital depreciation. We take the definition of depreciation to be broad—it could be individuals' skills literally declining or it could be obsolescence of their skills as the nature of their work changes. The distinction is not important in our context. 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 which imposes a shadow cost on leisure. 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 shadow cost is much less a concern because the horizon is shorter. Older workers may choose not to re-enter at a lower wage so they continue to stay out of the labor market.

While we show that our relatively simple model is able to explain the data, we also want to evaluate changes to the Social Security system. Our baseline model does not incorporate health or part time work. Since these may be important components in retiring, credible evaluation should account for them. We estimate a specification that allows for both health and part time work. In particular we allow the taste for leisure to depend on health and for this effect to increase with age. We show that while becoming unhealthy has a large effect on labor supply health shocks are relatively uncommon in the years in which labor supply declines (ages 50–65). As a result, health plays a relatively minor role in explaining the decline in labor supply late in life. We also include the possibility of part time labor supply that could lead to more gradual retirement.

We 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 evaluate the remedy 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 combinations of declining wages, increasing actuarial unfairness of the Social Security and pension system, and increasing tastes for leisure. However, as mentioned above, 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 the Social Security benefit by 20%, the previous literature takes the same age-wage profile as in the baseline model and re-estimates the retirement behavior under the new environment. 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, less generous Social Security benefits result in higher labor supply later in the life cycle, so workers adjust their investment over the life cycle, which results in a higher human capital level as well as higher labor supply earlier. In an experiment in which Social Security benefits are decreased by 20%, the measured wage levels are up to 5% higher between 60 and 70. Over the whole life cycle, human capital investment, total employment rates, measured average yearly wages, and total pre-tax labor income increase by 0.3%, 1.2%, 0.3%, and 1.2% respectively in the general model with health and the part time option.

Section 2 of our paper briefly reviews the most relevant literature. Section 3 introduces the model and Section 4 explains how it is estimated. Section 5 presents the estimates from the baseline model and Section 6 explains the extension to the more general model. Section 7 simulates the policy counterfactuals and Section 8 concludes.

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 often use the term "human capital model" to mean "Ben-Porath model." Heckman (1975, 1976) extends the model and presents 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. 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 much work empirically estimating the Ben-Porath model. Mincer (1974) 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) estimate the Ben-Porath model and incorporate it into an equilibrium model. They utilize the implication of the standard Ben-Porath model where at old ages the 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. Browning et al. (1999) survey much of this literature.[2]

Another type of human capital model, the learning-by-doing model, draws relatively more attention in empirical work. In the standard learning-by-doing model human capital accumulates exogenously, but only when an individual works. Thus workers only impact their human capital accumulation through the work decision. In these models, the total cost of leisure is not only the direct lost earnings at the current time, but also includes the additional lost future earnings from the lower level of human capital. 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 model with learning-by-doing. Blundell et al. (2016) is a more recent example. These papers assume an exogenously fixed retirement age. Keane and Wasi (2016) and Iskhakov and Keane (2020) extend these models to consider workers at older ages and show the models fit well.

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 contrast the Ben-Porath style model predictions of the EITC policy effects with those of the learning-by-doing model. While learning-by-doing fits better for more educated women, the Ben-Porath style model fits better for less educated women.

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 (e.g. Gustman and Steinmeier, 1986) or increasing disutility toward labor supply (e.g. 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

[2] Other more recent work includes Taber (2002), who incorporates progressive income taxes into the estimation, Kuruscu (2006), who estimates the model nonparametrically, and Wallenius (2011) who focuses on labor supply elasticities.

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 (2013) 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. Studying the 1999 pension reform in German, Gohl et al. (2020) find that this assumption may be wrong. Responding to an exogenous increase in early retirement age from 60 to 63, employed women aged 53–60 increase their human capital investment significantly. This will likely change the wage profile.

3 Model

3.1 Overview

The model is a finite time life-cycle model. The main features that individuals choose are

- Human capital investment
- Labor supply (extensive margin)
- Consumption/savings

We add several other features to the baseline model both for fitting the data and for realism

- Social Security benefits/taxes
- Exogenous marriage and spousal labor supply
- Bequest motive
- Consumption floor

Our extended model adds features

- Health status—including disability
- Part time work

3.2 Environment and Econometric Specification

Demographics

Time is discrete and measured in years. Each individual i lives from period $t = 0$ to $t = T$. We use i and t subscripts to be clear how parameters vary across individuals and time. At the beginning of the initial period, each individual is endowed with an initial asset $A_{i0} \in \mathbb{R}$ and an initial human capital level $H_{i0} \in \mathbb{R}^+$.

Our model of family behavior is similar to Adda et al. (2017) in that we model the labor status of one individual (in our case the male) taking marriage, divorce, and spousal earnings as exogenous. Specifically, family status is an exogenous discrete state variable that can take three different values. A single or divorced individual is denoted by $M_{it} = 0$, while a married individual is indicated by either $M_{it} = 1$ (spouse not working) or $M_{it} = 2$ (spouse working). Each individual is single at the beginning of the life cycle, $M_{i0} = 0$. The family status evolves exogenously following an age-dependent Markov transition matrix.

Preferences

In the baseline model we focus on the extensive margin of labor supply only, so at each period the individual decides either to work or not. The flow utility at period t is

$$u_t\left(c_{it}, \ell_{it}, ssa_{it}; M_{it}, \varepsilon_{it}, a_{i0}\right) = \psi\left(t, M_{it}\right) \frac{c_{it}^{1-\eta_c}}{1-\eta_c} + \gamma\left(a_{i0}, M_{it}, \varepsilon_{it}\right) \ell_{it} + \varpi\left(t\right) ssa_{it} \tag{1}$$

where c_{it} is total family consumption, $\ell_{it} \in \{0, 1\}$ is leisure, and ssa_{it} is a dummy variable indicating whether individual i starts claiming Social Security benefits at time t. We mention again that retirement is not modeled explicitly—it is a phrase that one can use loosely to describe the status $\ell_{it} = 1$ for older workers but we don't model this any different than any other period of non-employment.

The coefficient $\psi(t, M)$ shifts the marginal utility of consumption (e.g., Gourinchas and Parker, 2002) and is assumed a parametric form,

$$\psi\left(t, M\right) = \exp\left(\varphi_1 t + \varphi_2 t^2 + \varphi_3 t^3 + \varphi_4 \mathbf{1}\left\{M \neq 0\right\}\right) \tag{2}$$

Note that the shifter depends upon marital status.

The coefficient $\gamma\left(a_0, M, \varepsilon\right)$ represents taste for leisure and also depends on the family status. We use the parametric form

$$\gamma\left(a_0, M, \varepsilon\right) = exp\left(a_0 + \sum_{j=1}^{2} a_j \mathbf{1}\left\{M = j\right\} + \varepsilon\right) \tag{3}$$

where ε_{it} follows an independent and identical normal distribution with mean 0 and variance σ_ε^2. The a_0 is a component of unobserved heterogeneity that we will discuss below. A key part of our exercise is that we do not explicitly allow $\gamma\left(a_0, M, \varepsilon\right)$ to vary systematically across age.

The final term of flow utility, $\varpi(t)$, accounts for tastes for applying for Social Security benefits. The literature has documented two peaks of Social Security application at age 62 and 65. Rust and Phelan (1997) demonstrate that the resource constraint and health insurance constraint are two major factors contributing to the peaks at age 62 and 65, respectively. While it is beyond the scope of this paper to explain these patterns explicitly, it is important to account for them when modeling labor supply and savings decisions of people of this age. With this goal in mind we let the model fit these patterns by assuming an individual obtains additional utility from receiving the Social Security benefit, and thus this total additional flow utility at period t becomes

$$\varpi(t) = b_{62}\mathbf{1}\left\{t = 62\right\} + \left[b_{65} + b_{65t}\left(t - 65\right)\right]\mathbf{1}\left\{65 \le t \le 70\right\} \tag{4}$$

The first term (b_{62}) captures the effect of resource constraint as well as pension eligibility and the second term captures the "security value" of health insurance through employment, as studied in previous literature (e.g., Rust and Phelan, 1997; French, 2005; French and Jones, 2011).

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

$$b(A) = b_1 \frac{\left(b_2 + A\right)^{1-\eta_c}}{1 - \eta_c} \tag{5}$$

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

Human Capital

If a man chooses to work, $\ell_{it} = 0$, he decides on how much time, $I_{it} \in [0, 1]$, to invest in human capital and spends the rest, $1 - I_{it}$, at effective (or productive) work from which the wage income is earned. Human capital is produced according to the production function

$$H_{it+1} = (1-\delta) H_{it} + \xi_{it} \pi_i I_{it}^{\alpha_I} H_{it}^{\alpha_H} \tag{6}$$

where H_{it} is the human capital level at period t, ξ_{it} is an idiosyncratic shock to the human capital innovation, π_i is a form of unobserved heterogeneity and δ, α_I,and $\dot{a}_H$ are parameters. If an individual chooses not to work, he does not invest in human capital (so $I_{it} = 0$) and human capital depreciates at rate δ.

We assume ξ_{it} is i.i.d and follows a log-normal distribution,

$$\log(\xi_{it}) \sim \mathcal{N}\left(-\frac{\log\left(\sigma_\xi^2 + 1\right)}{2}, \log\left(\sigma_\xi^2 + 1\right)\right). \tag{7}$$

This specification yields a level of ξ_{it} with a mean of one and a variance of σ_ξ^2.

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_{it}. Thus pre-tax labor income at any point in time is

$$w_{it} = H_{it} (1 - \ell_{it}) (1 - I_{it}). \tag{8}$$

Social Security and Budget Constraint

While we have tried to keep the baseline model as simple as possible, the Social Security system in the U.S. is such a crucial part of later life economic decisions that we incorporate it into the model in great detail capturing all important components. 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 let ss_{it} be a state variable indicating whether a person began claiming prior to period t and as mentioned above ssa_{it} indicates the decision to start claiming benefits. Since claiming is irreversible, once $ss_{it} = 1$ then ssa_{it} is no longer a relevant choice variable. Thus the law of motion can be written as

$$ss_{it+1} = \begin{cases} 0 & t = 0 \\ \max\{ss_{it}, ssa_{it}\} & \text{otherwise} \end{cases}. \tag{9}$$

The claiming decision (ssa_{it}) is made separately from the labor supply decision (ℓ_{it}) 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 benefits ssb_{it} which is a function of the claiming age, the Average Indexed Monthly Earnings ($AIME_{it}$), and working behavior after claiming (through the earnings test). In practice we approximate the AIME and use the Social Security rules as of 2004. The benefit ssb_{it} is updated each year if an individual worked to account for the earnings test. Details are in Appendix C. This is incorporated into the budget constraint

$$A_{it+1} = A_{it} + \Upsilon_t \left(rA_{it}, w_{it}, y_{it}, ssb_{it} \right) - c_{it} + \tau_{it}, \tag{10}$$

where A_{it} stands for asset, r is the risk free interest rate, and y_{it} is spousal income. $\Upsilon_t(\cdot)$ is the after-tax income which is a function of positive capital income, wage income, spousal income (if applicable), the Social Security benefit (if applicable), and the tax code. Details can be found in Appendix C.

Spousal income takes the form

$$y_{it} = \zeta_{it} \mathbf{1} \left\{ M_{it} = 2 \right\}, \; \log \left(\zeta_{it} \right) \sim \mathcal{N} \left(\mu_{\zeta t}, \, \sigma^2_{\zeta t} \right) \tag{11}$$

where ζ_{it} is an age-dependent log-normal random variable.

Government transfers, τ_{it}, provide a consumption floor $\underline{c}$ as in Hubbard et al. (1995) so

$$\tau_{it} = \max \left\{ 0, \, \underline{c} - \left(A_{it} + \Upsilon_t \left(rA_{it}, w_{it}, y_{it}, ssb_{it} \right) - \underline{A}_{it+1} \right) \right\}, \tag{12}$$

where $\underline{A}_{it+1}$ is the asset lower bound at period $t+1$.[3]

We note that some of our model assumptions are strong and lead all human capital to be financed by the worker through foregone wages. In particular, with Equation (8), we deviate from the original Ben-Porath model by ignoring monetary inputs which would have to be subtracted from the right hand side. Relaxing various parts of this model could lead the firm to finance some of the human capital—for example search frictions, asymmetric information, or if some of the human capital is firm specific (see e.g. Acemoglu and Pischke, 1998, 1999; Sanders and Taber, 2012). However, separating the contributions of firms and workers to training empirically is notoriously difficult, if not impossible. As in Becker (1962) and Rosen (1972) we take human capital investment as a broad concept assuming that workers have to sacrifice some current earnings in order to increase future earnings. In addition to workers' contribution to trainings, our wage specification also captures career choices with a low starting wage but a steeper age-earnings profile. For

[3]We define the asset lower bound as the amount that each individual can pay back for sure before death, as in Aiyagari (1994). Note that from Equation (5), $b_2 + A_T$ can not be negative. Since the probability of not working at each period is positive, the lower bound is characterized by the non-negative consumption so that $A_T \geq -b_2$. Discounted to period t this gives $\underline{A}_{it} = -b_2 / \left(1+r\right)^{T-t+1}$.

example, a law school graduate could either start as an associate in a law firm with a relatively low starting wage but very high potential earnings in the future, or choose another career with a higher starting wage but a flatter age-earnings profile.

3.3 Solving the Model

Four random variables are realized each period: evolving family status, M_{it}, spousal income, y_{it}, the shock in taste for leisure, ε_{it}, and the human capital innovation shock, ξ_{it}. The timing of the model works as follows: between periods $t-1$ and t the y_{it-1} and ξ_{it-1} are drawn determining A_{it} and H_{it}, the Markov process determines M_{it}, and the leisure shock ε_{it} is realized. The agent then simultaneously chooses consumption, labor supply, human capital investment, and when relevant, Social Security application. All four shocks are i.i.d. conditional on M_{it-1} from the perspective of the econometrician and the agent—so agents have no private information about their value prior to their realizations.

The recursive value function for $t < T$ can be written as

$$V_t\left(X_{it}\right) = \max_{c,\ell,I,ssa} \left\{ u_t\left(c,\ell,ssa;M_{it},\varepsilon_{it},a_{i0},\pi_i\right) + \beta E\left[V_{t+1}\left(X_{it+1}\right) \mid X_{it},c,\ell,I,ssa\right]\right\} \tag{13}$$

subject to (9)–(12) where

$$X_{it} = \left\{M_{it},A_{it},H_{it},ss_{it},AIME_{it},ssb_{it},\varepsilon_{it};a_{i0},\pi_i\right\} \tag{14}$$

is the vector of state variables. Note that $AIME_{it}$ is only relevant prior to claiming $(ss_{it}=0)$ while ssb_{it} is only determined after claiming $(ss_{it}=1)$. That is, prior to claiming, $AIME_{it}$ increases over time, but ssb_{it} has not yet been determined. At the time an individual starts to claim, the benefit (ssb_{it}) is calculated and relevant for the rest of life, but since $AIME_{it}$ only enters the model through its contribution to ssb_{it}, once that has been determined $AIME_{it}$ is no longer relevant. The expectation in (13) is over the human capital innovation ξ_{it}, spousal income y_{it}, the leisure shock ε_{it+1}, and the Markov draw for the new family status M_{it+1}.

For $t = T$ we write

$$V_T\left(X_{iT}\right) = \max_{c,\ell,I,ssa} \left\{ u_T\left(c,\ell,ssa;M_{iT},\varepsilon_{iT},a_{i0},\pi_i\right) + \beta E\left[b(A_{iT+1}) \mid X_{iT},c,\ell,I,ssa\right]\right\} \tag{15}$$

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

Define $\widetilde{X}_{it}$ to be the set of state variables apart from ε_{it}. The optimal consumption $\mathcal{C}_{it0}\left(\widetilde{X}_{it}\right)$, investment $\mathcal{I}_{it0}\left(\widetilde{X}_{it}\right)$, and Social Security claiming $\mathcal{SSA}_{it0}\left(\widetilde{X}_{it}\right)$ decisions conditional on participating in the labor market ($\ell_{it}=0$) can be obtained from

$$\begin{aligned}&\left\{\mathcal{C}_{it0}\left(\widetilde{X}_{it}\right),\mathcal{I}_{it0}\left(\widetilde{X}_{it}\right),\mathcal{SSA}_{it0}\left(\widetilde{X}_{it}\right)\right\}\\&\equiv\underset{c,I,ssa}{\operatorname{argmax}}\left\{\psi\left(t,M_{it}\right)\frac{c^{1-\eta_c}}{1-\eta_c}+\varpi\left(t\right)ssa+\beta E\left[V_{t+1}\left(X_{it+1}\right)\mid\widetilde{X}_{it},c,\ell_{it}=0,I,ssa\right]\right\}\end{aligned}\tag{16}$$

and the conditional value function is

$$\begin{aligned}\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right)\equiv&\psi\left(t,M_{it}\right)\frac{\left(\mathcal{C}_{it0}\left(\widetilde{X}_{it}\right)\right)^{1-\eta_c}}{1-\eta_c}+\varpi(t)\mathcal{SSA}_{it0}\left(\widetilde{X}_{it}\right)\\&+\beta E\left[V_{t+1}\left(X_{it+1}\right)\mid\widetilde{X}_{it},\mathcal{C}_{it0}\left(\widetilde{X}_{it}\right),\ell_{it}=0,\mathcal{I}_{it0}\left(\widetilde{X}_{it}\right),\mathcal{SSA}_{it0}\left(\widetilde{X}_{it}\right)\right].\end{aligned}\tag{17}$$

Notice that since there is no serial correlation in the stochastic shocks of leisure, ε_{it}, the conditional policy and value functions defined in Equations (16) and (17) do not depend on it.

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

$$\begin{aligned}&\left\{\mathcal{C}_{it1}\left(\widetilde{X}_{it}\right),\mathcal{SSA}_{it1}\left(\widetilde{X}_{it}\right)\right\}\\&\equiv\underset{c,ssa}{\operatorname{argmax}}\left\{\psi\left(t,M_{it}\right)\frac{c^{1-\eta_c}}{1-\eta_c}+\varpi\left(t\right)ssa+\beta E\left[V_{t+1}\left(X_{it+1}\right)\mid\widetilde{X}_{it},c,\ell_{it}=1,I_{it}=0,ssa\right]\right\}\end{aligned}\tag{18}$$

and define the conditional value function to be

$$\begin{aligned}\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)\equiv&\psi\left(t,M_{it}\right)\frac{\left(\mathcal{C}_{it1}\left(\widetilde{X}_{it}\right)\right)^{1-\eta_c}}{1-\eta_c}+\varpi(t)\mathcal{SSA}_{it1}\left(\widetilde{X}_{it}\right)\\&+\beta E\left[V_{t+1}\left(X_{it+1}\right)\mid\widetilde{X}_{it},\mathcal{C}_{it1}\left(\widetilde{X}_{it}\right),\ell_{it}=1,I_{it}=0,\mathcal{SSA}_{it1}\left(\widetilde{X}_{it}\right)\right].\end{aligned}\tag{19}$$

The optimal labor supply solution is

$$\ell_{it}=\arg\max_{\ell\in\{0,1\}}\left\{\widetilde{V}_{t\ell}\left(\widetilde{X}_{it}\right)+\gamma(a_{i0},M_{it},\varepsilon_{it})\ell\right\}\tag{20}$$

This gives a convenient functional form for the expected value function. To see this note that

$$\varepsilon_t^*\left(\widetilde{X}_{it}\right) \equiv \log\left(\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)\right) - a_{i0} - \sum_{j=1}^{2} a_j \mathbf{1}\left\{M_{it} = j\right\} \tag{21}$$

is the cutoff value of ε_{it} that determines work (see Appendix A for derivation). Then it is easy to see that the optimal labor supply decision is

$$\ell_{it} = \mathbf{1}\left(\varepsilon_{it} \geq \varepsilon_t^*\left(\widetilde{X}_{it}\right)\right) \tag{22}$$

where $\mathbf{1}(\cdot)$ is the indicator function.

Using properties of log-normal random variables, we show in Appendix A that the expected value function is

$$\begin{aligned} E\left[V_t\left(X_{it}\right) \mid \widetilde{X}_{it}\right] =& \Phi\left(\frac{\varepsilon_t^*\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) + \left(1 - \Phi\left(\frac{\varepsilon_t^*\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)\right) \cdot \\ & \cdot \left[\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + exp\left(a_{i0} + \sum_{j=1}^{2} a_j \mathbf{1}\left\{M_{it} = j\right\} + \frac{\sigma_\varepsilon^2}{2}\right)\frac{\Phi\left(1 - \frac{\varepsilon_t^*\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{\varepsilon_t^*\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)}\right] \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution.

Finally note that each component of $\widetilde{X}_{it+1}$ is a known function of $\widetilde{X}_{it}$, c_{it}, ℓ_{it}, I_{it}, ssa_{it}, ξ_{it}, and y_{it} (or ζ_{it}), so to solve for

$$E\left[V_{t+1}\left(X_{it+1}\right) \mid X_{it}, c_{it}, \ell_{it}, I_{it}, ssa_{it}\right] = E\left[E\left[V_{t+1}\left(X_{it+1}\right) \mid \widetilde{X}_{it+1}\right] \mid \widetilde{X}_{it}, c_{it}, \ell_{it}, I_{it}, ssa_{it}\right]$$

we integrate over the distributions of ξ_{it}, ζ_{it}, and M_{it+1}.[4]

3.4 Unobserved Heterogeneity

We allow for unobserved heterogeneity in ability to learn (π_i), initial human capital (H_{i0}), and tastes for leisure (a_{i0}). For computational reasons we only have nine types determining the joint distribution of (a_{i0}, π_i). 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_{i0}, the mean and variance of π_i, 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

[4] $\widetilde{X}_{it+1}$ is the set of state variables apart from ε_{it+1} at period $t+1$. Explicitly, M_{it+1} is a variable we integrate over, A_{it+1} is determined in Equation (10), H_{it+1} is determined in Equation (6), ss_{it} is determined by Equation (9), $AIME_{it}$ and ssb_{it} are described in Appendix C, and a_{i0} and π_i do not change.

since we are only using nine points we are not assuming that the Gauss-Hermite is a good approximation of a normal, but rather taking this as the parametrization itself.

Since human capital is already a continuous state variable in our model, we can be more flexible in its initial value. We allow it to be correlated with (a_{i0}, π_i) through the functional form

$$H_{i0} = \exp\left(\gamma_0 + \gamma_{a_0} a_{i0} + \gamma_\pi \pi_i + \nu_i\right) \tag{23}$$

where $\nu_i \sim \mathcal{N}\left(0, \sigma^2_{H_0}\right)$ is an i.i.d normal random variable.

4 Estimation

The estimation of the model is carried out using a three-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 evolution of the state variables involving spouses. In the third and largest step, we estimate the remaining preference and production parameters of the model using indirect inference. The model is described by Equations (1)–(23) and we summarize the parameters here. The parameters determining unobserved heterogeneity are $\mu_{a_0}, \sigma_{a_0}, \mu_\pi, \sigma_\pi, \rho, \gamma_0, \gamma_{a_0}, \gamma_\pi$, and σ_{H_0}. The additional parameters related to preferences are the discount rate, β, the intertemporal elasticity of consumption, η_c, the consumption shifter, $(\varphi_1, \varphi_2, \varphi_3, \varphi_4)$, the taste for leisure, $a_1, a_2, \sigma_\varepsilon$, the bequest parameters, b_1 and b_2, and the Social Security claiming parameters b_{62}, b_{65}, and b_{65t}. Human capital production is determined by δ, α_I, α_H and σ_ξ. Parameters related to the budget constraint are the interest rate r and the consumption floor $\underline{c}$. There are other parameters used to determine family status and spousal earnings. Finally there are initial values for the state variables, assets, A_{i0}, and Averaged Indexed Monthly Earnings, $AIME_{i0}$.

4.1 Pre-set Parameters

The set of parameters pre-set in the first stage includes the interest rate, the time discount rate, initial wealth and initial AIME, consumption floor, and bequest shifter.

One period is defined as one year.[5] The initial period in our model corresponds to age 18 and ends at age 80.[6] 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

[5] 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.

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

Table 1: Normalized or pre-set parameters

Parameters		Normalized/Pre-set Values
Interest rate	r	0.03
Discount	β	0.97
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	222.0

[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 \$444,000 in 2004\$.

as $\beta = 0.97$. The consumption floor is set as $\underline{c} = 2.19$, as estimated in French and Jones (2011).[7]

The parameter which determines the curvature of the bequest function is set as $b_2 = 222$, as in French and Jones (2011).[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. In Appendix E we show that the results are robust to other alternatives.

4.2 Demographics

Given unobserved heterogeneity in our model, we can not obtain consistent estimates for most of the remaining parameters outside the model. The exception is spousal demographics because they are unrelated to unobserved heterogeneity.

We estimate the 3×3 Markov transition matrix at each age from the SIPP data, smoothed by a probit regression on the age quadruple. Panels (a)–(c) of Figure B1 plot the transition probabilities at each age, and Panel (d) displays the resulting distributions which are similar to patterns in the SIPP data and in the CPS data. For each age, we estimate the mean $\mu_{\zeta t}$ and standard deviation $\sigma_{\zeta t}$ of the logarithm of the positive spousal income in the SIPP data, and then smooth them by an age quadruple function. Panel (e) of Figure B1 plots the estimated $\mu_{\zeta t}$ and $\sigma_{\zeta t}$.

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

[8]It is equivalent to \$444,000 in 2004 U.S. dollar. We also tried estimating b_2 to see how the results changed and the estimate is 215.4 with a standard error of 21.4, so it is neither statistically nor economically significantly different from the value we set.

4.3 Estimation Procedure

We apply indirect inference to estimate the remaining parameters of interest, Θ, with

$$\Theta = \left\{ \underbrace{\mu_{a_0}, \sigma_{a_0}, \mu_\pi, \sigma_\pi, \rho, \gamma_0, \gamma_{a_0}, \gamma_\pi, \sigma_{H_0}}_{\text{heterogeneity}}, \underbrace{\eta_c, \varphi_1 \varphi_2, \varphi_3, \varphi_4}_{\text{c}}, \underbrace{a_1, a_2, \sigma_\varepsilon}_{\text{leisure}}, \underbrace{b_1}_{\text{bequest}}, \underbrace{b_{62}, b_{65}, b_{65t}}_{\text{SSA}}, \underbrace{\delta, \alpha_I, \alpha_H, \sigma_\xi}_{\text{human capital}} \right\}$$

according to the following procedure.

i) Calculate the auxiliary model from the data.

ii) Iterate on the following procedure for different values 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 auxiliary model from the simulation.

 (d) Calculate the distance between the simulated auxiliary model and the data auxiliary model.

4.4 Data and the Auxiliary Parameters

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. We use the SIPP because it is a large representative data set with a panel data element. To focus on as homogeneous a group as possible, the sample only includes male high school graduates. Estimation results for college graduates are presented in Appendix F.

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

$$W_t = H_t \left(1 - I_t\right) \tag{24}$$

in the model.

The primary four things that agents in our model choose are consumption, labor supply, human capital investment, and Social Security application. We obtain life-cycle data on the three of these that can be easily observed: consumption, labor supply, and Social Security application. Human capital is not observed directly, so we choose moments on measured wages. We match the life-cycle profile of measured wages and also life-cycle measured wages conditional on fixed effects as they look quite different and we want our model to be able to explain both. Since depreciation will play an important role in our results, we construct a measure of human capital decline following spells of non-employment. To measure persistence in employment we also collect data on the transition rates in and out of work.

In SIPP an individual is observed at most three times each year. Due to the seam bias problem in SIPP we only use measures of working and wages during the survey month. We use only years in which we observe the worker three times and if an individual works in two or three of the observations, he is categorized as working in the labor market, otherwise not.[9] We construct the hourly wage as the earnings in the survey month divided by the total number of hours worked in the survey month and average across the survey months in a year in which the respondent works.

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 supply patterns among very early labor market entrance as they transition from school to work as suggested by literature (Topel and Ward, 1992; Neal, 1999). Our model does over-predict the labor supply for those individuals.

Eight sets of moment conditions across different ages are chosen to assemble the auxiliary model. We use a total of 645,630 panel observations from 100,298 different respondents.

i) The employment rates (ER), ages 22–65.[10]

ii) The first moments of the logarithm of measured wages, ages 22–65.

iii) The first moments of the logarithm of measured wages after controlling for individual fixed effects, ages 22–65.[11]

iv) The second moments (standard deviation) of the logarithm of measured wages, ages

[9] Clearly this aggregation is imperfect as the model is simulated at an annual basis. Ideally we would simulate the model at the monthly level, but this is not computationally feasible. Our goal is to understand labor supply at the life-cycle frequency so abstracting from the monthly frequency does not seem first order.

[10] We focus on the employment rather than the labor force participation in both data and the model, as we do not have unemployment in the model.

[11] 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. We obtain the log measured wages after controlling for individual fixed effects using a first difference estimator rather than the fixed effect estimator as the identification is much clearer in the former.

22–65.

v) The first moments of adult equivalent consumption, ages 22–65.[12]

vi) The Social Security benefit application rates, ages 62 to 70.

vii) The overall transition probabilities averaged between age 35 and 50,[13]

   (a) from working to not working

   (b) from not working to working

viii) The average measured wage change rate after one nonemployment spell averaged between age 41 and 65.[14]

We match both age-measured wage profiles, with and without controlling for individual fixed effect as the two have quite different patterns.

Figures 1(a)–1(e) present the six profiles. Figure 1(a) plots the employment rates between age 22 and 65. Figure 1(b) plots two log measured wage profiles. The first one is the profile from the pooled sample, while the second one is the profile after controlling for individual fixed effects. The original log measured 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 1(c) shows the extent to which the variance of log measured wages increases with age. Figure 1(d) presents adult equivalent consumption profile while Figure 1(e) illustrates the two peaks at age 62 and 65 in the Social Security benefit application ages.

The most interesting result in Figures 1(a)–1(e) is the discrepancy between the age-measured wage profiles with and 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 (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 intu-

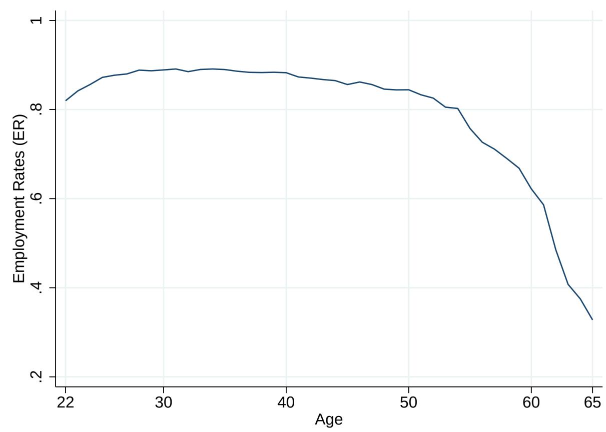
[12]The adult equivalent consumption profile is constructed from the Consumer Expenditure Survey as in Fernández-Villaverde and Krueger (2007).

[13]We want to focus on transition probabilities caused by heterogeneity rather than retirement, so we choose the prime working ages. Choosing a different age period, such as age 41 to 65, does not change the results in any significant way.

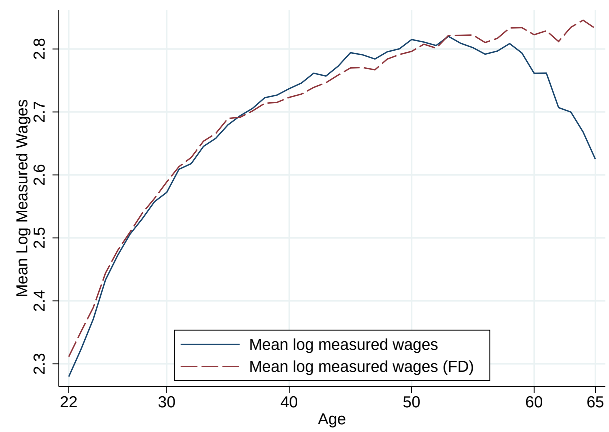
[14]We choose this close-to-retirement age group to emphasize the depreciation and minimize the investment channel. We estimate this parameter as the decline at 12 months using a local linear regression with a uniform kernel and a bandwidth of 7 months. We explored this at different months and do not find evidence against the constant rate of depreciation, but the standard errors are quite large.

Figure 1: Data moments and profiles

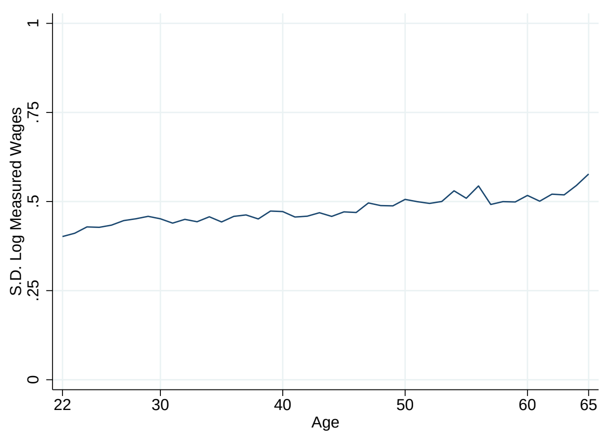
(a) Employment rates (SIPP)



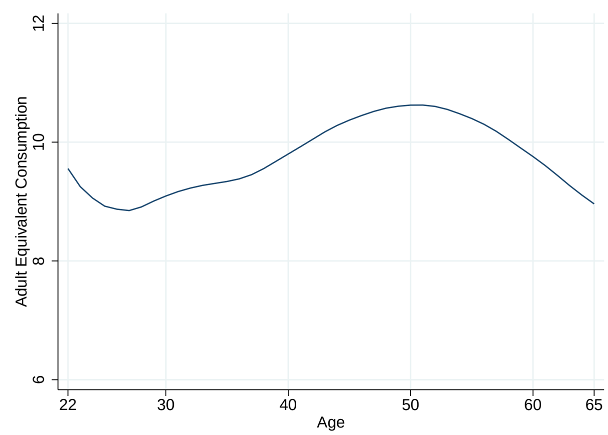
(b) Mean log measured wages (SIPP)



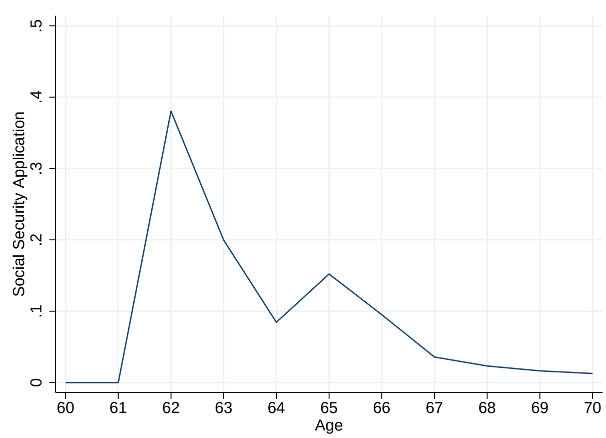
(c) S.D. log measured wages (SIPP)



(d) Adult equivalent consumption (CES)



(e) Social Security application (SIPP)



(f) Mean log measured wages (CPS)

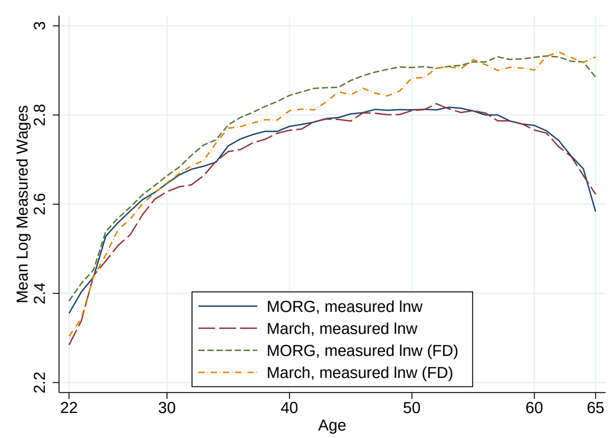


Table 2: Transitions moments

Models	Transition Probabilities[a] Working to Not Working	Not Working to Working	Wage Change Rate After One Nonemployment Spell[b]
Data	0.034	0.200	-0.071
Baseline model	0.040	0.219	-0.079
No depreciation at work, Ver 1	0.022	0.196	-0.065
No depreciation at work, Ver 2	0.033	0.240	-0.066
Model with health & part-time	0.028	0.167	-0.082

[a]The transition rate is the average transition probability between age 35 and 50.
[b]The average wage change rate after one nonemployment spell is the average change rate between age 41 and 65.

ition is that when the human capital depreciation outweighs the investment, wages start to decline which generates a hump-shaped profile. We show below that this is not necessarily the case as the decline in investment can offset the depreciation and we can fit the pattern of the wage profile after controlling for fixed effects. It does make fitting the pattern more challenging because we need to explain the decrease in labor supply later in life when there is little evidence that measured 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 using the method proposed in Madrian and Lefgren (2000), and we have a short panel with each individual interviewed twice, one year apart.[15] 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 1(f) presents the age-measured wage profiles with or without controlling for individual fixed effects for male high school graduates from the 1979–2018 CPS MORG data and the 1989–2018 IPUMS-CPS March data (Flood et al., 2021).[16] We find an even larger discrepancy in the age-measured wage profiles than in the SIPP data presented in Figure 1(b). In the model this profile corresponds to net earnings $H_t(1 - I_t)$.

The values for the remaining moments (vii and viii) can be seen in the first row of Table 2. One can see that there is substantial persistence in labor supply and that the measured wage change following a nonemployment spell is large.

[15]For MORG data, they are the fourth and eighth interview.
[16]Time fixed effects are filtered out, as described in Footnote 11.

5 Estimation Results

The estimates of the parameters are listed in Table 3. Of particular importance are the depreciation rate, δ, curvature in the human capital production function, α_I, and σ_ε which determines the elasticity of labor supply. Before discussing these parameter values we examine the fit of the model in Figures 2(a)–2(f). The fit of the model in the two overall transition probabilities and the average wage change rate after one nonemployment spell is presented in the first two rows in Table 2.[17]

The first point is that our parsimonious model can reconcile the main facts in the data: a small increase in labor supply/large increase in measured wages at the beginning of the life cycle along with the large decrease in labor supply/small decrease in measured wages at the end of the life cycle. The simulated employment rate increases slightly between age 22 and 30 as shown in Figure 2(a). More importantly, this simple model is able to generate a massive decline in labor supply between age 55 and 65, which fits the sharp decline of employment rates within that age period in the data and simultaneously the flat measured wage profile in the fixed effect specification.

Our model generates similar discrepancy between the log measured wages with and without controlling for individual fixed effects, as shown in Figures 2(c) and 2(b), and both profiles fit the data well. Log measured wages after filtering out individual fixed effects increase at a decreasing pace and does not decrease during the examined period (Figure 2(c)). On the other hand, Figure 2(b) shows that the original log measured wage profile presents a hump shape which resembles the data profile. The model also replicates the log measured wage variation as in the data (Figure 2(d)).[18]

Our model tracks the hump-shape and the level of the adult equivalent consumption profile reasonably well (Figure 2(e)), as well as the two peaks at 62 and 65 in the Social Security application (Figure 2(f)).[19] The model also generates the similar overall transition probabilities between working and not working and the average wage change rate after one nonemployment spell, as shown in Table 2.

[17] The overidentification test statistic is reported in the bottom of Table 3. The model is rejected at the 1% 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 life-cycle patterns.

[18] For the most part we find that these basic patterns happen within our unobserved types. The types are important for explaining the level of the standard deviation of log wages but do not play a key role in explaining the lifecycle patterns of human capital and labor force participation.

[19] We didn't force our model to fit the initial decline at young ages in the consumption profile of high school graduates for two reasons. First, the initial decline in the data needs further investigation and could be for reasons not present in our model (e.g., sponsored by or living with parents). Second, the consumption and leisure are additively separable in our model, and thus the shape of initial consumption does not affect the labor supply decision in the absence of binding borrowing constraints.

Table 3: Estimates in the baseline model[a]

Parameters		Estimates	Standard Errors
HC depreciation[b]	δ	0.086	(0.003)
HC production function: I factor	α_I	0.098	(0.009)
HC production function: H factor	α_H	0.051	(0.007)
Standard deviation of HC innovation	σ_ξ	0.027	(0.005)
Consumption: CRRA	η_c	4.034	(0.016)
Consumption shifter: coef on t $\left(\times 10^{-1}\right)$	φ_1	0.121	(0.010)
Consumption shifter: coef on t^2 $\left(\times 10^{-2}\right)$	φ_2	0.187	(0.008)
Consumption shifter: coef on t^3 $\left(\times 10^{-3}\right)$	φ_3	-0.042	(0.002)
Consumption shifter: coef on married	φ_4	1.216	(0.065)
Leisure: standard deviation of shock	σ_ε	0.197	(0.009)
Leisure: spouse not working	a_1	-0.240	(0.032)
Leisure: spouse working	a_2	-0.155	(0.017)
Bequest weight	b_1	60,600,532	(4,669,063)
Parameter heterogeneity[c]			
Leisure: mean of intercept	μ_{a_0}	-5.706	(0.047)
Leisure: standard deviation of intercept	σ_{a_0}	0.634	(0.039)
HC productivity, mean	μ_π	1.854	(0.019)
HC productivity, standard deviation	σ_π	0.744	(0.011)
Correlation between a_0 & π	ρ	-0.745	(0.025)
Initial human capital level at age 18			
Intercept	γ_0	1.475	(0.110)
Coefficient on a_0	γ_{a_0}	0.028	(0.005)
Coefficient on π	γ_π	0.627	(0.051)
Standard deviation of error term	σ_{H_0}	0.002	(0.004)
Additional Social Security Application effects			
Effect of resource constraint $\left(\times 10^{-3}\right)$	b_{62}	0.297	(0.028)
Effect of health insurance: constant $\left(\times 10^{-3}\right)$	b_{65}	0.008	(0.009)
Effect of health insurance: coef on t $\left(\times 10^{-3}\right)$	b_{65t}	0.138	(0.026)
χ^2 Statistic = 814[d]		Degrees of freedom = 207	

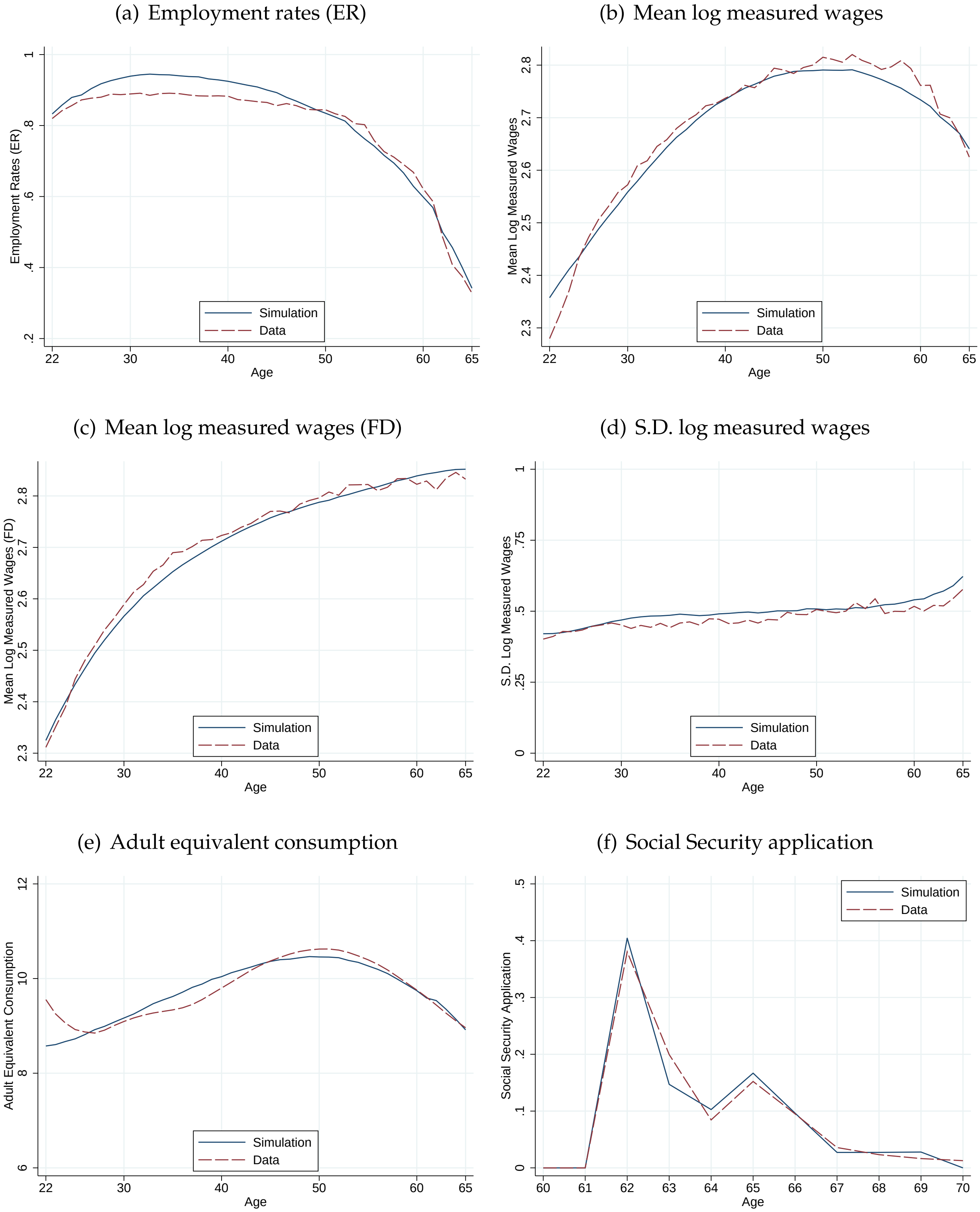
[a]Indirect inference estimates. Estimates use a diagonal weighting matrix. Standard errors are given in parentheses.

[b]HC: Human Capital.

[c]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.

[d]This is the J-statistic. The critical value of the χ^2 distribution is $\chi^2_{(207,0.01)} = 257$.

Figure 2: Fit of model



We obtain our fit of the life-cycle profiles of labor supply and log measured wages despite the lack of any explicit time-dependent preference of leisure, production or constraints in our model. A key feature of our model makes them possible: the combination of human capital depreciation and the separation between the effective labor and observed labor. We discuss these issues in the following subsection. We mention that precautionary savings also causes labor supply to fall late in life. Workers build up a buffer stock of assets which leads to lower labor supply late in life. As this is a common feature of many models we focus on the more novel aspects of our model.

Our focus in what follows is on understanding how our model fits these facts. Keane and Wasi (2016) are also able to fit the profile. Their model shares many features with ours so presumably much of this is relevant for their fit as well, though their model does differ in other ways so it won't be identical.

5.1 The Role of Human Capital Depreciation

Our estimate of human capital depreciation in the baseline model is quite high as it takes a large value to fit our data. A major issue in fitting the data is that as we can see in Figure 2(c) that in both the model and the data, once fixed effects are accounted for, measured wages are close to flat for ages 50–65 despite the fact that there is a large decrease in labor supply. There are two different aspects of depreciation that are important.

The first was mentioned in the introduction and is due to depreciation off the job. When workers do not work they can not invest in human capital and their human capital falls. It is directly reflected in the observed wage decline after one nonemployment spell in the data and we match it in the model. This yields a shadow cost of not working. Most importantly, the importance of this cost varies across the life cycle. When workers are in their mid-career the cost is high, but as they get older and the horizon gets shorter the importance of this shadow cost declines as does labor supply. Older workers take time off, their human capital declines, and they have even less reason to re-enter the labor market. This is one of the major driving forces of the decline in labor supply at older ages in spite of flat wages (after controlling for the fixed effects). The importance of the off-the-job depreciation can be illustrated by contrasting the simulated wage with the short-dash-dotted line labeled "log(H): all" in Figure 3(a) which presents the mean of $\log(H)$ but for the full population, not just workers. From the latter curve one sees at older ages (around 60) the actual human capital level has already depreciated to a relatively low level, even though the measured wage level is still quite high. This is due to the decline in investment that happens around that time, both from the decline in investment on the job but especially from not working at all. This is much less subtle and it is not surprising that the non-workers with low levels of human capital are not returning to the labor market.

The second feature is more subtle and comes from a point emphasized by Heckman et al. (1998a): measured wages are different than human capital. This distinction between human capital and measured wages can help explain some of this effect. For example, it is possible for wages to be flat but human capital to be declining. The reason is that wages are equal to $H_{it}(1 - I_{it})$ so it is possible that H_{it} is falling but I_{it} is falling as well to counteract it.

One can see the evidence of this second feature in Figure 3(a). The two lines to focus on for this point are the solid line which is the mean simulated log wage and the long-dash-dotted line labeled "log(H): working" which shows the potential log wage for workers. That is, they condition on the same set of individuals, but the former presents the mean of $\log(H_{it}(1 - I_{it}))$ while the latter presents the mean of $\log(H_{it})$. One can see that the human capital peaks around age 45 and is roughly flat during ages 45-55 and then starts to decline. By contrast, the measured wage keeps increasing after age 45 and peaks around 55, after which the measured wage starts declining slowly. This is because investment keeps decreasing. By age 62, however, since the worker has already allocated most of his time in effective working, there is little further room for such adjustment. This distinction can help explain the falls in labor supply at older ages. Figure 3(b) presents the investment profile in our model; the level and trend are very close to Figure 4 in Mulligan (1998), who calculates the time spent learning skills on the job at a 1976 study of time use by the Survey Research Center. The shape is also similar to that in Blundell et al. (2019) who find substantial training among older workers—though using data from the United Kingdom.

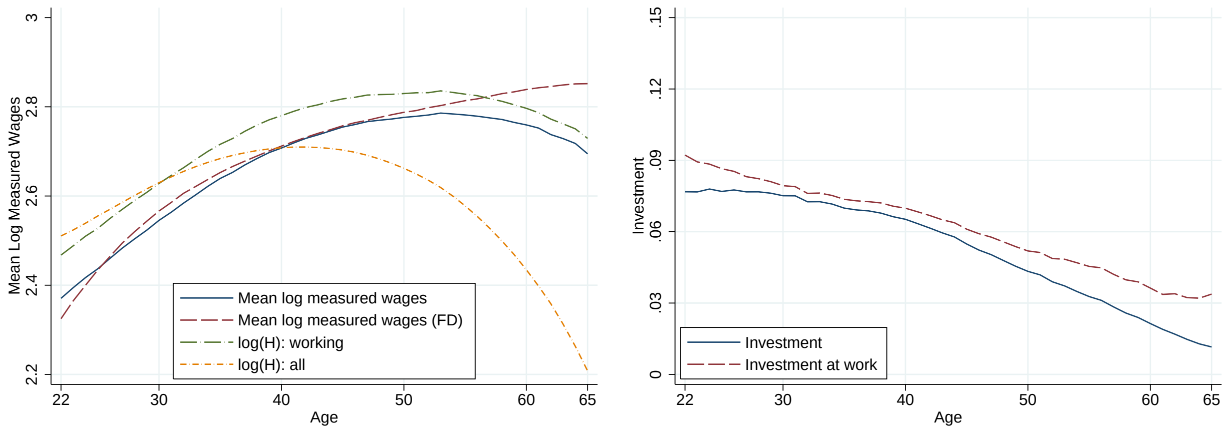
The relatively high value of investment late in the working career is also related to why we find a much smaller level of the human capital curvature parameter, α_I, compared to the literature summarized in Browning et al. (1999). The larger is α_I, the steeper is the decline in human capital investment 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 our models are quite different in a number of ways including the fact that this model includes leisure and in their model they set deprecation to zero.

To deliberately show the significance of the human capital depreciation in matching the labor supply profile and the two log measured wage profiles, we re-estimate the

Figure 3: Log measured wages, human capital, and investments

(a) Log measured wages and human capital

(b) Investment



model without depreciation. We tried estimating a model with no depreciation at all but our best fit of this model was still very poor so we do not discuss it. We focus on re-estimation of a less extreme case in which we continue to have it off the job but not while working.[20] In the latter alternative model, we assume human capital only depreciates if not working,

$$H_{it+1} = \begin{cases} H_{it} + \xi_{it}\pi_i I_{it}^{\alpha_I} H_{it}^{\alpha_H} & \text{if } \ell_{it} = 0 \\ (1-\delta)\, H_{it} & \text{if } \ell_{it} = 1 \end{cases} \tag{25}$$

and this model is labeled as the "no depreciation at work" model. Two versions of the estimation results of this model are listed in Table 4.

The fit of these two models is shown in the third and fourth rows of Table 2 and Figure 4. Note that both models still do capture much of the main features—partly as a result of depreciation off the job. That is we do see a decline in labor supply late in the life cycle. However, neither model is able to match the profiles of labor supply and log measured wages simultaneously. In particular, Version 1 is able to fit the log measured wages but not the employment rates (long-dashed lines), while Version 2 improves the fitting on the employment rates at the costs of worse fittings on the log measured wages (long-dash-dotted lines). Neither version is able to fit the wage variation well.[21]

[20] We have also re-estimated a model in which we allow the depreciation while working to differ from the depreciation off the job. The two estimates are very close to each other (0.0862 on the job vs 0.0855 off the job) so as our main specification we keep the version where they are a single parameter.

[21] We also tried looking at this in a different way by estimating very simple versions of this model with completely exogenous human capital and another with learning-by-doing (in which people internalize depreciation when making their labor supply decisions). We find a similar result—the exogenous model has trouble fitting the labor supply patterns. These results can be found in Appendix D.

Table 4: Two versions of estimates of model with no depreciation at work[a]

Parameters		Estimates Version 1	Version 2
HC depreciation[b]	δ	0.063	0.056
HC production function: I factor	α_I	0.650	0.669
HC production function: H factor	α_H	0.001	0.001
Standard deviation of HC innovation	σ_ξ	0.0001	0.0001
Consumption: CRRA	η_c	3.944	3.944
Consumption shifter: on t $\left(\times 10^{-1}\right)$	φ_1	1.198	1.198
Consumption shifter: on t^2 $\left(\times 10^{-2}\right)$	φ_2	-0.132	-0.132
Consumption shifter: on t^3 $\left(\times 10^{-3}\right)$	φ_3	-0.013	-0.013
Consumption shifter: coef on married	φ_4	0.066	0.066
Leisure: standard deviation of shock	σ_ε	0.071	0.107
Leisure: spouse not working	a_1	0.311	0.478
Leisure: spouse working	a_2	-1.017	-1.233
Bequest weight	b_1	83,015,680	83,015,680
Parameter heterogeneity[c]			
Leisure: mean of intercept	μ_{a_0}	-4.945	-4.906
Leisure: standard deviation of intercept	σ_{a_0}	0.664	0.807
HC productivity, mean	μ_π	0.537	0.519
HC productivity, standard deviation	σ_π	1.007	1.586
Correlation between a_0 & π	ρ	-0.804	-0.508
Initial human capital level at age 18			
Intercept	γ_0	2.477	2.453
Coefficient on a_0	γ_{a_0}	0.092	0.090
Coefficient on π	γ_π	0.739	0.765
Standard deviation of error term	σ_{H_0}	0.004	0.080
Additional Social Security Application effects			
Effect of resource constraint $\left(\times 10^{-3}\right)$	b_{62}	0.617	0.867
Effect of health insurance: constant $\left(\times 10^{-3}\right)$	b_{65}	0.471	0.526
Effect of health insurance: coef on t $\left(\times 10^{-3}\right)$	b_{65t}	0.473	0.543
χ^2 Statistic[d] (Degrees of freedom = 207)		1793	1864

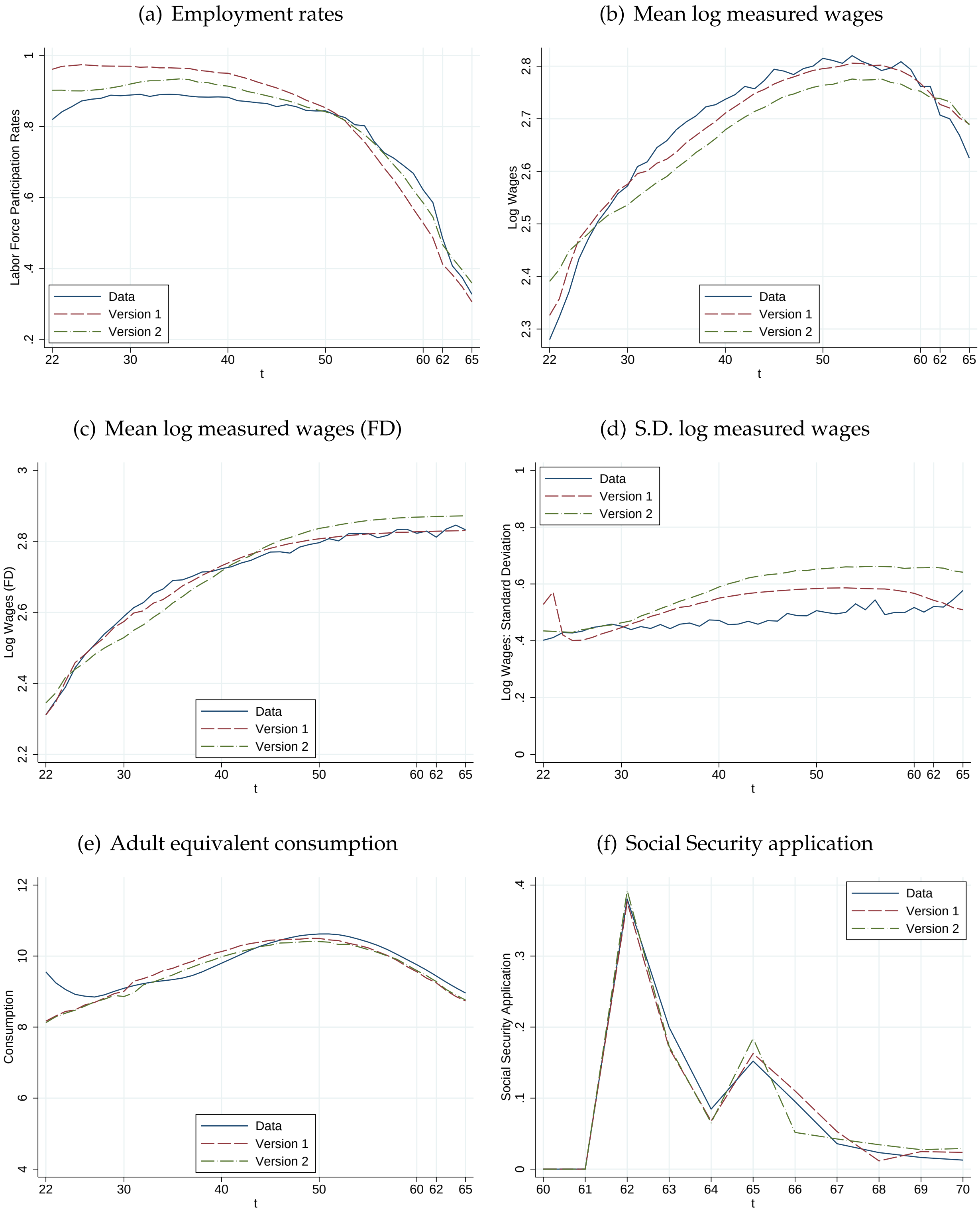
[a]Indirect inference estimates. Estimates use a diagonal weighting matrix. Standard errors are given in parentheses.

[b]HC: Human Capital.

[c]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. We restrict the lower bound of π to be 0.001 as we assume a positive marginal productivity of HC production.

[d]This is the J-statistic. The critical value of the χ^2 distribution is $\chi^2_{(207,0.01)} = 257$.

Figure 4: Fit of the alternative models with no depreciation at work



Given that we have shown that our estimate of a depreciation value $\delta = 0.086$ plays an important role explaining the pattern of wages and life-cycle labor supply, it is important to place this value into the range of estimates in the literature. This is not easily done as there is a very large range of estimates and none are directly comparable to our number. Some are larger than our 8.6% estimate and others are smaller. There are broadly three different literatures that estimate related parameters. The first of these 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%.[22] 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.[23] A third literature examines the effect of the length of an unemployment spell on the wage at rehire. Schmieder et al. (2016) 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 re-employment wages. They find that one extra month of unemployment leads to a decrease in wages of 0.8% which gives an annual rate close to our estimate of 8.6%. Keane and Wolpin (1997) is a similar style paper to ours and finds an analogue of our depreciation of 9.6% for blue-collar work which is close to our estimate. While it looks at women in England, Blundell et al. (2016) is also of similar style to our paper in the

[22] 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 leads to a decline in earnings of 25%. Kunze (2002), Gorlich and de Grip (2009), and Adda et al. (2017) all use German data (IAB employment sample and/or German Socio-economic panel). Kunze (2002) finds estimates of about 2–5% wage 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. Adda et al. (2017) find a range of estimates typically with small numbers but the largest being 6.9%.

[23] 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.

sense that it is a structural life-cycle model of labor supply and human capital formation. Interestingly, their analysis reveals a substantial depreciation of human capital ranging from 6% to 11%.

5.2 Elasticity of Labor Supply

The key parameter in our model that determines the elasticity of labor supply is σ_ε but its value is hard to interpret. In this subsection, we provide a measure to help the reader judge the magnitude. 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. However, to compare our elasticity to something somewhat similar to what is estimated in the literature we construct a counterpart 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 employment rate using the baseline model and dividing by the difference in the measured wage. We should emphasize that this is not the actual elasticity in our model because it does not account for the shadow cost of time. We call it the analogue to the empirical elasticity (aee) and define it formally below.

Let h_t^b be the employment rate at age t in the baseline model and h_t^t be the employment rate at age t (denoted by the subscript) in the simulation in which we increase the rental rate at age t (denoted by the superscript) by 10%. Then we define that elasticity to be

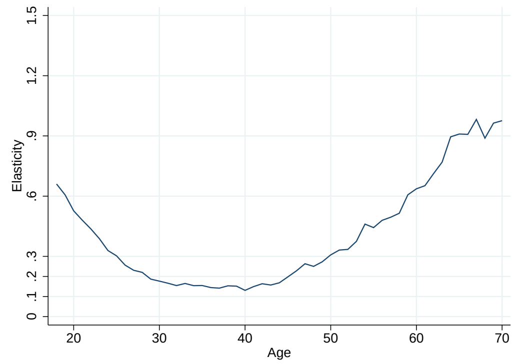
$$aee \equiv \frac{\log\left(h_t^t\right) - \log\left(h_t^b\right)}{\log(w_t^t) - \log(w_t^b)} \tag{26}$$

Note that while we call it the analogue of the empirical elasticity, it is not precisely that either as we have assumed that the changes are perfectly anticipated. Our main goal is to give readers some sense of the magnitude and we find this to be a convenient summary statistic.

This summary statistic is plotted in Figure 5. One sees a U shape: labor supply appears to be much more elastic at older ages and at near labor market entry than in the age range 30-50. This basic pattern is similar to similar plots in Keane and Wasi (2016) that also show U shaped elasticities.[24] In our case (and presumably theirs as well), this is due in large part to the shadow cost of leisure. The shadow cost is substantially larger for young workers than for older workers since the older workers have a shorter time horizon. As

[24] Our levels are not directly comparable to theirs.

Figure 5: Analogue to the empirical elasticity of labor supply



a result, the labor supply of young workers is less responsive to temporary wage shocks than is the labor supply of older workers. Like Keane and Wasi (2016) it is also due to the density of the tastes for leisure γ_t. When the probability of working is closer to 50% the density of people close to being indifferent will be larger which results in a larger elasticity.

For individuals under age 60 these estimates are very close to the estimates of labor supply elasticities found in the literature—though our definition of labor supply is not identical to them so they are not precisely the same. For example, the early literature estimates the Frisch elasticity being 0.09 (Browning et al., 1999), 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 retirement ages, Rogerson and Wallenius (2013) suggest that the intertemporal elasticity of substitution (IES) is 0.75 or greater given empirically reasonable level of nonconvexities or fixed costs. The average of our estimates between ages 55 and 65 is remarkably close to theirs.

6 Roles of Health, Disability and Part Time

We have intentionally kept our model simple to show that a simple model can explain the dramatic fall in labor supply at the end of the life cycle. However, our next goal is to simulate policy counterfactuals. While we find it useful to show we can fit the model with a parsimonious model, there are other important features that we feel are needed to make the policy counterfactuals credible. 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) where we include disability as part of health. If the primary reason for retirement is health or disability, its omission might seriously distort our results. The second feature we incorporate is part time work. Most importantly, we have assumed that retirement involves moving from full time work to no work. Working part time could make that pattern more gradual. Our next version of the model incorporates both of these features and we refer to it as the extended model.

6.1 Health and Disability

Our main innovation is to allow for an additional state variable—health status, $S_{it} \in \{e, g, b, d\}$, with e being in excellent health, g in good health, b in bad health, and d being disabled. We model the disability state as absorbing; it also makes one eligible for certain benefits, including the Social Security Disability Insurance or Supplemental Security Income (Appendix C.4). Each individual is assumed to have good health from the beginning of the first period up to age 49, $S_{it} = g,\ t \leq 49$. After age 49, the health status evolves exogenously according to a time-dependent probability transition matrix, and is realized at the beginning of each period before any choice is made. This process is estimated outside the model.[25]

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

$$\gamma\left(a_0, M, t, S, \varepsilon\right) = exp\left(a_0 + a_M + a^0_{hS} + a^1_{hS}t + \varepsilon_{it}\right). \tag{27}$$

That is, individuals with excellent, bad, or disabled health status have a different taste for leisure than those with good health and this difference changes as they age.[26] We normalize $a^0_{hg} = 0$ and we assume that leisure taste only changes for non-healthy people—that is, we assume $a^1_{he} = a^1_{hg} = 0$ but estimate a^1_{hb} and a^1_{hd}.

To estimate these five new parameters, $\left\{a^0_{he}, a^0_{hb}, a^0_{hd}, a^1_{hb}, a^1_{hd}\right\}$, we include three more sets of moment conditions: the difference in employment rates between workers with

[25]The health transition matrix is estimated from the CPS data. We include the health status from age 50 for two reasons. First, most individuals have excellent or good health before age 50. Second, this simplification reduces computation time. Panels (a)–(c) of Figure B2 plot the health transition probabilities at each age; Panel (d) plots the distribution of four health status. Between age 55 and 64, 11.7% of individuals are disabled, close to the actual SSDI enrollment ratio (10.9%) from the SSA administrative data in 2004 (Autor and Duggan, 2006).

[26]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 in this extended model with health. The main point of this subsection is to estimate a more general model to improve the credibility of the counterfactual exercises, so even though we are favoring the model with health by allowing this extra flexibility, health has a relatively minor role.

excellent health and workers with good health, the employment rate difference between workers with good health and workers with bad health, as well as the difference between workers with bad health and those with disability, across ages from 50 to 65. The data moments are derived from the CPS March data and the raw patterns can be seen in Panel (g) of Figure 6.

6.2 Part Time Work

We also include the part time work as a leisure choice in the extended model. At each period, an individual decides to work full time ($\ell_{it} = 0$), or to work part time ($\ell_{it} = p$), or not to work ($\ell_{it} = 1$). As we did for health we allow the utility from working part time to vary across ages and is to be estimated.[27] One can see from Figure 6(h) that part time work is uncommon for the sample we study. Working part time means spending half time in the labor market and the other half time at leisure. We let ϱ_t be the parameter vector that determines part time utility which varies across ages but not across individuals and we assume the utility of leisure associated with part time work is $\gamma\left(a_0, M, t, S, \varepsilon\right) \varrho\left(t\right)$, where

$$\varrho\left(t\right) = \frac{1}{1 + \exp\left(-a_{\varrho 0} - a_{\varrho 1} t - a_{\varrho 2} t^2\right)}.$$

We restrict this variable to be in the unit interval so the utility of leisure from part time work lies between no work and full time work. To estimate these three new parameters, $\left\{a_\varrho^0, a_\varrho^1, a_\varrho^2\right\}$, we include the part time employment rate at each age from 22 to 65 as additional moments.

If an individual chooses to work part time, the investment in human capital is $I_{it} \in \left[0, \frac{1}{2}\right]$ and the effective work time is $\frac{1}{2} - I_{it}$, with wage earning

$$w_{it} = H_{it} \cdot \left(\frac{1}{2} - I_{it}\right).$$

The solution is analogous to the baseline model with binary labor supply choices. The optimal labor supply solution is[28]

$$\ell_{it} = \arg \max_{\ell \in \{0,p,1\}} \left\{ \widetilde{V}_{t\ell}\left(\widetilde{X}_{it}\right) + \gamma_{it}\left(\mathbf{1}\left\{\ell = 1\right\} + \varrho_{it}\mathbf{1}\left\{\ell = p\right\}\right)\right\} \tag{28}$$

[27] In this sense our main goal is to test the robustness of our model to inclusion of part time work rather than explain part time work per se.

[28] Note now the $\widetilde{X}_{it}$ includes the health status S_{it} defined previously.

where

$$\left\{\mathcal{C}_{itp}\left(\widetilde{X}_{it}\right),\mathcal{I}_{itp}\left(\widetilde{X}_{it}\right),\mathcal{SSA}_{itp}\left(\widetilde{X}_{it}\right)\right\}$$
$$\equiv \underset{c,I,ssa}{\operatorname{argmax}}\left\{\psi_{tM_{it}}\frac{c^{1-\eta_c}}{1-\eta_c}+\varpi_t\left(ssa\right)+\beta E\left[V_{t+1}\left(X_{t+1}\right)\middle|\,\widetilde{X}_{it},c,\ell_{it}=p,I,ssa\right]\right\} \tag{29}$$

$$\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)\equiv\psi_{tM_{it}}\frac{\left(\mathcal{C}_{itp}\left(\widetilde{X}_{it}\right)\right)^{1-\eta_c}}{1-\eta_c}+\varpi_t\left(\mathcal{SSA}_{itp}\left(\widetilde{X}_{it}\right)\right)$$
$$+\beta E\left[V_{t+1}\left(X_{it+1}\right)\middle|\,\widetilde{X}_{it},\mathcal{C}_{itp}\left(\widetilde{X}_{it}\right),\ell_{it}=p,\mathcal{I}_{itp}\left(\widetilde{X}_{it}\right),\mathcal{SSA}_{t,p}\left(\widetilde{X}_{it}\right)\right] \tag{30}$$

The details of solving this model depend on the values of the three values of $\widetilde{V}_{t\ell}\left(\widetilde{X}_{it}\right)$ as well as ϱ_t. We discuss the details in Appendix A. In the case in which all three options may be chosen,[29] the model is like an ordered probit with cutoffs

$$\ell_t=\begin{cases}0, & \varepsilon_{it}<\varepsilon^*_{t1}\left(\widetilde{X}_{it}\right),\\ p, & \varepsilon^*_{t1}\left(\widetilde{X}_{it}\right)<\varepsilon_{it}<\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right),\\ 1, & \varepsilon_{it}>\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)\end{cases}$$

where

$$\varepsilon^*_{t1}\left(\widetilde{X}_{it}\right)=\log\left(\frac{\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right)-\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)}{\varrho_t}\right)-a_{i0}-\sum_{j=1}^{2}a_j\mathbf{1}\left\{M_{it}=j\right\}$$
$$\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)=\log\left(\frac{\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)-\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)}{1-\varrho_t}\right)-a_{i0}-\sum_{j=1}^{2}a_j\mathbf{1}\left\{M_{it}=j\right\}.$$

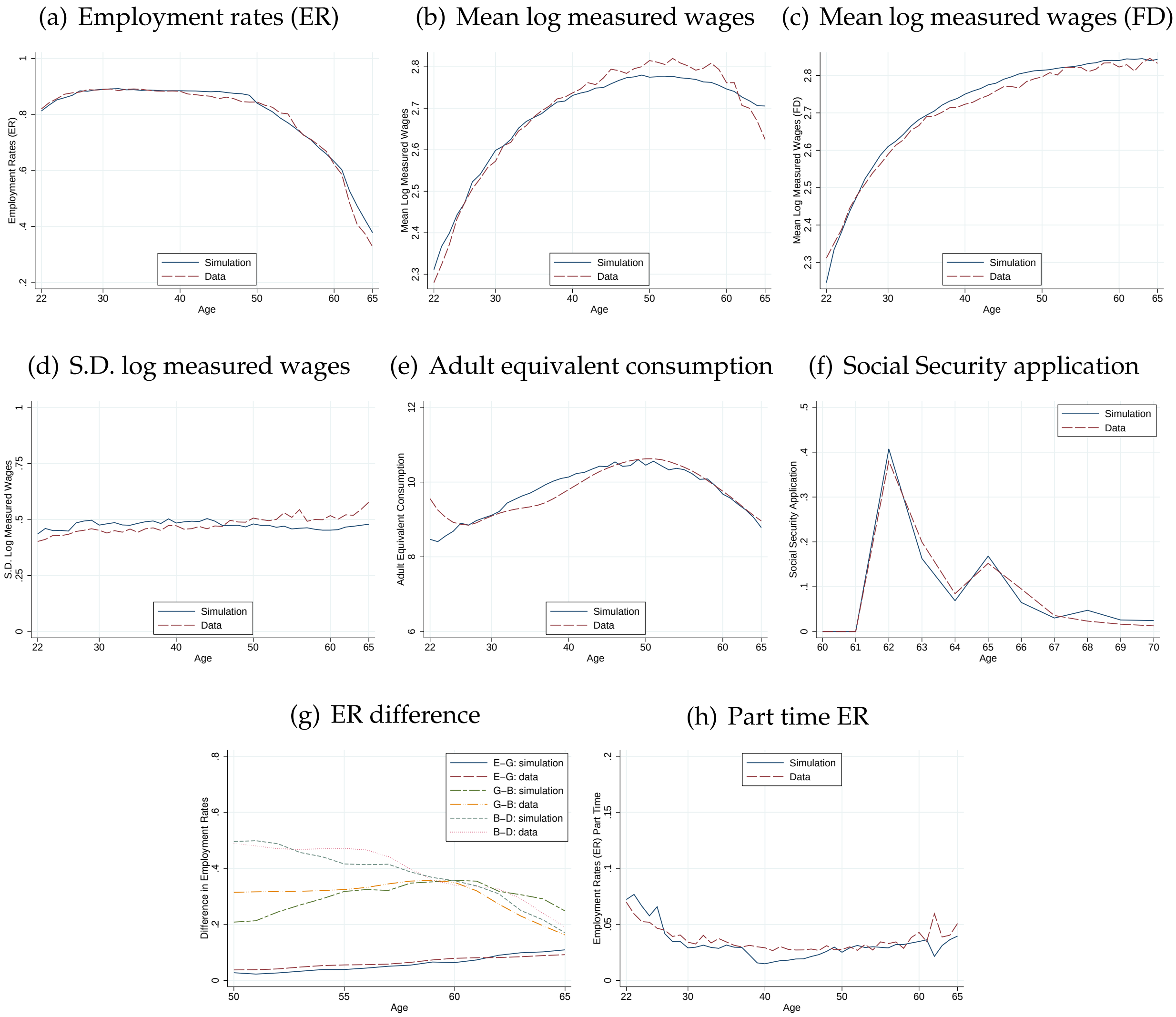
The expected value function still has a closed form, but it is complicated and given in Appendix A.

6.3 Estimation and Investigation

Parameter estimates from our extended model with health, disability and a part time option are presented in Table 5. The fit of the model is presented in Figure 6 and Table

[29] There are combination of parameters, state variables, and tastes in which part time would not be chosen for any realization of ε_{it}.

Figure 6: Fit of the extended model with health, disability and part time option



(a) Employment rates (ER) (b) Mean log measured wages (c) Mean log measured wages (FD)

(d) S.D. log measured wages (e) Adult equivalent consumption (f) Social Security application

(g) ER difference (h) Part time ER

2. Including health and part time leads to a similar fit for the base profiles (a)–(f) and the additional moments for health in (g) and part time enrollment in (h) fit well. We fit both the age profile of the relationship between health and labor supply as well as the slight increase in part time work prior to retirement.[30]

We conduct two sets of experiments to investigate the importance of health and the part time option on life-cycle labor supply. First, as either health or human capital could potentially explain retirement,[31] to control for health we simulate a counterfactual in which there was no health change. We eliminate the importance of health for individ-

[30]This is something Iskhakov and Keane (2020) had trouble matching.

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

Table 5: Estimates in the extended model with health, disability and part time option[a]

Parameters		Estimates	S.E.
HC depreciation[b]	δ	0.094	(0.001)
HC production function: I factor	α_I	0.119	(0.004)
HC production function: H factor	α_H	0.093	(0.003)
Standard deviation of HC innovation	σ_ξ	0.048	(0.002)
Consumption: CRRA	η_c	3.960	(0.020)
Consumption shifter: coef on t $(\times 10)$	φ_1	0.195	(0.007)
Consumption shifter: coef on t^2 $(\times 10^2)$	φ_2	0.103	(0.003)
Consumption shifter: coef on t^3 $(\times 10^3)$	φ_3	-0.033	(0.001)
Consumption shifter: coef on married	φ_4	1.704	(0.061)
Leisure: standard deviation of Shock	σ_ε	0.453	(0.013)
Leisure: spouse not working	a_1	-0.259	(0.011)
Leisure: spouse working	a_2	-0.694	(0.025)
Leisure: excellent health	a^0_{he}	-0.257	(0.011)
Leisure: bad health	a^0_{hb}	0.270	(0.016)
Leisure: bad health time trend	a^1_{hb}	0.011	(0.001)
Leisure: disabled	a^0_{hd}	2.947	(0.073)
Leisure: disabled time trend	a^1_{hd}	0.016	(0.001)
Part time utility: constant	a^0_ϱ	-1.315	(0.028)
Part time utility: coef on t $(\times 10)$	a^1_ϱ	0.251	(0.006)
Part time utility: coef on t^2 $(\times 10^2)$	a^2_ϱ	-0.013	(0.001)
Bequest weight	b_1	27,450,042	(1,039,117)
Parameter heterogeneity[b]			
Leisure: mean of intercept	μ_{a_0}	-5.871	(0.064)
Leisure: standard deviation of intercept	σ_{a_0}	2.308	(0.066)
HC productivity, mean	μ_π	1.826	(0.025)
HC productivity, standard deviation	σ_π	0.642	(0.023)
Correlation between a_0 & π	ρ	-0.562	(0.027)
Initial human capital level at age 18			
Intercept	γ_0	1.732	(0.099)
Coefficient on a_0	γ_{a_0}	0.234	(0.014)
Coefficient on π	γ_π	1.061	(0.062)
Standard deviation of error term	σ_{H_0}	0.013	(0.011)
Additional Social Security Application effects			
Effect of resource constraint $(\times 10^3)$	b_{62}	0.327	(0.040)
Effect of health insurance: constant $(\times 10^3)$	b_{65}	0.012	(0.002)
Effect of health insurance: coef on t $(\times 10^3)$	b_{65t}	0.306	(0.034)
χ^2 Statistic = 857[c]		Degrees of freedom = 291	

[a]Indirect inference estimates. Estimates use a diagonal weighting matrix. Standard errors are given in parentheses. HC: Human Capital.

[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 value of the χ^2 distribution is $\chi^2_{(291,0.01)} = 350$.

uals over 50 in two different ways—1) we do not allow their health to worsen and 2) we eliminate the interaction between health and preferences for work. Specifically, the first experiment restricts the health status an individual had at age 50 to remain for the rest of their life. In addition to fixing the health status at age 50, for individuals with bad/disabled health status on and after age 50, the second experiment assumes that their taste for leisure does not change with age. That is, letting t^* be the time period when the individual turns 50, we assume that the taste for leisure is now

$$\gamma\left(a_0, M, t, S, \varepsilon\right) = exp\left(a_0 + a_M + a^0_{hS} + a^1_{hS} \cdot 50 + \varepsilon_{it}\right) \tag{31}$$

and $S_{it} = S_{it^*}$ for $t > 50$. We then re-solve the modified model and simulate the life-cycle profile for each individual using the same estimates from the aforementioned extended model with health and the part time option.[32] The labor supply profiles of these two experiments are plotted as the long-dashed lines in Figure 7. If the health condition does not change with age, workers do supply more labor, in both experiments. The average difference in labor supply between the first counterfactual and the extended model with health is 15.5%. When we assume the taste of leisure does not vary with age for all health status, the difference in labor supply between the second counterfactual and the extended model with health is only slightly larger, 16.9%. Therefore, the main feature driving the results is health status itself not the parameterization of the utility function. Overall, these experiments imply that in our extended model health is a factor influencing retirement, but not the primary driver. 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.

The relatively small effect comes from the fact that bad health is relatively uncommon, not from the fact that it doesn't affect retirement. To see this, we show that at the individual (as opposed to aggregate) level, disability does induce an immediate and permanent decline in labor supply. We do this by assuming a worker's health status become excellent (or good/band/disabled) permanently at age 50. Similar labor supply profiles are plotted in Figure 7. These counterfactuals illustrate that upon becoming disabled, which is permanent, most workers will retire immediately and permanently.

In the second set of experiments, we investigate the effect of having a part time option

[32] 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.

Figure 7: Sensitivity to heath preferences: employment rates

(a) Health status no change after age 50

(b) Health status and taste of leisure no change after age 50

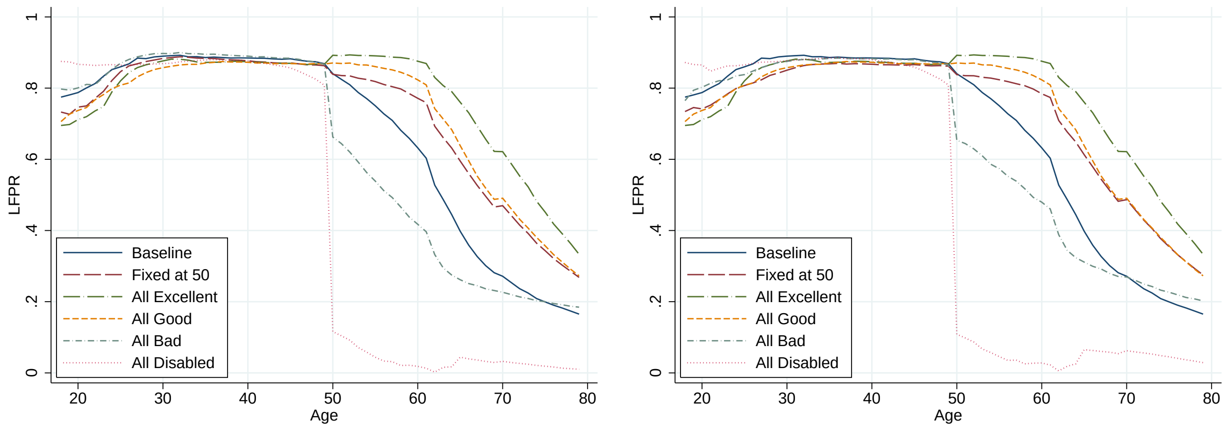
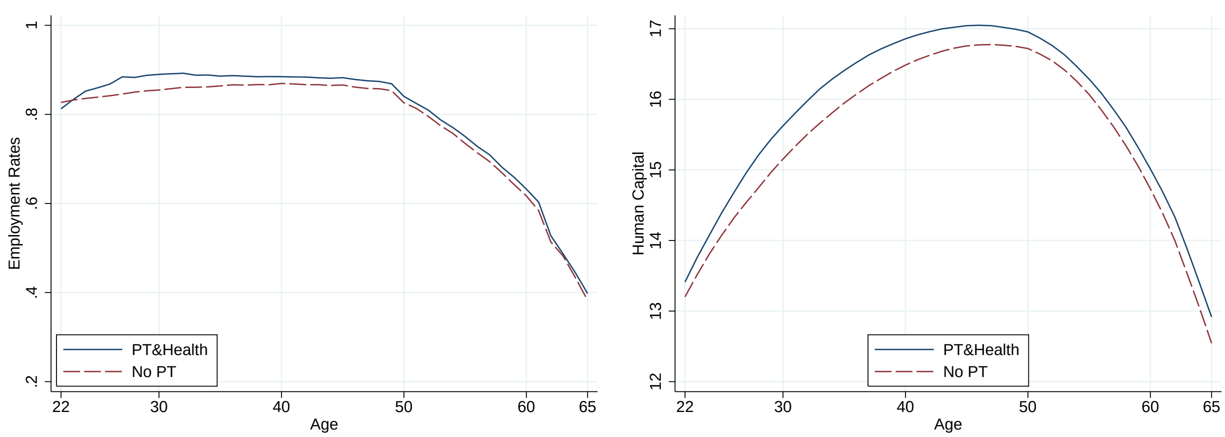


Figure 8: Sensitivity to part time option: turn off the part time option

(a) Employment rates

(b) Human Capital levels



by simulating a counterfactual that removes the option. Figure 8 presents the profiles of labor supply and the human capital. It appears that removing the part time option does not change the retirement pattern significantly, suggesting that the more flexible labor supply arrangement is not a major factor in understanding retirement.

7 Changes in Tax and Social Security

The preceding sections show that the model fits the life-cycle profiles of labor supply and log measured wages in the data well. In this section, we use the model to predict how changes in the tax or Social Security systems 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 tax and Social Security rules.

The policy experiments are

i) Increase taxes proportionally by 50%
ii) Eliminate Social Security earnings test
iii) Increase the Normal Retirement Age to 67
iv) Reduce Social Security benefits by 20%
v) Eliminate Social Security taxes
vi) Eliminate Social Security benefits
vii) Eliminate Social Security system (both taxes and benefits)

It is important to recognize that we are focusing on men with exactly 12 years of education. A full evaluation would require incorporating the other demographic groups as well.

The results of these experiments are summarized in Columns 1–7 of Table 6, where Panel A is for the extended model with health and part time work and Panel B is for the baseline model. All numbers are summations or averages throughout the life cycle (from age 18 to 80). Since we put more credibility on the extended model we focus on Panel A.[33]

Column 1 shows the result from the first experiment. 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 extended model with health and part time, the income effect dominates the substitution effect and an average individual works six-and-a-half additional months over the life cycle, equivalent to 1.3% of the total lifetime labor supply.[34] Most novel is the effect on human capital

[33]To make these two models comparable, we keep the SSDI and SSI benefit unchanged in all experiments.

[34]In our model leisure and consumption are separable. In the simplest static form of that model without

Table 6: Effects of changing taxes or Social Security rules

	0	1		2		3		4		5		6		7	
	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: Extended Model with Health and Part Time Option															
Labor Supply	41.910	0.542	1.293	0.123	0.293	0.489	1.167	0.495	1.182	-1.177	-2.809	2.464	5.880	0.698	1.664
LS (FT eqv)[d]	0.658	0.008	1.199	0.002	0.322	0.007	1.048	0.008	1.271	-0.018	-2.795	0.042	6.346	0.014	2.193
LS (PT)[e]	0.032	0.003	10.434	-0.003	-8.074	0.002	7.652	-0.001	-3.275	-0.001	-2.705	0.020	62.722	0.018	57.399
LS (FT)	0.642	0.006	0.968	0.003	0.532	0.006	0.883	0.009	1.385	-0.018	-2.797	0.032	4.933	0.005	0.809
Effective Labor	0.601	0.008	1.301	0.002	0.355	0.007	1.125	0.009	1.419	-0.016	-2.673	0.039	6.484	0.014	2.404
Pre-tax Income	10.359	0.183	1.769	0.013	0.129	0.096	0.927	0.120	1.162	-0.373	-3.598	0.672	6.484	0.148	1.426
Average lnw	2.664	0.005	0.195	0.003	0.100	0.004	0.158	0.003	0.099	-0.004	-0.138	0.013	0.491	-0.020	-0.769
Human Capital	14.014	0.148	1.059	0.010	0.072	0.077	0.553	0.077	0.551	-0.333	-2.373	0.596	4.255	0.194	1.385
Investment	0.056	0.0001	0.115	0.0001	0.032	0.0001	0.224	0.0002	0.307	-0.002	-4.098	0.003	4.876	0.0001	0.031
Consumption	8.985	-0.268	-2.978	0.001	0.013	-0.003	-0.038	-0.055	-0.610	0.442	4.919	-0.337	-3.754	0.117	1.308
Panel B: Baseline Model															
Labor Supply	41.278	0.948	2.296	0.210	0.508	0.422	1.022	0.721	1.748	-3.241	-7.852	4.900	12.871	0.287	0.696
Effective Labor	0.613	0.014	2.287	0.003	0.504	0.006	1.037	0.011	1.773	-0.047	-7.733	0.073	11.965	0.006	0.904
Pre-tax Income	10.225	0.405	3.959	0.056	0.551	0.125	1.223	0.225	2.201	-1.064	-10.403	1.412	13.813	0.065	0.639
Average lnw	2.604	0.025	0.947	0.005	0.180	0.005	0.207	0.012	0.464	-0.062	-2.389	0.045	1.729	-0.008	-0.318
Human Capital	13.957	0.324	2.318	0.049	0.353	0.098	0.699	0.170	1.221	-0.931	-6.672	1.033	7.404	0.122	0.877
Investment	0.042	0.001	2.415	0.0002	0.578	0.0003	0.804	0.001	1.381	-0.004	-9.583	0.004	10.513	0.001	2.305
Consumption	9.000	-0.230	-2.551	0.004	0.050	-0.027	-0.301	-0.037	-0.416	0.245	2.724	-0.147	-1.636	0.057	0.634

[a]The "Level" column refers to the annual value averaged over the whole life cycle, except the "Labor Supply" which is the total number of years worked over the whole life cycle. For example, in the baseline model, the total labor supply is 41.278 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 labor supply is 0.210 years higher than that in the baseline model across the whole life cycle from 18 to 80.

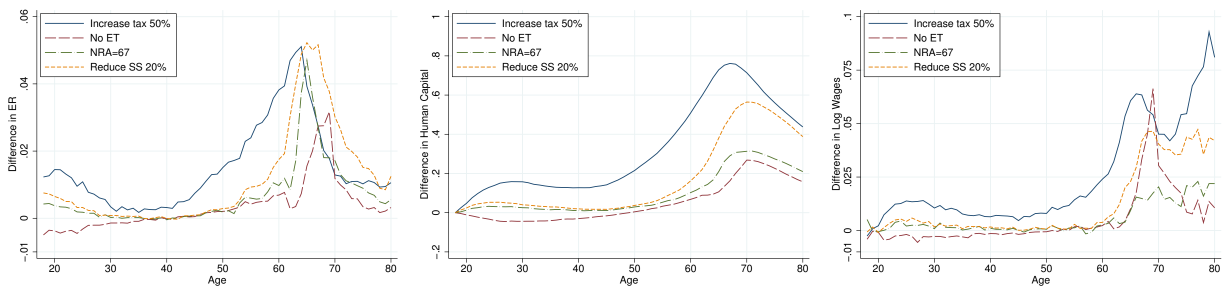
[c]The "%Δ" column refers to the percentage of the difference in the "ΔLevel" column relative to the level in the baseline model. For example, in the "No Earnings Test" case, the labor supply increases by 0.210 years which is equivalent to 0.508% of the labor supply in the baseline model.

[d]The "LS (FT eqv)" is the full time equivalent labor supply where working part time is counted as 0.5.

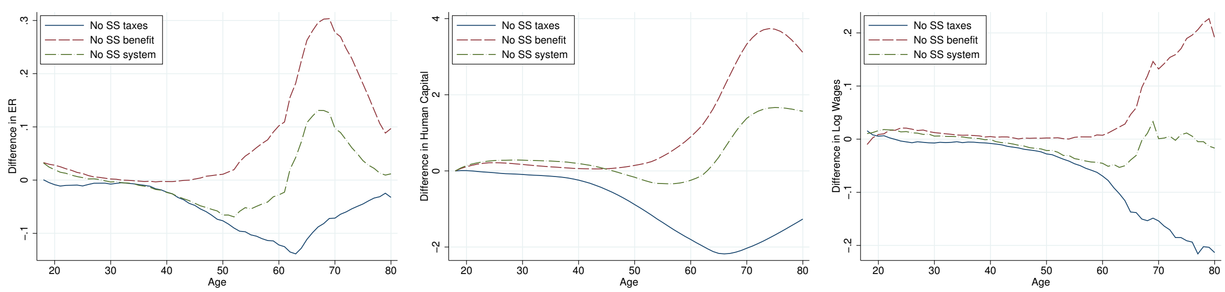
[e]The "LS (PT)" refers to the labor supply of part time workers.

Figure 9: Policy experiments: change taxes or Social Security benefits

(a) Change taxes or Social Security benefits: difference in employment rates (left), human capital (middle), and log measured wages(right)



(b) Remove Social Security taxes or benefits: difference in employment rates (left), human capital (middle), and log measured wages (right)



investment which increases by 0.1%, leading to 1.1% increase in the human capital level and 0.5% increase in the measured wages. 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 effective labor increases by six months or 1.3% and the pre-tax income increases by almost 2%.[35] Annual consumption reduces by 3%. Note that these effects are averaged across the life cycle. Figure 9(a) shows how the effects change at different ages.

The manner in which Social Security rules affect labor supply and wages is of central interest to policy makers. The remaining six experiments are devoted to answering these questions. In the first three we manipulate the current Social Security rules (Columns 2–4) while in the last three we decompose the distortionary effects of the current Social Security system (Columns 5–7).

human capital, whether the income or substitution effect dominates depend on whether η_c is larger or smaller than one. We estimate η_c to be around 4 which is well within the estimates in the literature, so it is not surprising that the income effect dominates the substitution effect.

[35] 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.

First we remove the Social Security earnings test, which is effective between age 62 and 70. In the second one, we delay Normal Retirement Age (NRA) by two years: the new NRA is age 67 in this counterfactual experiment while it is age 65 in the baseline model.[36] In the third one, we reduce the Social Security benefit proportionally by 20%. The results are presented in Columns 2–4 in Table 6 and in Figure 9(a). Removing the Social Security earnings test between ages 62 and 70 has a smaller effect on all variables; delaying the normal retirement age by two years, has a larger impact, and reducing the generosity of the Social Security benefit has a similar magnitude to delaying the normal retirement age (though the effects are slightly higher). For instance, in the extended model they increase the labor supply by roughly one-and-a-half, six, and six months, respectively. One important feature is that while the largest changes in the labor supply happen later in the life cycle when the policy change is directly effective, the policy influences choices over the whole life cycle, as indicated in Figure 9(a). For instance, delaying the Normal Retirement Age or universally reducing Social Security benefit induce substantially higher employment rates as well as more human capital investment, resulting persistently higher human capital levels and therefore higher wages at old ages. As seen in the three panels of Figure 9(a), the wage difference is negligible before age 60 but increases substantially after that, reaching 2% and 5%, respectively, around age 67. Our results are echoed in Gohl et al. (2020) who estimate a related effect directly and find that employed women aged 53–60 increase their human capital investment substantially when the early retirement age is increased from 60 to 63 in Germany. Ignoring such a human capital or wage response in experiments involving retirement policy will most likely introduce bias. The budget calculation in Table D3 shows that these three experiments reduce the Social Security deficit by 0.8%, 26.5%, and 43.2%, respectively.

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 5 of Table 6 we keep the Social Security benefit but eliminate the Social Security taxes (the payroll taxes);[37] in Column 6 we remove the Social Security benefit completely but keep the Social Security taxes; in Column 7 we remove the entire Social Security system, that is, both the Social Security taxes and the benefit. Removing the Social Security taxes in the extended model induces an average individual to supply 2.8% less labor or fourteen months. This is not surprising because removing the Social Security taxes is essentially a universal cut in the tax rate. In our tax hike counterfactual,

[36] Note that when we do this we adjust the claiming "norm" captured in Equation (4) to 67 as well. We have also run this counterfactual without changing this. It yields qualitatively similar results but larger in magnitude.

[37] The income taxes are still effective.

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. The increase in the labor supply is 5.9%, which is higher than the 2.8% 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 7 indicates that eliminating the current Social Security system increases average labor supply by 1.7% or eight months over the life cycle. Such observation is also mentioned qualitatively in Gustman and Steinmeier (1986) and Rust and Phelan (1997). Figure 9(b) shows that the changes in the labor supply and log wages are most pronounced at old ages when either taxes or benefit is removed in the Social Security system.

Another point worth emphasizing is that in almost every policy counterfactual, the changes in the endogenously determined wage levels are substantial. This is especially true at old ages (between age 60 and 75):[38] as high as 2%–7% when removing the earnings test or delaying NRA by two years or reducing Social Security benefit, almost 15% when removing Social Security benefits, almost −20% when removing Social Security taxes, or −5% when removing the entire Social Security system. These are caused by changes in the human capital levels as a result of higher or lower investment. This makes the importance of endogenizing human capital clear. Ignoring the human capital investment channel would generate substantial bias in terms of predicting labor supply at old ages in similar experiments.

Panel B of Table 6 presents the results of experiments from the baseline model. The responses to the policy changes are qualitatively similar to the extended model across all experiments.

8 Conclusion

This paper develops and estimates a rich life-cycle model that merges a Ben-Porath style human capital framework with a neoclassical style model of endogenous labor supply and uses it to examine changes in the Social Security system. We use it to study what is typically referred to as retirement in the literature without treating retirement as fundamentally different than a no-work decision in either the data or the model. In the model, each individual chooses consumption, labor supply, human capital investment, and Social Security application. Investment in human capital generates wage growth over the life cycle, while depreciation of human capital is the main force generating a decline in

[38] The employment rate is very low after age 75 so the wage comparison is less interesting.

working for older workers. We show that the parsimonious model is able to fit the main features of life-cycle labor supply, measured wages (with and without fixed effects) as well as retirement. In particular we can fit both the large increase in measured wages and small changes in labor supply at the beginning of the life cycle along with the small changes in measured wages but large changes in labor supply at the end.

Despite the fact that our framework does not rely on age or time varying preference or production function parameters, our model is consistent with a rather small and empirically plausible labor supply elasticity that rises with age. To show the importance of depreciation in explaining the result we re-estimate the model without allowing depreciation on the job and show the model can not fit the data as well. We also estimate an extension of the model allowing for both health shocks and part time option. While these factors are relevant, they are not the main factors driving retirement. The model is also robust to several robustness checks in which we vary pre-set parameters.

We use the estimated model to simulate the impacts of various policy changes. While prior work typically takes the wage process as given and focuses on the retirement decision, we are able to model the effect of the policy change on the wage process and the labor supply decisions. As we show in our model, less generous Social Security benefits result in higher labor supply later in the life cycle, so workers adjust their investment over the life cycle. This results in a higher human capital level as well as higher labor supply earlier in the life cycle. The bottom line is that modeling labor supply and human capital decisions jointly is critical in an analysis of the effects of policy changes. While presumably other factors would be important for explaining other features of labor markets, endogenous labor supply is critical for understanding life-cycle human capital investment and life-cycle human capital investment is critical for understanding life-cycle labor supply.

References

Acemoglu, Daron and Jörn-Steffen Pischke, "Why Do Firms Train? Theory and Evidence*," *The Quarterly Journal of Economics*, 02 1998, *113* (1), 79–119.

_ **and** _ , "The structure of wages and investment in general training," *Journal of political economy*, 1999, *107* (3), 539–572.

Adda, Jérôme, Christian Dustmann, and Katrien Stevens, "The Career Costs of Children," *Journal of Political Economy*, 2017, *125* (2), 293–337.

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

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

Autor, David H and Mark G Duggan, "The growth in the social security disability rolls: a fiscal crisis unfolding," *Journal of Economic Perspectives*, 2006, *20* (3), 71–96.

Becker, Gary, "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, Oct. 1962, *70* (5), 9–49.

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.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathon Shaw, "Female Labour Supply, Human Capital and Welfare Reform," *Econometrica*, September 2016, *84* (5), 1705–1753.

_ **, Monica Costa-Dias, David Goll, and Costas Meghir**, "Wages, Experience and Training of Women over the Lifecycle," April 2019. Unpublished Manuscript.

Browning, Martin, 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, p. 73–132.

Fernández-Villaverde, Jesús and Dirk Krueger, "Consumption over the life cycle: Facts from consumer expenditure survey data," *The Review of Economics and Statistics*, 2007, *89* (3), 552–565.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry, "Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]," Technical Report, Minneapolis, MN 2021. https://doi.org/10.18128/D030.V9.0.

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.

Gohl, Niklas, Peter Haan, Elisabeth Kurz, and Felix Weinhardt, "Working Life and Human Capital Investment: Causal Evidence from Pension Reform," 2020. IZA Discussion Paper No. 12891.

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

Gourinchas, Pierre-Olivier and Jonathan A Parker, "Consumption over the life cycle," *Econometrica*, 2002, *70* (1), 47–89.

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," in N. Terleckyj, ed., *Household production and consumption*, Columbia University Press, 1975, p. 99–138.

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

Heckman, James J, Lance J Lochner, and Petra E Todd, "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond," *Handbook of the Economics of Education*, 2006, *1*, 307–458.

Heckman, James J., 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, p. 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.

Iskhakov, Fedor and Michael Keane, "Effects of taxes and safety net pensions on life-cycle labor supply, savings and human capital: The case of Australia," *Journal of Econometrics*, 2020.

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

Keane, Michael P and Kenneth I Wolpin, "The effect of parental transfers and borrowing constraints on educational attainment," *International Economic Review*, 2001, *42* (4), 1051–1103.

_ **and Nada Wasi**, "Labour supply: the roles of human capital and the extensive margin," *The Economic Journal*, 2016, *126* (592), 578–617.

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, *Schooling, Experience and Earnings*, New York: NBER Press, 1974.

_ **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.

Mulligan, Casey B, "Substitution over time: another look at life-cycle labor supply," *NBER macroeconomics annual*, 1998, *13*, 75–134.

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

Rosen, Sherwin, "Learning and Experience in the Labor Market," *The Journal of Human Resources*, Summer 1972, *7* (3), 362–342.

— , "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.

Sanders, Carl and Christopher Taber, "Life-Cycle Wage Growth and Heterogeneous Human Capital," *Annual Review of Economics*, 2012, *4* (1), 399–425.

Schmieder, Johannes F., Till von Wachter, and Stefan Bender, "The Effect of Unemployment Benefits and Nonemployment Durations on Wages," *American Economic Review*, March 2016, *106* (3), 739–77.

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: How large is the bias?," *Review of Economic Dynamics*, 2011, *14* (4), 577–591.

Appendix A Value Function Derivations

The following fact will be useful in this section:

If ε_{it} is normal with 0 expected value and variance of σ_ε^2 then

$$E\left(e^{\varepsilon_{it}} \mid \varepsilon_{it} \geq \varepsilon^*\right) = e^{\frac{\sigma_\varepsilon^2}{2}} \frac{\Phi\left(\sigma_\varepsilon - \frac{\varepsilon^*}{\sigma_\varepsilon}\right)}{1-\Phi\left(\frac{\varepsilon^*}{\sigma_\varepsilon}\right)} \tag{A1}$$

where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution.

We first solve the labor supply model and the value function for the baseline model and then for the model that includes part time option.

A.1 Baseline Model

For simplicity assume that

$$\widetilde{a}_{it} \equiv a_{i0} + a_1 \mathbf{1}\left\{M_{it}=1\right\} + a_2 \mathbf{1}\left\{M_{it}=2\right\} \tag{A2}$$

so that we can write

$$\gamma_{it} = exp\left(\widetilde{a}_{it} + \varepsilon_{it}\right) \tag{A3}$$

Then as we state in the text as long as $\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) > \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)$,

$$\begin{aligned} \ell_{it} =& \mathbf{1}\left\{\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + \gamma_{it} \geq \widetilde{V}_{t0}\left(\widetilde{X}_{it}\right)\right\} = \mathbf{1}\left\{\widetilde{a}_{it} + \varepsilon_{it} \geq \log\left(\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)\right)\right\} \\ =& \mathbf{1}\left\{\varepsilon_{it} \geq \varepsilon_{it}^*\left(\widetilde{X}_{it}\right)\right\} \end{aligned} \tag{A4}$$

where

$$\varepsilon_t^*\left(\widetilde{X}_{it}\right) \equiv \log\left(\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)\right) - \widetilde{a}_{it}. \tag{A5}$$

Note that if $\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) \leq \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)$ then the individual would never choose to work.

The only difference between X_{it} and $\widetilde{X}_{it}$ is that ε_{it} is included in X_{it}, so

$$\begin{aligned}
E\left[V_t\left(X_{it}\right)\middle|\widetilde{X}_{it}\right] &= Pr\left(\varepsilon_{it} < \varepsilon_t^*\left(\widetilde{X}_{it}\right)\right) E\left[\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right)\middle|\widetilde{X}_{it}, \varepsilon_{it} < \varepsilon_t^*\left(\widetilde{X}_{it}\right)\right] \\
&\quad + Pr\left(\varepsilon_{it} \geq \varepsilon_t^*\left(\widetilde{X}_{it}\right)\right) E\left[\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + e^{\widetilde{a}_{it}} e^{\varepsilon_{it}}\middle|\widetilde{X}_{it}, \varepsilon_{it} \geq \varepsilon_t^*\left(\widetilde{X}_{it}\right)\right] \\
&= \Phi\left(\frac{\varepsilon_t^*\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) + \left(1-\Phi\left(\frac{\varepsilon_t^*\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)\right)\left[\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + e^{\widetilde{a}_{it}+\frac{\sigma_\varepsilon^2}{2}} \frac{\Phi\left(\sigma_\varepsilon - \frac{\varepsilon_t^*\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)}{1-\Phi\left(\frac{\varepsilon_t^*\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)}\right]
\end{aligned}$$

A.2 Part Time

We continue to use the simplified expression for $\widetilde{a}_{it}$ as defined above and including the health,

$$\widetilde{a}_{it} \equiv a_{i0} + a_1 \mathbf{1}\left\{M_{it}=1\right\} + a_2 \mathbf{1}\left\{M_{it}=2\right\} + \sum_{j=1}^{4} \mathbf{1}\left\{S_{it}=j\right\}\left(a_{sj}^0 + a_{sj}^1 t\right) \tag{A6}$$

First consider the labor supply decision. In this case

$$\ell_t = \begin{cases} 0 & \widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) > \max\left\{\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) + \gamma_{it}\varrho_t,\ \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + \gamma_{it}\right\} \\ p & \widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) + \gamma_{it}\varrho_t > \max\left\{\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right),\ \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + \gamma_{it}\right\} \\ 1 & \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + \gamma_{it} > \max\left\{\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right),\ \widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) + \gamma_{it}\varrho_t\right\} \end{cases} \tag{A7}$$

or

$$\ell_t = \begin{cases} 0, & \gamma_{it} < \min\left\{\frac{\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)}{\varrho_t},\ \widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)\right\} \\ p, & \frac{\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)}{\varrho_t} < \gamma_{it} < \frac{\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)}{1-\varrho_t}, \\ 1, & \gamma_{it} > \max\left\{\frac{\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)}{1-\varrho_t},\ \widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)\right\} \end{cases} \tag{A8}$$

One can see that there are a number of different cases to consider. Since γ_{it} is log normal, ties will be irrelevant so we abstract from them.

The most interesting case is that in which

$$0 < \frac{\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)}{\varrho_t} < \widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) < \frac{\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)}{1-\varrho_t} \tag{A9}$$

as this is the only case where all three possibilities happen with positive probability.

In this case

$$\ell_t = \begin{cases} 0, & \varepsilon_{it} < \varepsilon^*_{t1}\left(\widetilde{X}_{it}\right), \\ p, & \varepsilon^*_{t1}\left(\widetilde{X}_{it}\right) < \varepsilon_{it} < \varepsilon^*_{t2}\left(\widetilde{X}_{it}\right), \\ 1, & \varepsilon_{it} > \varepsilon^*_{t2}\left(\widetilde{X}_{it}\right) \end{cases} \tag{A10}$$

where

$$\varepsilon^*_{t1}\left(\widetilde{X}_{it}\right) = \log\left(\frac{\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)}{\varrho_t}\right) - \widetilde{a}_{it} \tag{A11}$$

$$\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right) = \log\left(\frac{\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) - \widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)}{1-\varrho_t}\right) - \widetilde{a}_{it} \tag{A12}$$

To derive the value function recall that for any random variable Y

$$E\left(Y \mid a \le Y < b\right) = \frac{E\left(Y \mid Y \ge a\right) Pr\left(Y \ge a\right) - E\left(Y \mid Y \ge b\right) Pr\left(Y \ge b\right)}{Pr\left(a \le Y < b\right)}$$

Using this and the log normal result, in this case with all three possibilities (i.e. A9)

$$\begin{aligned}
& E\left[V_t\left(X_{it}\right) \middle| \widetilde{X}_{it}\right] \\
=& Pr\left(\varepsilon_{it} < \varepsilon^*_{t1}\left(\widetilde{X}_{it}\right)\right) E\left[\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) \middle| \widetilde{X}_{it}, \varepsilon_{it} < \varepsilon^*_{t1}\left(\widetilde{X}_{it}\right) 0\right] \\
& + Pr\left(\varepsilon^*_{t1}\left(\widetilde{X}_{it}\right) < \varepsilon_{it} \le \varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)\right) E\left[\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) + \varrho_t e^{\widetilde{a}_{it}} e^{\varepsilon_{it}} \middle| \widetilde{X}_{it}, \varepsilon^*_{t1}\left(\widetilde{X}_{it}\right) < \varepsilon_{it} \le \varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)\right] \\
& + Pr\left(\varepsilon_{it} \ge \varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)\right) E\left[\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + e^{\widetilde{a}_{it}} e^{\varepsilon_{it}} \middle| \widetilde{X}_{it}, \varepsilon_{it} \ge \varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)\right] \\
=& \Phi\left(\frac{\varepsilon^*_{t1}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right) \widetilde{V}_{t0}\left(\widetilde{X}_{it}\right) + \left[\Phi\left(\frac{\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right) - \Phi\left(\frac{\varepsilon^*_{t1}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)\right] \cdot \\
& \cdot \left[\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right) + \varrho_t e^{\widetilde{a}_{it} + \frac{\sigma_\varepsilon^2}{2}} \frac{\Phi\left(\sigma_\varepsilon - \frac{\varepsilon^*_{t1}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right) - \Phi\left(\sigma_\varepsilon - \frac{\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)}{\Phi\left(\frac{\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right) - \Phi\left(\frac{\varepsilon^*_{t1}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)}\right] \\
& + \left(1 - \Phi\left(\frac{\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)\right) \left[\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right) + e^{\widetilde{a}_{it} + \frac{\sigma_\varepsilon^2}{2}} \frac{\Phi\left(\sigma_\varepsilon - \frac{\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)}{1 - \Phi\left(\frac{\varepsilon^*_{t2}\left(\widetilde{X}_{it}\right)}{\sigma_\varepsilon}\right)}\right]
\end{aligned} \tag{A13}$$

The other possibilities are special cases of this for which fewer than three possibilities are possible.

If

$$\frac{\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)-\widetilde{V}_{t1}\left(\widetilde{X}_{it}\right)}{1-\varrho_t}<\frac{\widetilde{V}_{t0}\left(\widetilde{X}_{it}\right)-\widetilde{V}_{tp}\left(\widetilde{X}_{it}\right)}{\varrho_t} \tag{A14}$$

then part time is not an option and we return to the basic model. In other cases, other options will disappear which simplifies the expression.

Appendix B Demographics

B.1 Family Status and Spousal Income

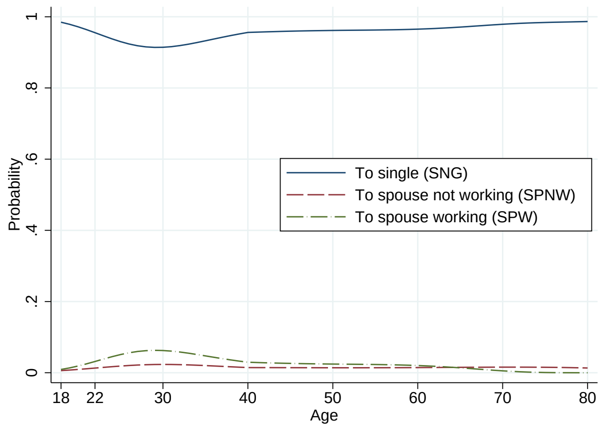
There are three family statuses: single or divorced (SNG), married with spouse not working (SPNW), and married with spouse working (SPW). Figure B1 plot the transition probabilities (Panels (a)–(c)) at each age. The resulting distribution (Panel (d)) is similar to patterns in the SIPP data and in the CPS data. For those married with spouse working, Panel (e) of Figure B1 plots the estimated mean ($\mu_{\zeta t}$) and standard deviation ($\sigma_{\zeta t}$) of the logarithm of the spousal income.

B.2 Health Status

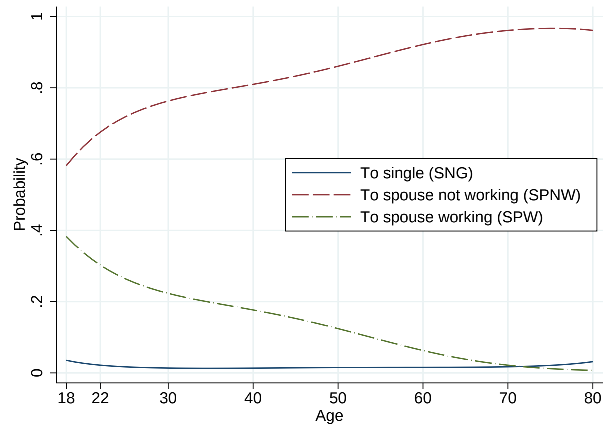
There are four health status: excellent, good, bad, and disabled. The disability state is absorbing. Figure B2 plot the health transition probabilities (Panels (a)–(c)) and the distribution (Panel (d)) at each age.

Figure B1: Family status transitions, distribution, and spousal income

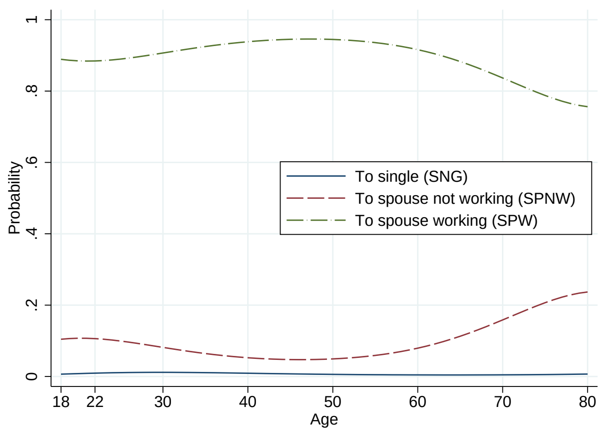
(a) From single (SNG)



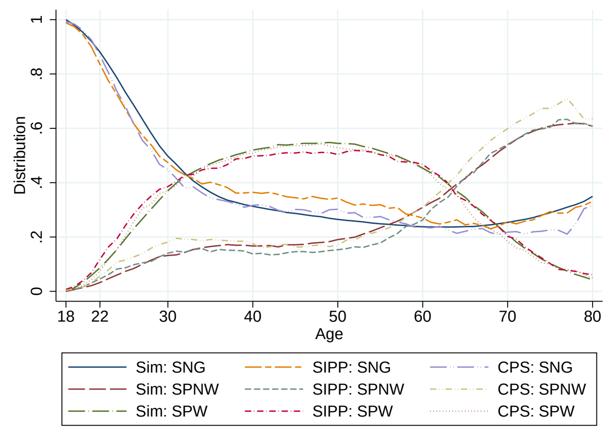
(b) From married and spouse not working (SPNW)



(c) From married and spouse working (SPW)



(d) Distribution



(e) Log of spousal income: mean and S.D.

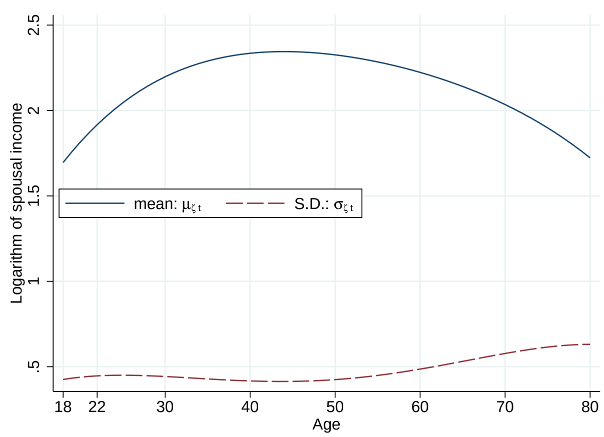
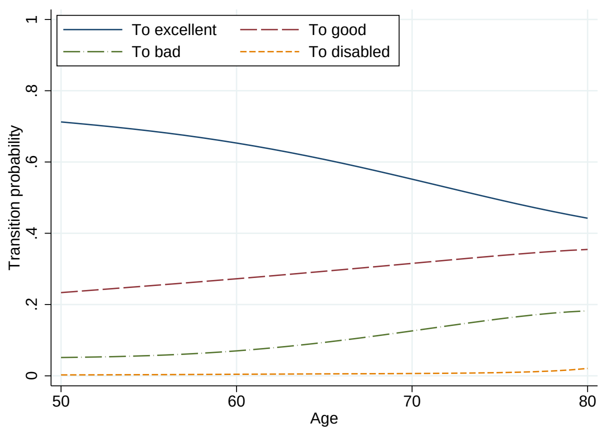
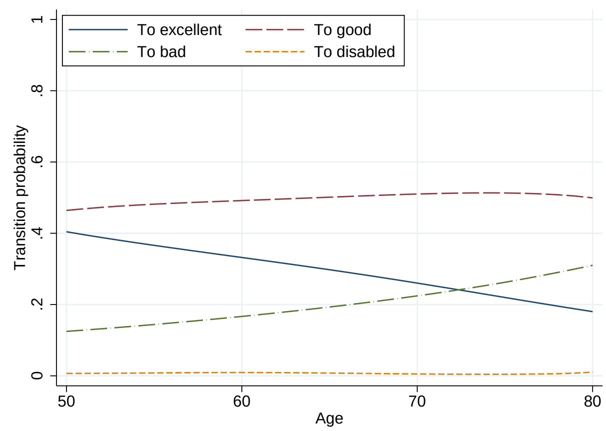


Figure B2: Health status transitions and distribution

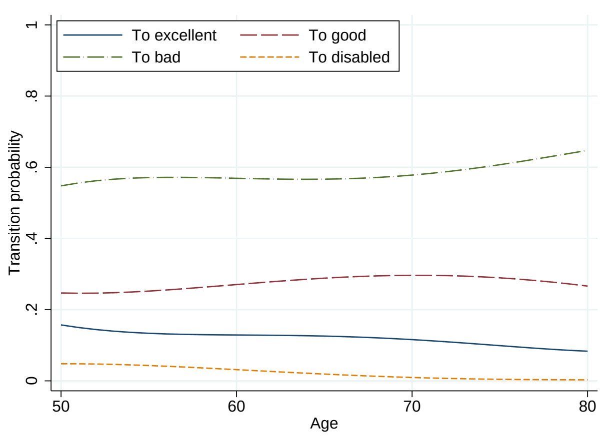
(a) From excellent health



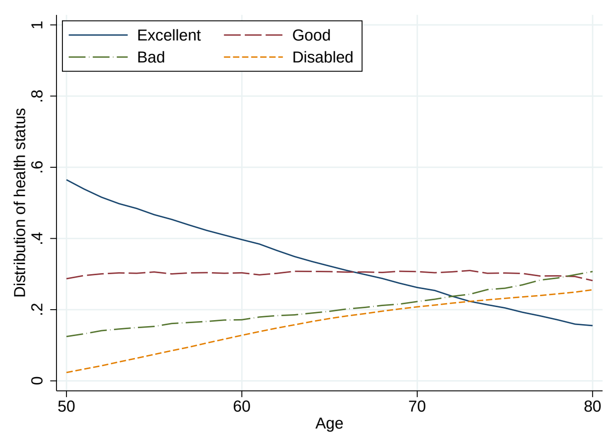
(b) From good health



(c) From bad health



(d) Health distribution



Appendix C Taxes and Social Security

We use tax codes and Social Security rules in the year of 2004, except earnings test where we follow the rules in 1999.[39]

There are two different kinds of taxes that the worker's wage income is subject to, namely the payroll taxes and the federal income taxes. We ignore state income taxes. The payroll taxes include the Social Security portion, 6.2% capped at $87,900, and the Medicare tax, 1.45% 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 deduction for head of household is $7,150. These all together generate the tax codes used in the paper in the following formula,

$$\mathrm{Y}_{it} = v\left(Y_{it}^{o} + ssb_{it}^{taxable}\right) + ssb_{it}^{pre} - ssb_{it}^{taxable} \tag{C1}$$

where Y_{it} is total net income, $Y_{it}^{o} = \max\{rA_{it}, 0\} + w_{it} + y_{it}$ is the gross income, ssb_{it}^{pre} is pre-tax Social Security benefit, $ssb_{it}^{taxable}$ is the taxable part of Social Security benefit, and v is the after tax income as a function of pre-tax income. It is presented in Table C1.

The pre-tax Social Security benefit amount, ssb_{it}^{pre}, is determined by the current age t, the age when the individual first starts receiving the benefit, t_{ssa}, and the entire earning history up to current age t. It is summarized in the following formula,

$$ssb_{it}^{pre} = \begin{cases} 0 & \text{if } ss_{it} = ssa_{it} = 0 \\ ssb_{it} - ET_{it} & \text{if } ss_{it} + ssa_{it} > 0 \;\&\; t < 70 \\ ssb_{it} & \text{if } ss_{it} + ssa_{it} > 0 \;\&\; t \geq 70 \end{cases} \tag{C2}$$

where

$$ssb_{it} = \begin{cases} \Pi_{j=t_{ssa}}^{t-1}\left(1 + DRC_{ij}^{ET}\right) \cdot PIA_{it} \cdot \left(1 - \frac{65 - t_{ssa}}{15}\right) & \text{if } 62 \leq t_{ssa} < 65 \\ \Pi_{j=t_{ssa}}^{t-1}\left(1 + DRC_{ij}^{ET}\right) \cdot PIA_{it} \cdot [1 + 0.06 \cdot (t_{ssa} - 65)] & \text{if } 65 \leq t_{ssa} < 70 \\ PIA_{it} \cdot [1 + 0.06 \cdot (69 - 65)] & \text{if } t_{ssa} \geq 70 \end{cases} \tag{C3}$$

is the eligible Social Security benefit and the DRC_{ij}^{ET} is the Delayed Retirement Credit due to benefits withdrew at the earnings test.

The remainder of the section describes how each component in Equations (C1)–(C3) is

[39] Before 2000, the earnings test applies to ages before 70. Since 2000, the earnings test is eliminated after reaching NRA. All information about Social Security benefits in this section is extracted from http://www.ssa.gov.

Table C1: Wage income tax codes (in 2004$).

Marginal Tax Rate	Pre-tax (Y)	Post-tax Income $\text{Y}_t = v_0 + v_1 (Y - v_2)$
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	$\geq 329,351$	$229,371.63 + 0.6355 (Y - 329,350)$

defined.

C.1 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 are equal to the Primary Insurance Amount (PIA), which is a function of Average Indexed Monthly Earnings (AIME),

$$\begin{aligned} PIA_{it} &= 0.9 \cdot \min \{bp_1, AIME_{it}\} + 0.32 \cdot \min \{bp_2 - bp_1, \max \{0, AIME_{it} - bp_1\}\} \\ &\quad + 0.15 \cdot \max \{0, AIME_{it} - bp_2\} \end{aligned} \tag{C4}$$

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

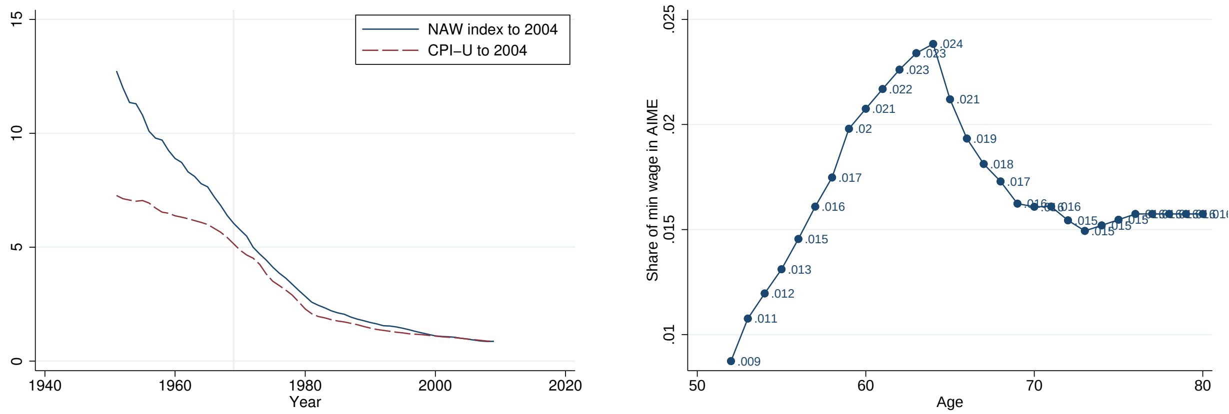
The AIME is computed as the monthly average earning of the 35 years with highest inflation-adjusted earnings. Only earnings subject to the Social Security tax are used in the calculation and therefore AIME is capped. The included earning in a specific year is adjusted for wage inflation by multiplying 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 C1, 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 entire annual earning history, which is computationally infeasible. Instead we apply an approximating

Figure C1: Wage Index, CPI, and minimum wage share in AIME

(a) Relative (to Year 2004) indices of National Average Wage Index and CPI-U

(b) Share of minimum wage in AIME, assuming starting working from age 16



method, taking into account the wage growth pattern over the life cycle,

$$AIME_{it+1} = AIME_{it} + \max\left\{0,\ \frac{sse_{it}}{35 \times 12} - share_{min}(t) \cdot AIME_{it}\right\} \tag{C5}$$

where $sse_{it} = \min\{H_{it}(1-\ell_{it})(1-I_{it}),\ \bar{sse}\}$ is included earning, capped at $\bar{sse} = \$87,900$. The $share_{min}$ is the share of minimum wage in AIME. Figure C1 plots the estimated $share_{min}(t)$ from the SIPP data for age 52 to 80, assuming starting working age of 18, and $share_{min}(t < 52)$=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. The delayed retirement credit (DRC) of 6% is given to the applicant for each delayed year up to age 69.[40] No DRC is given for applicants at age 70 or older.

C.2 The Social Security Earnings Test

We use the Social Security earnings test rules in 1999. The Social Security benefits could be withheld partly or totally if the worker is earning income while taking the Social

[40] 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.

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). The following formula summarizes the earnings test,

$$ET_{it} = \begin{cases} \min\left\{ssb_{it},\ \max\left\{0,\ \frac{Y^o_{it}-10885}{2}\right\}\right\} & \text{if } ss_{it} + ssa_{it} > 0 \ \& \ 62 \leq t < 65 \\ \min\left\{ssb_{it},\ \max\left\{0,\ \frac{Y^o_{it}-17575}{3}\right\}\right\} & \text{if } ss_{it} + ssa_{it} > 0 \ \& \ 65 \leq t < 70 \end{cases} \tag{C6}$$

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 but not actuarially fair for individuals aged 65 or older. This is summarized by the following formula,

$$DRC^{ET}_{it} = \frac{ET_{it}}{ssb^e_{it}} \times \begin{cases} 0.067 & \text{if } ss_{it} + ssa_{it} > 0 \ \& \ 62 \leq t < 65 \\ 0.06 & \text{if } ss_{it} + ssa_{it} > 0 \ \& \ 65 \leq t < 70 \end{cases}$$

C.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. Defining $\widetilde{Y}_{it} = Y^o_{it} + \frac{1}{2} ssb_{it}$, the taxable part of Social Security benefits is calculated as

$$ssb^{taxable}_{it} = \begin{cases} 0, & \text{if } \widetilde{Y}_{it} \leq 25000 \\ \min\left\{0.85 \cdot ssb_{it}, \frac{1}{2}\min\left\{ssb_{it}, \widetilde{Y}_{it} - 25000, 9000\right\} \right. & \text{otherwise} \\ \qquad \left. + \ 0.85 \max\left\{0, \widetilde{Y}_{it} - 34000\right\}\right\} & \end{cases} \tag{C7}$$

C.4 Disability Benefits

The Social Security Disability Insurance (SSDI) and the Supplemental Security Income (SSI) programs are the two largest Federal programs that provide assistance to people with disabilities.

To be qualified for the SSDI, workers cannot earn employment income higher than the disability thresholds, namely the Substantial Gainful Activity (SGA), which are \$810 per month for nonblind persons and \$1350 for blind persons in 2004. We use the SGA of \$810. Before NRA, the SSDI benefit is based on AIME. Upon reaching the NRA, SSDI benefits are automatically converted to the normal Social Security benefits and are not subject to the SGA earnings restriction anymore. Workers with disabilities may also receive the Supplemental Security Income (SSI). SSDI and SSI combined guarantee a minimum monthly benefit of \$564. The following formula summarizes the disability benefits,

$$SSDI_{it} + SSI_{it} = \begin{cases} 0 & \text{if } t < 65 \text{ \& } Y^o_{it} \geq 9720 \\ \max\{6768,\ PIA_{it}\} & \text{if } t < 65 \text{ \& } Y^o_{it} < 9720 \\ ssb_{it} & \text{if } t \geq 65 \end{cases} \tag{C8}$$

C.5 Practical Implementation

We adjust the state variables $AIME_{it}$ (relevant for $ss_{it} = 0$) and ssb_{it} (relevant for $ss_{it} = 1$) to reflect all aforementioned factors which affects current or future Social Security benefits, including the Early Retirement, the Delayed Retirement Credit, and the benefit increase due to earnings test. Specifically, we use the following formula when numerically solving the life-cycle model.

Prior to claiming, $AIME_{it}$ is updated according to Equation (C5). Note that this is not exact but is an approximation.

In the first year that an individual begins to claim (when $ss_{it} = 0$, $ssa_{it} = 1$), the initial ssb_{it} is calculated as

$$ssb_{it} = PIA_{it} \cdot \left[1 - \left(\frac{65 - t}{15}\right) \cdot \mathbf{1}\{t < 65\}\right] \cdot \left[1 + 0.06 \cdot \min(4, t - 65) \cdot \mathbf{1}\{t > 65\}\right] \tag{C9}$$

where PIA_{it} comes from Equation (C4).

When $ss_{it} + ssa_{it} > 0$, the Social Security benefit is calculated as

$$ssb^{pre}_{it} = ssb_{it} - ET_{it} \cdot \mathbf{1}\{t < 70\} \tag{C10}$$

The ssb_{it} is then updated as

$$ssb_{it+1} = ssb_{it} \cdot \left(1 + \frac{ET_{it}}{ssb_{it}} \cdot \{0.067 \cdot \mathbf{1}\{t < 65\} + 0.06 \cdot \mathbf{1}\{65 \leq t < 70\}\}\right) \tag{C11}$$

The earnings test is calculated from Equation (C6).

Appendix D 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), with an age-dependent human capital production function. 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

$$H_{it+1} = (1-\delta)\, H_{it} + \xi_{it}\pi \left(1 + \alpha_1 t + \alpha_2 t^2\right) \tag{D1}$$

where t is potential experience. Notice that this is very close to our standard model from Equation (6). 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. For this reason, we remove the moment condition of wage change rate after one nonemployment spell when estimating this exogenous model.

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

$$H_{it+1} = (1-\delta)\, H_{it} + (1-\ell_{it})\, \xi_{it}\pi \left(1 + \alpha_1 t + \alpha_2 t^2\right). \tag{D2}$$

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 supply decision. When individuals do not work, their human capital depreciates at rate δ.[41]

In Section 5 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

[41] We estimated another two different versions of production functions in (D1) and (D2): one replacing $\left(1 + \alpha_1 t + \alpha_2 t^2\right)$ with $H_{it}^{\alpha_H}$ and the other with $\left(1 + \alpha_1 H_{it} + \alpha_2 H_{it}^2\right)$ in the second term at the right hand side. The results are very similar.

supply but small decrease in wages at the end. The first is human capital depreciation: when workers stop working their human capital falls. The second is the distinction between measured 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 D1 and the fit of the two models is presented in Figure D1. We first discuss the completely exogenous model. As expected, it is difficult for this model to fit both the labor supply and the two wage profiles at the same time. Precautionary savings and Social Security lead to income effects where labor supply can fall late in life in the exogenous model. 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 large 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 analogue to the empirical elasticity as above and present it in Figure D2. The exogenous model requires a substantially larger elasticity at most ages.

By contrast the learning-by-doing model fits the data well. The elasticity of labor supply is much closer to the baseline model than it is to the exogenous model—as one can see from Figure D2, or from the fact that σ_ε takes on a value 0.027 as opposed to 0.002 in the exogenous model, still lower than the value of 0.197 in the baseline model. This results in a higher elasticity of labor supply at early ages than the baseline model. The key to understanding this difference is human capital. When the human capital rental rate increases at age t, in the learning-by-doing model workers are able to adjust their labor supply decision throughout the whole life cycle, which is more efficient than the exogenous model and consequently induces a smaller elasticity of labor supply. On the other hand, the baseline model gives a worker 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 even more efficiently than the learning-by-doing model, and therefore a smallest elasticity especially at early ages. It is important to note here that we did not try a wide range of exogenous or learning-by-doing models; we just did a comparison between our baseline model and an exogenous or 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 difference between human capital and measured wages.

Table D3 presents the budget calculation from the seven policy experiments for these two alternative human capital models in Panel C and D.

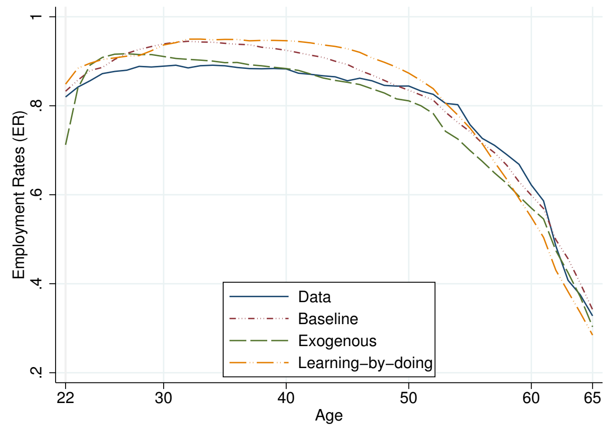
Table D1: Estimates of alternative human capital models[a]

Parameters		Exogenous[b] Estimates	Exogenous[b] S.E.	Learning-by-Doing[c] Estimates	Learning-by-Doing[c] S.E.
HC depreciation[d]	δ	0.116	(0.007)	0.086	(0.002)
HC production function: on t	α_1	0.015	(0.001)	-0.002	(0.001)
HC production function: on t^2 $(\times 10^4)$	α_2	-2.985	(0.213)	0.046	(0.007)
Standard deviation of HC innovation	σ_ξ	0.921	(0.053)	0.012	(0.004)
Consumption: CRRA	η_c	4.100	(0.025)	3.745	(0.028)
Consumption shifter: on t $(\times 10)$	φ_1	0.455	(0.016)	0.462	(0.056)
Consumption shifter: on t^2 $(\times 10^2)$	φ_2	0.067	(0.003)	0.076	(0.013)
Consumption shifter: on t^3 $(\times 10^3)$	φ_3	-0.035	(0.001)	-0.030	(0.002)
Consumption shifter: coef on married	φ_4	1.484	(0.072)	0.034	(0.005)
Leisure: standard deviation of shock	σ_ε	0.002	(0.001)	0.027	(0.008)
Leisure: spouse not working	a_1	1.234	(0.062)	0.057	(0.014)
Leisure: spouse working	a_2	-0.257	(0.010)	-0.873	(0.051)
Bequest weight	b_1	67,773,144	(6,610,869)	7,802,931	(933,463)
Parameter heterogeneity					
Leisure: mean of intercept	μ_{a_0}	-6.169	(0.064)	-5.275	(0.078)
Leisure: standard deviation of intercept	σ_{a_0}	0.108	(0.012)	0.368	(0.030)
HC productivity, mean	μ_π	1.896	(0.097)	1.703	(0.010)
HC productivity, standard deviation	σ_π	0.650	(0.038)	0.688	(0.009)
Correlation between a_0 & π	ρ	0.030	(0.006)	-0.480	(0.053)
Initial human capital level at age 18					
Intercept	γ_0	1.234	(0.302)	1.195	(0.130)
Coefficient on a_0	γ_{a_0}	0.055	(0.011)	-0.066	(0.019)
Coefficient on π	γ_π	0.668	(0.140)	0.550	(0.079)
Standard deviation of error term	σ_{H_0}	0.034	(0.050)	0.389	(0.043)
Additional Social Security Application effects					
Effect of resource constraint $(\times 10^3)$	b_{62}	0.218	(0.037)	0.293	(0.052)
Effect of health insurance: constant $(\times 10^3)$	b_{65}	0.062	(0.023)	0.057	(0.053)
Effect of health insurance: coef on t $(\times 10^3)$	b_{65t}	0.205	(0.044)	0.272	(0.043)
χ^2 Statistic[e]		1820		871	
Degrees of freedom		206		207	

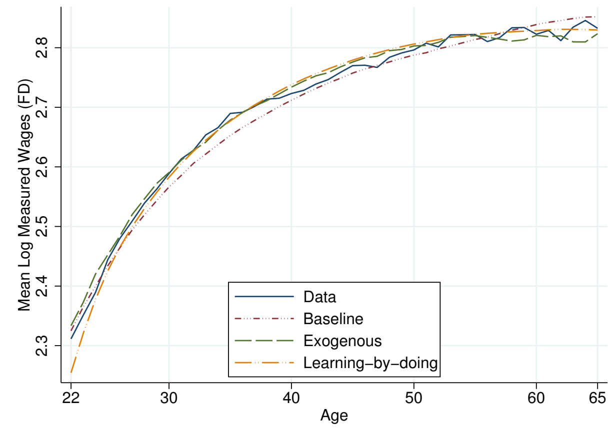
[a]Indirect inference estimates. Estimates use a diagonal weighting matrix. Standard errors are given in parentheses.
[b]In the exogenous model, the human capital production function is $H_{t+1} = (1-\delta) H_t + \xi_t \pi \left(1 + \alpha_1 t + \alpha_2 t^2\right)$.
[c]In the learning-by-doing model, 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: Human Capital.
[e]This is the J-statistic. The critical values of the χ^2 distribution are $\chi^2_{(206,0.01)} = 256$, $\chi^2_{(207,0.01)} = 257$.

Figure D1: Exogenous and learning-by-doing models moments

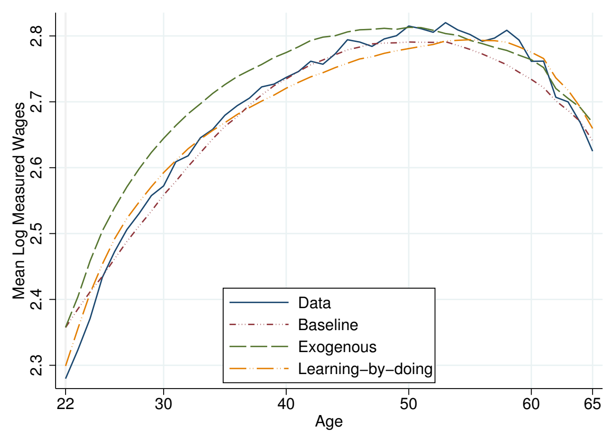
(a) Employment rates



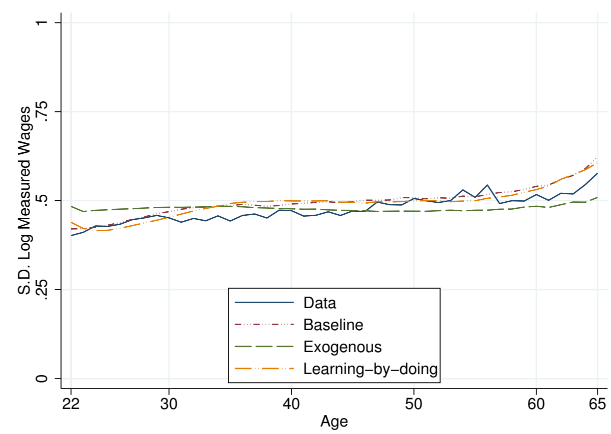
(b) Mean log measured wages (FD)



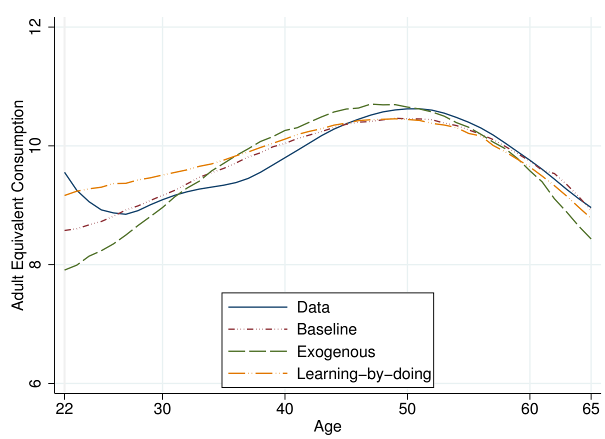
(c) Mean log measured wages



(d) S.D. log measured wages



(e) Adult equivalent consumption



(f) Social Security application

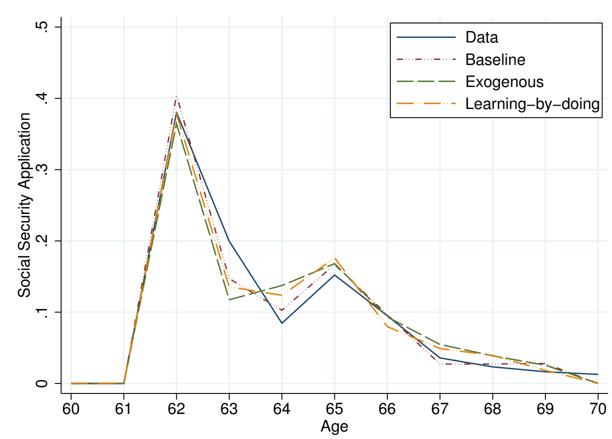


Figure D2: Analogue to the Empirical Elasticity of Labor Supply: Exogenous model and Learning-by-doing model

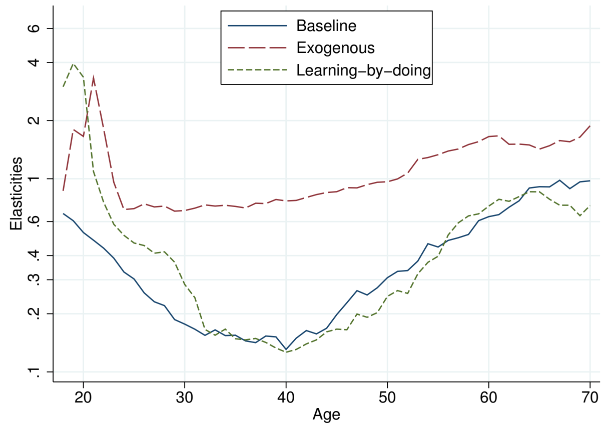


Table D2: Transitions from alternative human capital models[a]

	Models	Transition Probabilities: Working to Not Working	Transition Probabilities: Not Working to Working	Wage Change Rate After One Nonemployment Spell[b]
1	Data	0.034	0.200	-0.071
2	Exogenous model	0.047	0.248	0.041[c]
3	Learning-by-doing Model	0.018	0.173	-0.081

[a]The transition rate is the average transition probability between age 35 and 50.

[b]The average wage change rate after one nonemployment spell is the average change rate between age 41 and 65.

[c]The moment is not targeted.

Table D3: Budget calculation for experiments for the four human capital models

	0	1		2		3		4		5		6		7	
	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: Model with Health and Part Time															
Total Output	1305.185	23.085	1.769	1.688	0.129	12.098	0.927	15.161	1.162	-46.962	-3.598	84.628	6.484	18.614	1.426
SS Tax	80.921	1.431	1.769	0.105	0.129	0.750	0.927	0.940	1.162	-80.921	-100	5.247	6.484	-80.921	-100
Medicare Tax	18.925	0.335	1.769	0.024	0.129	0.175	0.927	0.220	1.162	-18.925	-100	1.227	6.484	-18.925	-100
SS Benefit	191.031	0.013	0.007	-0.628	-0.329	-23.220	-12.155	-38.258	-20.027	-2.306	-1.207	-191.031	-100	-191.031	-100
SS Deficit[d]	91.184	-1.753	-1.922	-0.757	-0.830	-24.146	-26.480	-39.417	-43.228	97.540	106.971	-197.505	-216.600	-91.184	-100
Income Tax	225.831	84.317	37.336	1.367	0.605	1.080	0.478	3.555	1.574	-39.043	-17.289	23.169	10.260	-22.674	-10.040
SSDI Benefit	16.374	0.201	1.226	-0.046	-0.281	4.324	26.409	0.025	0.154	-0.045	-0.278	-0.165	-1.006	-0.070	-0.426
Panel B: Baseline Model															
Total Output	1288.339	51.007	3.959	7.101	0.551	15.761	1.223	28.359	2.201	-134.023	-10.403	177.958	13.813	8.233	0.639
SS Tax	79.877	3.162	3.959	0.440	0.551	0.977	1.223	1.758	2.201	-79.877	-100	11.033	13.813	-79.877	-100
Medicare Tax	18.681	0.740	3.959	0.103	0.551	0.229	1.223	0.411	2.201	-18.681	-100	2.580	13.813	-18.681	-100
SS Benefit	199.828	1.303	0.652	-1.291	-0.646	-24.891	-12.456	-39.955	-19.995	-4.179	-2.091	-199.828	-100	-199.828	-100
SS Deficit[d]	101.270	-2.599	-2.566	-1.834	-1.811	-26.097	-25.770	-42.124	-41.596	94.379	93.195	-213.442	-210.765	-101.270	-100
Income Tax	239.235	93.076	38.906	1.846	0.772	3.585	1.499	6.672	2.789	-61.028	-25.510	38.856	16.242	-34.533	-14.435
Panel C: Exogenous Model															
Total Output	1249.544	33.515	2.682	0.882	0.071	10.028	0.803	17.067	1.366	-74.308	-5.947	93.381	7.473	-3.844	-0.308
SS Tax	77.350	2.073	2.680	0.054	0.070	0.622	0.805	1.057	1.367	-77.350	-100	5.783	7.477	-77.350	-100
Medicare Tax	18.118	0.486	2.682	0.013	0.071	0.145	0.803	0.247	1.366	-18.118	-100	1.354	7.473	-18.118	-100
SS Benefit	207.971	2.408	1.158	-0.083	-0.040	-25.880	-12.444	-40.888	-19.660	-4.140	-1.991	-207.971	-100	-207.971	-100
SS Deficit[d]	112.503	-0.151	-0.134	-0.150	-0.133	-26.648	-23.686	-42.193	-37.504	91.328	81.178	-215.108	-191.202	-112.503	-100
Income Tax	224.351	84.393	37.617	0.531	0.237	2.899	1.292	4.880	2.175	-49.263	-21.958	28.464	12.687	-29.895	-13.325
Panel D: Learning-by-doing Model															
Total Output	1253.798	36.121	2.881	1.351	0.108	10.183	0.812	20.000	1.595	-92.479	-7.376	122.451	9.766	-0.543	-0.043
SS Tax	77.735	2.240	2.881	0.084	0.108	0.631	0.812	1.240	1.595	-77.735	-100	7.592	9.766	-77.735	-100
Medicare Tax	18.180	0.524	2.881	0.020	0.108	0.148	0.812	0.290	1.595	-18.180	-100	1.776	9.766	-18.180	-100
SS Benefit	207.410	2.643	1.274	-0.076	-0.036	-25.768	-12.424	-40.986	-19.761	-4.884	-2.355	-207.410	-100	-207.410	-100
SS Deficit[d]	111.494	-0.120	-0.108	-0.179	-0.160	-26.547	-23.810	-42.516	-38.133	91.031	81.647	-216.777	-194.429	-111.494	-100
Income Tax	214.309	80.054	37.354	0.358	0.167	2.879	1.343	5.185	2.419	-50.174	-23.412	31.769	14.824	-27.688	-12.920

[a]The "Level" column refers to the aggregate value over the whole life cycle. For example, in the baseline model, the total output that an average individual produce between 18 and 80 is $1,288 thousand. All numbers in this column are in the unit of $1000.

[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 total output is $7.101 thousand higher than that in the baseline model across the whole life cycle from 18 to 80. All numbers in this column are in the unit of $1000.

[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 total output increases by $7.101 thousand which is equivalent to 0.551% of the total output in the baseline model. All numbers in this column are in the unit of %.

[d]"SS Deficit" = SS Benefit - SS Tax - Medicare Tax. It is the deficit for the government budget.

Appendix E Robustness Check

In Subsection 4.1 we set some of the parameters 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 β from 0.97 to 0.96 but increase the interest r from 0.03 to 0.04; (4) increase the initial asset A_0 from 0.0 to $50,000$. In each case, all other pre-set parameters are kept the same as the baseline model, and then we re-estimate all of the parameters of the model. The estimation results are listed in Table E2; the moments are plotted in Figure E1 and listed in Table E1 as well.

In all cases the simulated moments fit the data moments quite well. Varying pre-set parameters does change the estimated values of some parameters, but in all variants our model generates simulated auxiliary model which matches data auxiliary model quite well.

We perform another exercise. In general we try to fit the wage and employment profiles without directly relying on age dummies. As part of our empirical exercise we also fit the consumption profile and the Social Security application age effects. The question is whether we need these age effects to fit wages and employment. To address that we re-estimate the model without the consumption specific or Social Security application specific time effects.[42] We also no longer target consumption or the application age. These results are also shown in Figure E1. It should be clear that we do an excellent job fitting wages and labor supply without these. The fit on consumption is not very good. The fit of the application age is substantially worse than the baseline—but is actually not bad given that it is not targeted at all.

We also investigate the effects of borrowing constraint. Gradually tightening the borrowing constraint initially has minimal effects on the fitting of the estimation or the counterfactual requests. For instance, setting the borrowing limit to 25% of that in the baseline model does not change the results materially. In this case the estimated CRRA coefficient is smaller, $\eta_c = 3.450$, which is consistent with the statement made in Keane and Wolpin (2001) (Page 1078). Further tightening the borrowing constraint deteriorates the fitting of the model, especially on the consumption profile. This exercise illustrates that we need either a borrowing option (as in our model) or a higher initial asset (as in Keane and Wolpin, 2001).

[42] We set $\varphi_1 = \varphi_2 = \varphi_3 = 0$ in Equation (2) and $b_{62} = b_{65} = b_{65t} = 0$ in Equation (4).

Figure E1: Fit of alternative models

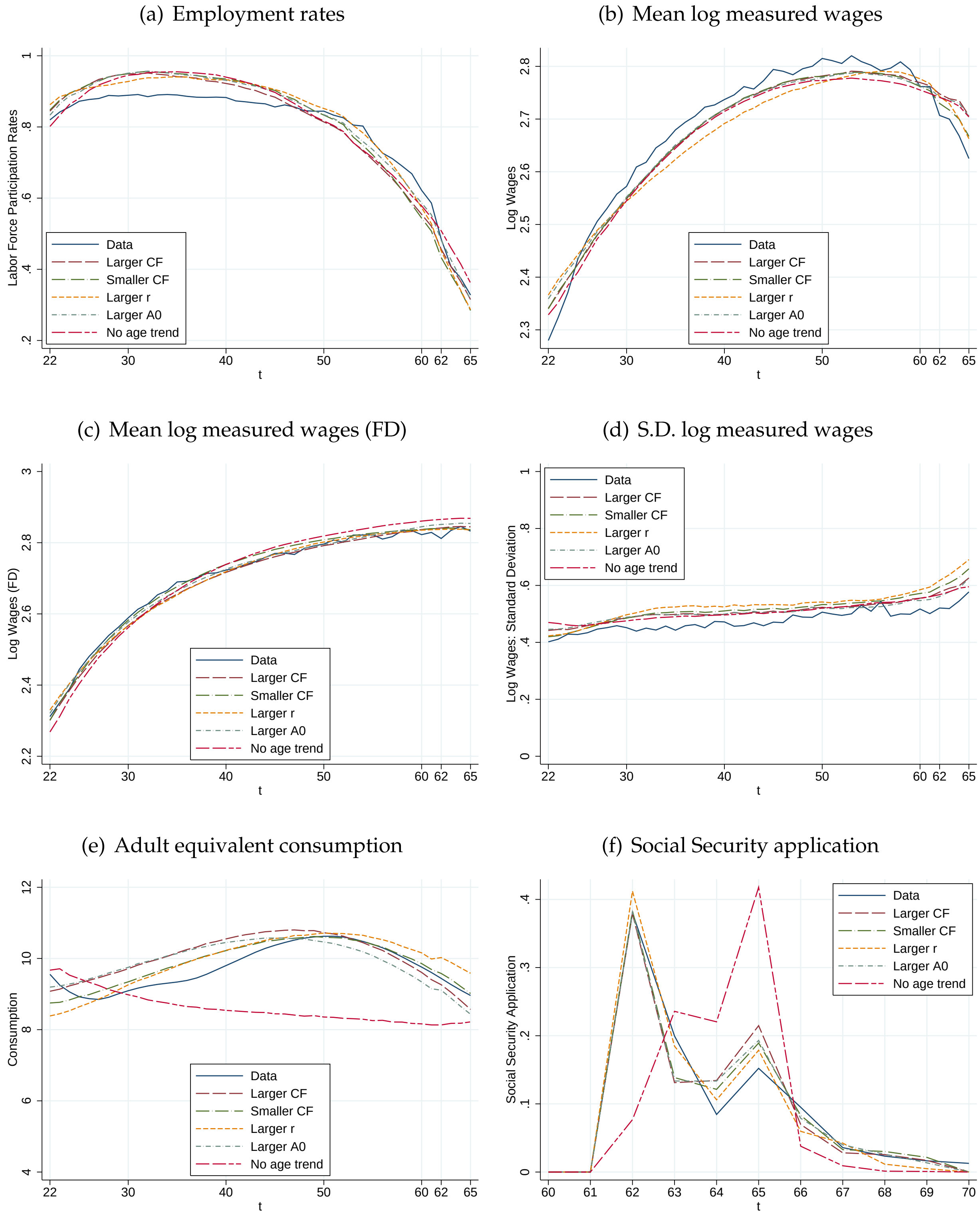


Table E1: Transitions from alternative models[a]

	Models	Transition Probabilities: Working to Not Working	Transition Probabilities: Not Working to Working	Wage Change Rate After One Nonemployment Spell[b]
	Data	0.034	0.200	-0.071
1	Larger $\underline{c}$	0.036	0.217	-0.081
2	Smaller $\underline{c}$	0.029	0.212	-0.081
3	Larger r smaller δ	0.025	0.218	-0.081
4	Larger A_0	0.035	0.220	-0.080
5	No age trend	0.036	0.240	-0.070

[a]The transition rate is the average transition probability between age 35 and 50.

[b]The average wage change rate after one nonemployment spell is the average change rate between age 41 and 65.

Table E2: Estimates in the model variants[a]

		1	2	3	4	5
MODEL SPECIFICATIONS		Larger $\underline{c}$	Lower $\underline{c}$	Change δ, r	Larger A_0	No age trend
Interest rate	r			0.04		
Discount	β			0.96		
Initial wealth	A_0				$50,000$	
Consumption floor	$\underline{c}$	2.5	1.8			
HC depreciation[b]	δ	0.087	0.087	0.085	0.091	0.077
HC production function: I factor	α_I	0.088	0.104	0.090	0.088	0.071
HC production function: H factor	α_H	0.058	0.075	0.064	0.055	0.044
Standard deviation of HC innovation	σ_ξ	0.015	0.006	0.004	0.175	0.012
Consumption: CRRA	η_c	3.957	4.015	4.020	3.994	4.031
Consumption shifter: coef on t $\left(\times 10^{-1}\right)$	φ_1	0.436	0.141	0.120	0.154	0
Consumption shifter: coef on t^2 $\left(\times 10^{-2}\right)$	φ_2	0.090	0.174	0.177	0.130	0
Consumption shifter: coef on t^3 $\left(\times 10^{-3}\right)$	φ_3	-0.035	-0.040	-0.038	-0.036	0
Consumption shifter: coef on married	φ_4	0.763	1.168	1.456	1.298	0.188
Leisure: standard deviation of shock	σ_ε	0.178	0.162	0.106	0.181	0.170
Leisure: spouse not working	a_1	-0.136	0.052	0.272	-0.253	-0.364
Leisure: spouse working	a_2	-0.212	-0.231	-0.387	-0.145	-0.371
Bequest weight	b_1	23,918,044	37,269,064	103,870,248	60,683,056	41,321,532
Parameter heterogeneity[c]						
Leisure: mean of intercept	μ_{a_0}	-5.689	-5.770	-5.632	-5.870	-6.434
Leisure: standard deviation of intercept	σ_{a_0}	0.664	0.680	0.518	0.580	0.595
HC productivity, mean	μ_π	1.778	1.785	1.756	1.864	1.685
HC productivity, standard deviation	σ_π	0.692	0.701	0.706	0.727	0.670
Correlation between a_0 & π	ρ	-0.759	-0.587	-0.613	-0.764	-0.354
Initial human capital level at age 18						
Intercept	γ_0	1.519	1.556	1.500	1.522	1.539
Coefficient on a_0	γ_{a_0}	0.050	0.028	0.022	0.040	0.053
Coefficient on π	γ_π	0.665	0.577	0.616	0.630	0.731
Standard deviation of error term	σ_{H_0}	0.054	0.044	0.003	0.005	0.007
Additional Social Security Application effects						
Effect of resource constraint $\left(\times 10^{-3}\right)$	b_{62}	0.205	0.270	0.138	0.179	0
Effect of health insurance: constant $\left(\times 10^{-3}\right)$	b_{65}	0.015	0.015	0.031	0.013	0
Effect of health insurance: coef on t $\left(\times 10^{-3}\right)$	b_{65t}	0.109	0.133	0.146	0.107	0

[a] Indirect inference estimates. Estimates use a diagonal weighting matrix.

[b] HC: Human Capital.

[c] 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.

Appendix F College Graduates

We estimate the model for college graduates separately as well. The results are shown in Table F2. One can see that the fits are quite well as shown in Table F1 and Figure F1.

Table F1: Transitions for college graduates[a]

		Transition Probabilities		Wage Change Rate
	Models	Working to Not Working	Not Working to Working	After One Nonemployment Spell[b]
12	Data	0.018	0.296	-0.095
13	Model	0.018	0.304	-0.112

[a]The transition rate is the average transition probability between age 35 and 50.

[b]The average wage change rate after one nonemployment spell is the average change rate between age 41 and 65.

Table F2: Estimates in the baseline model for college graduates[a]

Parameters		Estimates	Standard Errors
HC depreciation[b]	δ	0.118	(0.004)
HC production function: I factor	α_I	0.116	(0.012)
HC production function: H factor	α_H	0.236	(0.011)
Standard deviation of HC innovation	σ_ξ	0.00001	(0.00001)
Consumption: CRRA	η_c	3.811	(0.036)
Consumption shifter: coef on t $\left(\times 10^{-1}\right)$	φ_1	-0.047	(0.007)
Consumption shifter: coef on t^2 $\left(\times 10^{-2}\right)$	φ_2	0.273	(0.024)
Consumption shifter: coef on t^3 $\left(\times 10^{-3}\right)$	φ_3	-0.062	(0.003)
Consumption shifter: coef on married	φ_4	2.480	(0.252)
Leisure: standard deviation of shock	σ_ε	0.230	(0.021)
Leisure: spouse not working	a_1	0.128	(0.025)
Leisure: spouse working	a_2	-0.094	(0.017)
Bequest weight	b_1	79,749,696	(9,824,487)
Parameter heterogeneity[c]			
Leisure: mean of intercept	μ_{a_0}	-5.736	(0.101)
Leisure: standard deviation of intercept	σ_{a_0}	0.845	(0.087)
HC productivity, mean	μ_π	2.192	(0.038)
HC productivity, standard deviation	σ_π	0.861	(0.026)
Correlation between a_0 & π	ρ	0.030	(0.007)
Initial human capital level at age 18			
Intercept	γ_0	1.536	(0.112)
Coefficient on a_0	γ_{a_0}	-0.068	(0.020)
Coefficient on π	γ_π	0.386	(0.034)
Standard deviation of error term	σ_{H_0}	0.006	(0.005)
Additional Social Security Application effects			
Effect of resource constraint $\left(\times 10^3\right)$	b_{62}	0.361	(0.068)
Effect of health insurance: constant $\left(\times 10^3\right)$	b_{65}	0.039	(0.027)
Effect of health insurance: coef on t $\left(\times 10^3\right)$	b_{65t}	0.109	(0.016)
χ^2 Statistic = 2141[d]		Degrees of freedom = 187	

[a]Indirect inference estimates. Estimates use a diagonal weighting matrix. Standard errors are given in parentheses.

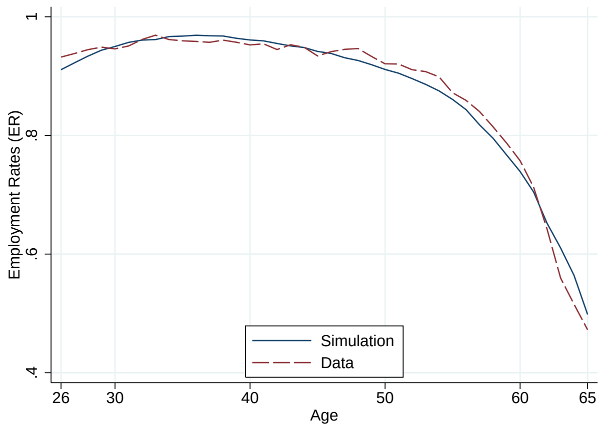
[b]HC: Human Capital.

[c]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.

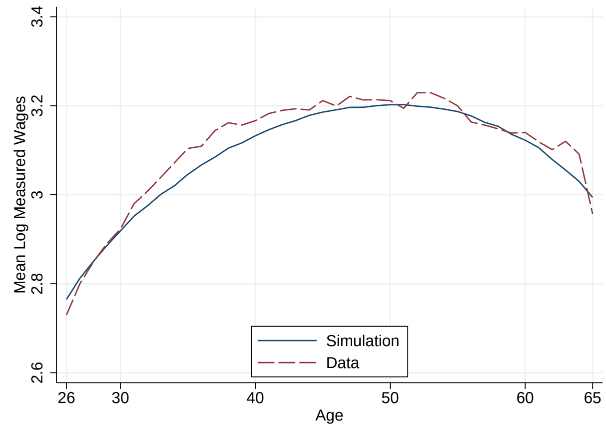
[d]This is the J-statistic. The critical value of the χ^2 distribution is $\chi^2_{(207,0.01)} = 257$.

Figure F1: Fit of model with college graduates

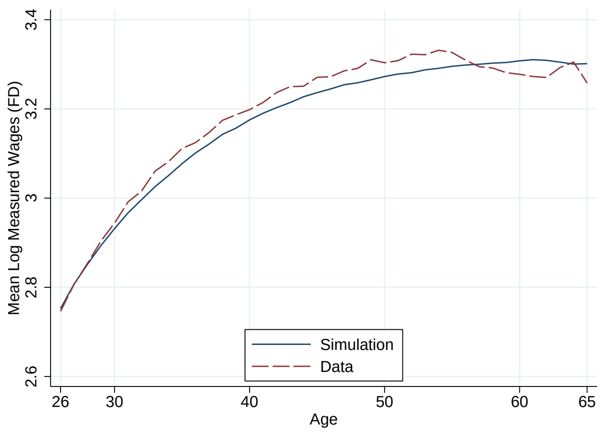
(a) Employment rates



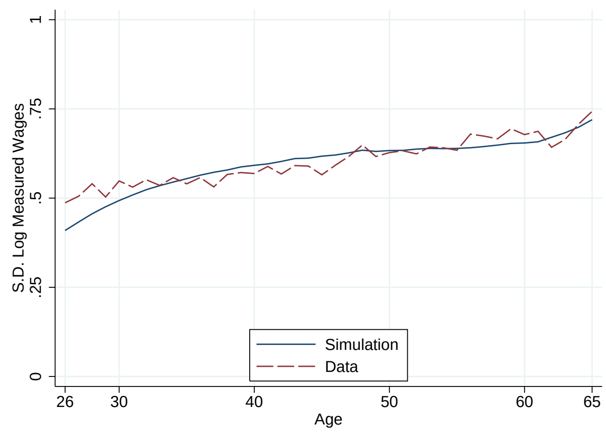
(b) Mean log measured wages



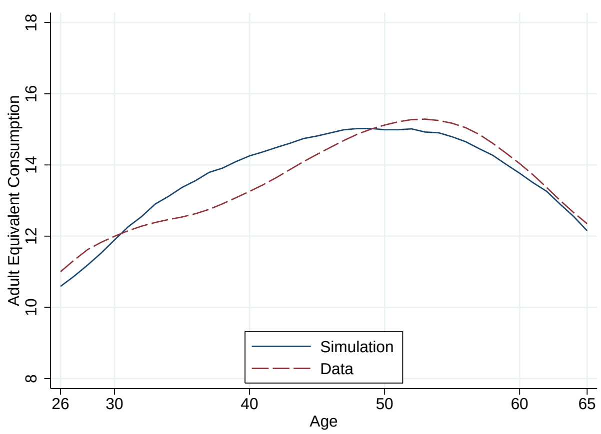
(c) Mean log measured wages (FD)



(d) S.D. log measured wages



(e) Adult equivalent consumption



(f) Social Security application