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

Xiaodong Fan

Monash University

Ananth Seshadri

University of Wisconsin–Madison

Christopher Taber

University of Wisconsin–Madison

We estimate a life-cycle model of consumption, human capital investment, and labor supply. The interaction between human capital and labor supply toward the end of the life cycle is most novel. The estimates replicate the main features of the data, in particular the large increase in wages and small increase in labor supply at the beginning of the life cycle and the small decrease in wages but large decrease in labor supply toward the end. We show that incorporating human capital is critical when analyzing changes to Social Security.

I. 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

We thank James Heckman, four anonymous referees, seminar participants at the Society for Economic Dynamics, the University of Tokyo, Yale University, the Society of Labor Economists, the Centre for Economic Policy Analysis and Research, the University of Reading,

Electronically published December 14, 2023

Journal of Political Economy, volume 132, number 1, January 2024.

https://doi.org/10.1086/726232

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 of Heckman (1975, 1976), there has been little effort integrating these two important paradigms. A goal of this paper is 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 decision separate from labor supply in the model or 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.

A novel aspect of our paper is examining policy effects on human capital toward 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, Lochner, and Taber 1998a; Manuelli, Seshadri, and Shin 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 supply profiles of male high school graduates from the Survey of Income and Program Participation

Royal Holloway, the Catholic University of Milan, the University of Minnesota, the University of Toulouse, Cornell University, and Johns Hopkins University for helpful comments and suggestions. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure, which is supported by the Australian government. All remaining errors are our own. This paper was edited by James Heckman.

[1] Iskhakov and Keane (2021) do look at retirement rules—but in the Australian context, which is a very different system.

(SIPP). We are able to replicate the main features of the data with a parsimonious life-cycle model in which the taste for leisure does not depend on age or experience. 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. Depreciation plays a crucial role for two reasons.

The first and quantitatively more important role of depreciation arises in its interaction with the labor supply decision. In a simple model without human capital depreciation, there is no a priori reason for workers to concentrate their leisure toward 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. Alternatively, 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 reenter at a lower wage, so they continue to stay out of the labor market.

The second role of depreciation is more subtle but perhaps more interesting. It comes from a point emphasized by Heckman, Lochner, and Taber (1998a) that in a Ben-Porath-style model wages are different from human capital, as workers are paid only for their productivity net of investment time. As a result, the model can yield wages that do not change with age at older ages while human capital declines because investment is also falling. This declining human capital also helps reconcile the falling labor supply late in life despite relatively flat wages. This arises from an interaction between the depreciation and the elasticity of human capital investment with respect to time. When the elasticity is lower, human capital investment spreads across the life cycle rather than being concentrated at the beginning. The combination of low elasticity and relatively large depreciation leads the model to produce relatively flat wages yet declining human capital.

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 determining the labor supply of older workers, 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. Previous 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 reestimates 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 to be 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 ages 60 and 70. Over the whole life cycle, human capital investment, human capital level, total employment rates, measured average log wages, and total pretax labor income increase by 0.12%, 0.23%, 0.53%, 0.13%, and 0.71%, respectively, in the general model with health and the part-time option.

Section II of this paper briefly reviews the most relevant literature. Section III introduces the model, and section IV explains how it is estimated. Section V presents the estimates from the baseline model, and section VI explains the extension to the more general model. Section VII simulates the policy counterfactuals, and section VIII concludes.

II. 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, Seshadri, and Shin (2012) calibrate a Ben-Porath model to include the endogenous retirement decision. All three models are deterministic.

Relative to the success in theory, there has not 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 (see a survey of the literature in Heckman, Lochner, and Todd 2006). Early work on explicit estimation of the Ben-Porath model was done by Heckman (1975, 1976), Haley (1976), and Rosen (1976). Heckman, Lochner, and Taber (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 anymore when the retirement is uncertain, where each individual always has an incentive to invest a positive amount in human capital. Browning, Hansen, and Heckman (1999) survey much of this literature.[2]

Another type of human capital model, the learning-by-doing model, has been utilized relatively more often in empirical work. In the standard learning-by-doing model, human capital accumulates exogenously but only when an individual works. Thus, workers impact their human capital accumulation only through their labor supply decision. In these models, the total cost of leisure is not only the direct lost earnings at the current time but also 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 Panel Study of Income Dynamics (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) are a more recent example. These papers assume an exogenously fixed retirement age. Keane and Wasi (2016) and Iskhakov and Keane (2021) extend these models to consider workers at older ages and show that the models fit well.

Heckman, Lochner, and Cossa (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 earned

[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.

income tax credit (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 by either 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 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, Rogerson, and Wallenius (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 Germany, 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 presumably change the wage profile.

III. Model

A. Overview

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

- human capital investment,
- labor supply (extensive margin), and
- 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, and
- consumption floor.

Our extended model adds the following features:

- health status—including disability—and
- part-time work.

B. *Environment and Econometric Specification*

1. 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 age/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, Dustmann, and Stevens (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.

2. Preferences

In the baseline model, we focus on the extensive margin of labor supply only, so at each period the individual decides whether to work. The flow utility at period t is

$$u_t(c_{it}, \ell_{it}, f_{it}; M_{it}, \varepsilon_{it}, a_{i0}) = \psi(t, M_{it})\frac{c_{it}^{1-\eta_c}}{1-\eta_c} + \gamma(a_{i0}, M_{it}, \varepsilon_{it})\ell_{it} + \varpi(t)\mathrm{ssa}_{it}, \quad (1)$$

where c_{it} represents total family consumption, $\ell_{it} \in \{0, 1\}$ represents 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 do not model this any different than any other period of nonemployment.

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

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

where $\mathbf{1}\{\cdot\}$ is the indicator function. Note that the shifter depends on marital status.

The coefficient $\gamma(a_{i0}, M_{it}, \varepsilon_{it})$ represents taste for leisure and also depends on the family status. We use the parametric form

$$\gamma(a_{i0}, M_{it}, \varepsilon_{it}) = \exp\left(a_{i0} + \sum_{j=1}^{2} a_j \mathbf{1}\{M_{it} = j\} + \varepsilon_{it}\right), \tag{3}$$

where ε_{it} follows an independent and identical normal distribution with mean zero and variance σ_ε^2. The a_{i0} is a component of unobserved heterogeneity that we discuss below. A key part of our exercise is that we do not explicitly allow $\gamma(a_{i0}, M_{it}, \varepsilon_{it})$ 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 ages 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 ages 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.[3] With this goal in mind, we let the model fit these patterns by assuming that an individual obtains additional utility from receiving the Social Security benefit, and thus this total additional flow utility at period t becomes

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

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

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

[3] However, it is not clear whether to take this as fixed with policy changes. We explore this when we perform counterfactuals.

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

3. 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} represents the human capital level at period t, ξ_{it} represents an idiosyncratic shock to the human capital innovation, π_i represents a form of unobserved heterogeneity, and δ, α_I, and α_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 that ξ_{it} is independent and identically distributed (i.i.d.) and follows a lognormal distribution,

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

This specification yields an expected value of one for the level of ξ_{it} 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, pretax labor income at any point in time is

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

4. Social Security and Budget Constraint

While we have tried to keep the baseline model as simple as possible, the Social Security system in the United States 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 before period t, and as mentioned above, ssa_{it} indicates the decision to start claiming benefits. Since claiming is irreversible, once $\text{ss}_{it} = 1$ then ssa_{it} is no longer a relevant choice variable. Thus, the law of motion can be written as

$$\mathrm{ss}_{i1} = 0, \mathrm{ss}_{it+1} = \max\{\mathrm{ss}_{it}, \mathrm{ssa}_{it}\}. \tag{9}$$

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 of 2004. The benefit ssb_{it} is updated each year if an individual worked to account for the earnings test. This is incorporated into the budget constraint

$$A_{it+1} = A_{it} + \Upsilon_t(rA_{it}, w_{it}, y_{it}, \mathrm{ssb}_{it}) - c_{it} + \tau_{it}, \tag{10}$$

where A_{it} stands for asset, r represents the risk-free interest rate, and y_{it} represents spousal income. The $\Upsilon_t(\cdot)$ represents 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 (apps. A–F are available online).

Spousal income takes the form

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

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

Government transfers, τ_{it}, provide a consumption floor $\underline{c}$ as in Hubbard, Skinner, and Zeldes (1995) so that

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

where $\underline{A}_{it+1}$ represents the asset lower bound at period $t + 1$.[4]

We note that some of our model assumptions are strong. One example is that we assume that human capital is financed by the worker completely 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 whether 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

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

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 to increase future earnings. In addition to workers' contribution to training, our wage specification also captures career choices with a low starting wage but a steeper age-earnings profile. For 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.

C. Solving the Model

Four random variables are realized each period: 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 the value they had before their realizations.

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

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

subject to (1)–(12), where

$$X_{it} = \{M_{it}, A_{it}, H_{it}, \mathrm{ss}_{it}, \mathrm{AIME}_{it}, \mathrm{ssb}_{it}, \varepsilon_{it}; a_{i0}, \pi_i\} \tag{14}$$

is the vector of state variables. Note that AIME_{it} is relevant only before claiming ($\mathrm{ss}_{it} = 0$), while ssb_{it} is determined only after claiming ($\mathrm{ss}_{it} = 1$). That is, before 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} enters the model only 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

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

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

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

and the conditional value function for working is defined as

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

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.

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

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

and define the conditional value function for not working to be

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

The optimal labor supply solution is

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

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

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

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

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

Using properties of lognormal random variables, we show in appendix A that the expected value function is

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

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

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

$$E[V_{t+1}(X_{it+1}) \mid X_{it}, c_{it}, \ell_{it}, I_{it}, \text{ssa}_{it}] = E[E[V_{t+1}(X_{it+1}) \mid \tilde{X}_{it+1}] \mid \tilde{X}_{it}, c_{it}, \ell_{it}, I_{it}, \text{ssa}_{it}]$$

we integrate over the distributions of ξ_{it}, ζ_{it}, and M_{it+1}.[5]

D. *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 have only 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 (μ_{a_0}, σ_{a_0}, μ_π, σ_π, ρ). We emphasize that since we are using only nine points, we are not assuming that the Gauss-Hermite is a good approximation of a normal but rather taking this as the parameterization 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

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

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

[5] $\tilde{X}_{it+1}$ 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 eq. (10), H_{it+1} is determined in eq. (6), ss_{it} is determined by eq. (9), AIME_{it} and ssb_{it} are described in app. C, and a_{i0} and π_i do not change.

IV. Estimation

The estimation of the model is carried out using a three-step strategy. First, we preset 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 most important 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 μ_{a_0}, σ_{a_0}, μ_{π}, σ_{π}, ρ, γ_0, γ_{a_0}, γ_{π}, 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, σ_{ε}; 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 AIME_{i0}.

A. *Preset Parameters*

The set of parameters preset 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 1 year.[6] The initial period in our model corresponds to age 18 and ends at age 80.[7] The early-retirement age is 62, and the normal retirement age (NRA) is 65. The risk-free real interest rate is set as $r = 0.03$, and the time discount rate is set as $\beta = 0.97$. The consumption floor is set as $\underline{c} = 2.19$, as estimated in French and Jones (2011).[8]

The parameter that determines the curvature of the bequest function is set as $b_2 = 222$, as in French and Jones (2011).[9] We assume that all individuals start their adult life with no wealth and zero level of AIME at age 18. These normalized or preset parameters are summarized in table 1. In appendix E, we show that the results are robust to other alternatives.

[6] Midyear 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 major issue.

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

[8] $\underline{c} = 4{,}380/2{,}000 = 2.19$ since we normalize the total time endowment for labor supply at 1 period as one.

[9] It is equivalent to $444,000 in 2004 US dollars. 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.

TABLE 1
Normalized or Preset Parameters

Parameter	Definition	Normalized/Preset Value
r	Interest rate	.03
β	Discount	.97
A_0	Initial wealth[a]	.0
AIME_0	Initial AIME[a]	.0
$\underline{c}$	Consumption floor	2.19
b_2	Bequest shifter	222.0

Note.—The consumption floor is equivalent to $4,380 in 2004 US dollars, since we normalize the total time endowment for labor supply at 1 period—which is 2,000 hours—as one. The bequest shifter is equivalent to $444,000 in 2004 US dollars.

[a] The initial age is 18.

B. Demographics

Given unobserved heterogeneity in our model, we cannot 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 a quartic function of ages. The results are presented in appendix B. Figures B1*a*–B1*c* (figs. B1, B2, C1, D1, D2, E1–E3, and F1 are available online) plot the transition probabilities at each age, and figure B1*d* displays the resulting distributions, which are similar to patterns in the SIPP data and the CPS data. For each age, we estimate the mean $\mu_{\xi t}$ and standard deviation $\sigma_{\xi t}$ of the logarithm of the positive spousal income in the SIPP data and then smooth them by an age quadruple function. Figure B1*e* plots the estimated $\mu_{\xi t}$ and $\sigma_{\xi t}$.

C. 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.

D. Data and the Auxiliary Parameters

Our primary dataset is the 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 dataset with a panel data element. To focus on as homogeneous a group as possible, the sample includes only male high school graduates. Estimation results for college graduates are presented in appendix F and look broadly similar.

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

$$W_{it} = H_{it}(1 - I_{it}) \tag{24}$$

in the model. The 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 nonemployment. To measure persistence in employment, we also collect data on the transition rates in and out of work.

An individual in SIPP is observed at most three times each year. Due to the seam bias problem in SIPP, we use only 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.[10] 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.

[10] 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.

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 overpredict 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, ages 22–65.[11]
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.[12]
iv) The second moments (standard deviation) of the logarithm of measured wages, ages 22–65.
v) The first moments of adult equivalent consumption, ages 22–65.[13]
vi) The Social Security benefit application rates, ages 62–70.
vii) The overall transition probabilities averaged between ages 35 and 50,[14]
   (*a*) from working to not working or
   (*b*) from not working to working.
viii) The average measured wage change rate after one nonemployment spell averaged between ages 41 and 65.[15]

[11] 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.

[12] To construct these moments, we first regress log wage on the age dummies and survey year dummies and obtain the predicted log wage, denoted z. We pick a base age (30) and calculate the average predicted log wage at the base age for each year, denoted $\bar{z}_{a,j}$, where a represents the base age and j represents the 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 $\Delta\bar{z}_{a,j}$. Finally, we calculate the difference between the original log wage and $\Delta\bar{z}_{a,j}$ and define the result as $\log \tilde{W}_{it}$, 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.

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

[14] 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 ages 41–65, does not change the results in any significant way.

[15] 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.

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 ages 22 and 65. Figure 1*B* plots two log measured wage profiles. The first is the profile from the pooled sample, while the second 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 ages 22 and 65. Figure 1*C* shows the extent to which the variance of log measured wages increases with age. Figure 1*D* presents the adult equivalent consumption profile, while figure 1*E* illustrates the two peaks at ages 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 datasets, including the National Longitudinal Survey of Older Men data (Johnson and Neumark 1996), the PSID data (Rupert and Zanella 2015), and the Health and Retirement Survey 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 one's sixties or late sixties. One could argue that this evidence is not consistent with the traditional human capital model since the traditional human capital model would predict a hump-shaped wage. The intuition is that when the human capital depreciation outweighs the investment, wages start to decline, which generates a hump-shaped profile. 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, 1 year apart.[16] We construct a similar short panel from the CPS March Annual Social and Economic Supplement files. 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

[16] For MORG data, they are the fourth and eighth interviews.

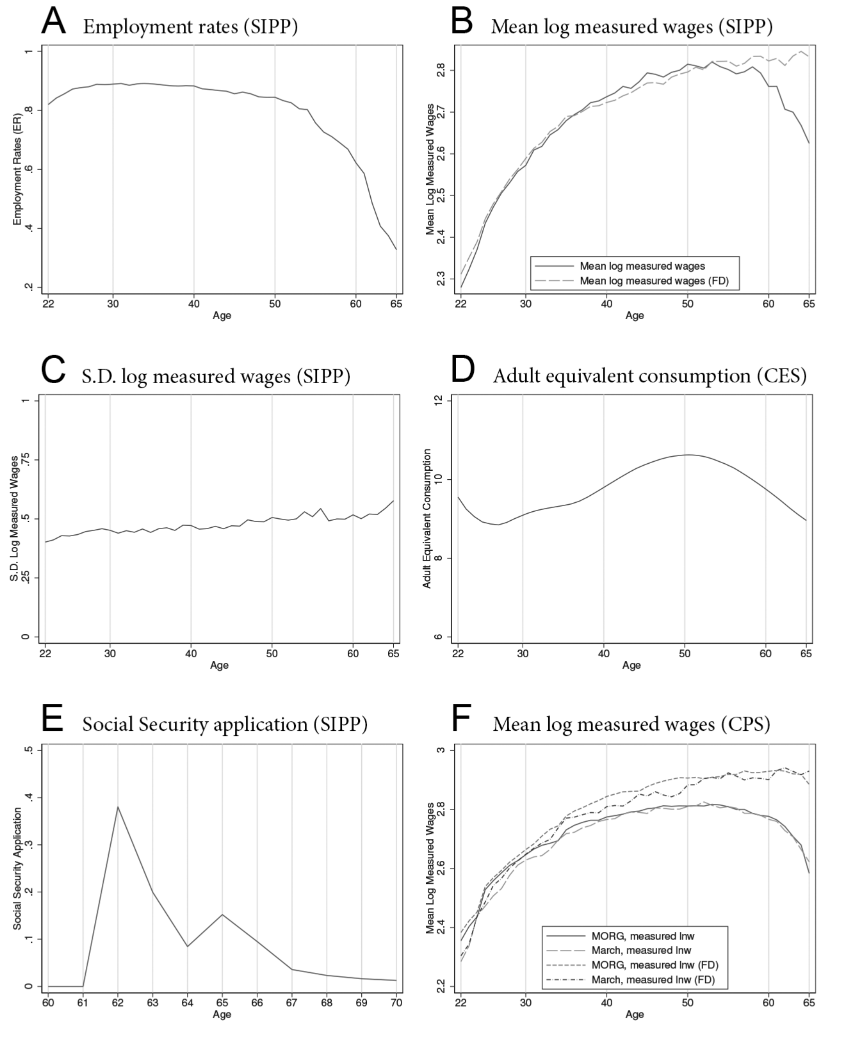

FIG. 1.—Data moments and profiles. All panels show variables changing over age. Panel *A* presents a dummy variable for employment from the SIPP—defined to be one if workers are working at least two of the three survey months in a year and zero otherwise. The solid line in panel *B* shows the raw mean logged wage for sample members in the SIPP who worked during the survey month, and the dashed line shows the age dummies from a regression on age dummies with individual-specific fixed effects using a first difference estimator. Panel *C* shows the raw standard deviation of the SIPP wage from the sample in *B*. Panel *D* uses the CES to show how the adult equivalent consumption evolves over age. Panel *E* plots the age of Social Security application from the SIPP. Panel *F* is analogous to *B* but uses the March CPS and MORG CPS.

TABLE 2
Transition Moments

Model	Transition Probability: Working to Not Working	Transition Probability: Not Working to Working	Wage Change Rate after One Nonemployment Spell
Data	.034	.200	−.071
Baseline model	.028	.214	−.079
No depreciation at work, version 1	.022	.196	−.065
No depreciation at work, version 2	.029	.201	−.056
Model with health and part-time	.030	.178	−.082

Note.—The transition rate is the average transition probability of working to not working in consecutive years between ages 35 and 50. The average wage change rate after one nonemployment spell is the average difference in log wages immediately following the nonemployment spell compared with immediately preceding the nonemployment spell between ages 41 and 65. The first row of data comes from the SIPP. The rest are simulated—each row corresponding to a different specification.

the 1979–2018 CPS MORG data and the 1989–2018 IPUMS-CPS March data (Flood et al. 2021).[17] 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 $H_{it}(1 - I_{it})$.

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.

V. 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 figure 2 and table 2.[18]

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

[17] Time fixed effects are filtered out, as described in n. 12.

[18] The overidentification test statistic is reported at the bottom of table 3. The model is rejected at the 1% level. The fact that we reject it 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.

TABLE 3
PARAMETER ESTIMATES IN THE BASELINE MODEL

Parameter	Symbol	Estimate	Standard Error
Human capital depreciation	δ	.087	.004
Human capital production function: *I* factor	α_I	.102	.012
Human capital production function: *H* factor	α_H	.050	.006
Standard deviation of human capital innovation	σ_ξ	.004	.001
Consumption: constant relative risk aversion	η_c	4.036	.027
Consumption shifter: coefficient on t ($\times 10^{-1}$)	φ_1	.320	.072
Consumption shifter: coefficient on t^2 ($\times 10^{-2}$)	φ_2	.165	.018
Consumption shifter: coefficient on t^3 ($\times 10^{-3}$)	φ_3	−.043	.001
Consumption shifter: coefficient on married	φ_4	.898	.160
Leisure: standard deviation of shock	σ_ε	.133	.018
Leisure: spouse not working	a_1	.201	.023
Leisure: spouse working	a_2	−.404	.053
Bequest weight	b_1	59,171,174	8,545,284
Parameter heterogeneity:			
Leisure: mean of intercept	μ_{a_0}	−5.628	.058
Leisure: standard deviation of intercept	σ_{a_0}	.697	.054
Human capital productivity, mean	μ_π	1.882	.026
Human capital productivity, standard deviation	σ_π	.750	.009
Correlation between a_0 and π	ρ	−.604	.063
Initial human capital level at age 18:			
Intercept	γ_0	1.467	.090
Coefficient on a_0	γ_{a_0}	.024	.008
Coefficient on π	γ_π	.577	.040
Standard deviation of error term	σ_{H_0}	.007	.013
Additional Social Security Application effects:			
Effect of resource constraint ($\times 10^{-3}$)	b_{62}	.285	.037
Effect of health insurance: constant ($\times 10^{-3}$)	b_{65}	.034	.021
Effect of health insurance: coefficient on t ($\times 10^{-3}$)	b_{65t}	.158	.032
χ^2 statistic[a] (df = 207)		809	

NOTE.—Indirect inference estimates. Estimates use a diagonal weighting matrix. 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.

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

cycle. The simulated employment rate increases slightly between ages 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 ages 50 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 a 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 do not decrease during the examined period (fig. 2*C*). Alternatively, figure 2*B* shows that the original log measured wage profile presents

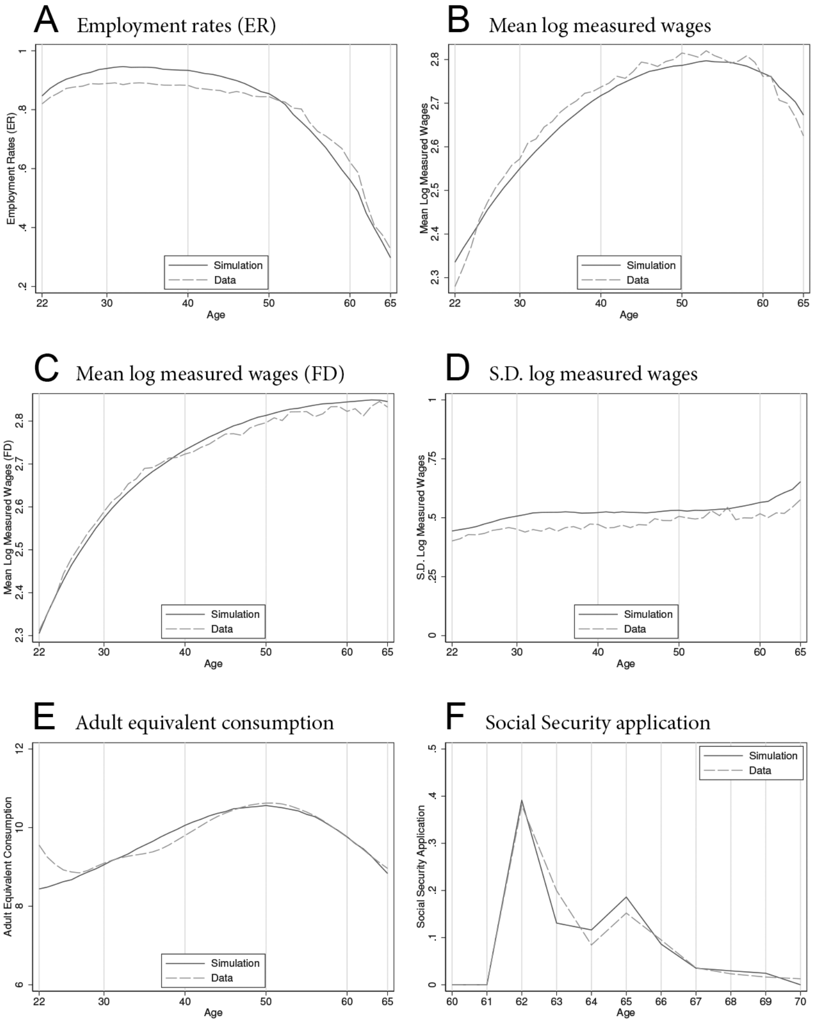

Fig. 2.—Fit of the baseline model. The dashed line in each panel is directly from the data and identical to what is shown in figure 1. The solid line simulates the same statistic over age from the baseline model. All figures show variables changing over age. Panel *A* presents a dummy variable for employment. Panel *B* shows the raw mean logged wage. Panel *C* shows the age dummies from a regression on age dummies with individual-specific fixed effects using a first-difference estimator. Panel *D* shows the raw standard deviation of wages from the sample in *B*. Panel *E* shows how the adult equivalent consumption evolves over age. Panel *F* plots the age of Social Security application.

a hump shape, which resembles the data profile. The model also replicates the log measured wage variation as in the data (fig. 2*D*).[19]

Our model tracks the hump shape and the level of the adult equivalent consumption profile reasonably well (fig. 2*E*), as well as the two peaks at ages 62 and 65 in the Social Security application (fig. 2*F*).[20] 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.

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. To be clear, we do have age effects in consumption and in Social Security take-up decisions. In the models, those features are important for fitting the consumption profile and the exact pattern of Social Security but not for the key patterns in our paper: the life-cycle wage and labor supply patterns. We show this explicitly by reestimating our model without these age-dependent features and fit these profiles well—but not the hump shape in consumption or the exact timing of Social Security application. These results can be found in figure E1.

A key feature of our model makes this fit 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 acknowledge that precautionary savings also cause 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.[21]

A. The Role of Human Capital Depreciation

Our estimate of human capital depreciation in the baseline model of 0.087 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*, in the data, once fixed effects

[19] 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 life-cycle patterns of human capital and labor force participation.

[20] We did not 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.

[21] Our focus in what follows is 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 will not be identical.

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. As we mention in the introduction, there are two different aspects of depreciation that are important.

The first is due to depreciation off the job. In the model, when a man does not work he cannot invest in human capital and his human capital falls. This idea is directly reflected in the data through the observed wage decline after one nonemployment spell, which we match. This future wage decline constitutes a shadow cost of not working. Most importantly, the magnitude of this cost varies across the life cycle. When workers are in their midcareer, 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. When older workers take time off, their human capital declines and they have even less reason to reenter 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 (labeled "Mean log measured wages") 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 only workers. From the latter curve, one sees at older ages (around 60) that 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 and (especially) from not working at all. When one looks at the large decline in human capital

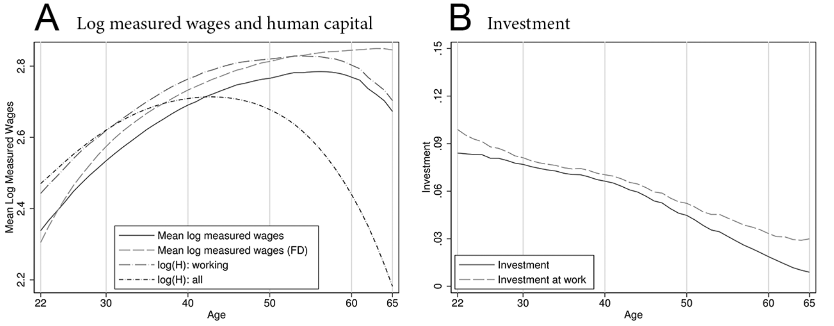

FIG. 3.—Log measured wages, human capital, and investments. These panels are all simulated life-cycle profiles from the baseline model. In panel *A*, the solid and dashed lines show the simulations we use to match wages in figure 2; the short-dash-dotted line and dash-dotted line show the log of human capital for all men and workers only, respectively. Panel *B* presents the average level of investment, I_{it}, across the life cycle for all men and for workers only.

between ages 50 and 60, it is not at all surprising that labor supply has declined.

The second feature is more subtle but perhaps more interesting. It arises from a point emphasized by Heckman, Lochner, and Taber (1998a): measured wages are different than human capital. This distinction between human capital and measured wages can help reconcile 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. The magnitudes of these effects are dictated by two parameters of the model: δ and α_I. The magnitude of the decline in H_{it} is determined primarily by the depreciation of human capital δ. The level of I_{it} around retirement age depends in large part on the size of the elasticity parameter of human capital investment α_I. In particular, the larger the elasticity, the more sensitive the investment to age and the steeper 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. So for an increase in $(1 - I_{it})$ to offset the relatively large value of depreciation, α_I needs to be relatively small. Our value of 0.102 is small compared with others in the literature (see Browning, Hansen, and Heckman 1999).[22] For example, Heckman, Lochner, and Taber (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 depreciation to zero.

One can see the evidence of this feature in figure 3*A*. The two lines to focus on are the solid line labeled "Mean log measured wages," which shows $\log(H_{it}(1 - I_{it}))$ for workers, and the long-dash-dotted line labeled "$\log(H)$: working," which shows $\log(H_{it})$ for workers. One can see that human capital for workers peaks around age 45, 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 is a feature that helps explain the fall in labor supply at older ages.

To further show the implication of our model for investment, figure 3*B* presents the investment profile in our model, which shows significant investment at older ages. The level and trend are very close to figure 4 in Mulligan (1998; reproduced in fig. E3), who calculates the time spent learning

[22] The low estimate of α_I comes from the combination of wage and employment patterns. To further confirm this, we estimated a model where we restrict $\alpha_I > 0.6$ and reestimate the baseline model. The estimates are presented in col. 6 of table E2 (tables C1, D1–D3, E1, E2, and F1 are available online)—we are not able to obtain a good fit with the data moments, as shown in fig. E2.

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. (2021), who find substantial training among older workers—though using data from the United Kingdom.

To show the significance of the human capital depreciation in matching the labor supply profile and the two log measured wage profiles, we reestimate the model without depreciation. We tried estimating a model with no depreciation at all but our best fit of this model was very poor so we do not discuss it. This again highlights the importance of our first channel. To understand the second channel, we reestimate a less extreme case in which we continue to have depreciation off the job but not while working.[23] In the latter alternative model, we assume that human capital depreciates only 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. Neither fit the data, but they miss in different ways. The fits of these two specifications are shown in the third and fourth rows of table 2 and figure 4; one can see that they miss different features of the data. Note that both models still do capture much of the main pattern—partly as a result of depreciation off the job. That is, we do see a large increase in wages at the beginning of the life cycle and 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 to fit the decline in labor supply it requires a larger labor supply elasticity (standard deviation of leisure shock falls from 0.133 to 0.071), which prevents it from fitting the labor supply level at the beginning of the life cycle. Version 2 improves the fit on the employment rates at the beginning by putting selection effects into the model at the cost of a worse fit on all three wage measures.[24]

Given that we have shown that our estimate of a depreciation value $\delta = 0.087$ plays an important role in explaining the pattern of wages and life-cycle labor supply, it is important to place this value into the range of

[23] We have also reestimated 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 (0.0872 on the job vs. 0.0865 off the job), so as our main specification we keep the version where they are a single parameter.

[24] 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 app. D.

TABLE 4
Two Versions of Estimates of Model with No Depreciation at Work

Parameter	Symbol	Estimate: Version 1	Estimate: Version 2
Human capital depreciation	δ	.063	.044
Human capital production function: I factor	α_I	.650	.614
Human capital production function: H factor	α_H	.001	.003
Standard deviation of human capital innovation	σ_ξ	.0001	.001
Consumption: constant relative risk aversion	η_c	3.944	3.946
Consumption shifter: on t ($\times 10^{-1}$)	φ_1	1.198	1.176
Consumption shifter: on t^2 ($\times 10^{-2}$)	φ_2	−.132	−.114
Consumption shifter: on t^3 ($\times 10^{-3}$)	φ_3	−.013	−.015
Consumption shifter: coefficient on married	φ_4	.066	.216
Leisure: standard deviation of shock	σ_ε	.071	.059
Leisure: spouse not working	a_1	.311	.343
Leisure: spouse working	a_2	−1.017	−1.727
Bequest weight	b_1	83,015,683	85,824,816
Parameter heterogeneity:			
Leisure: mean of intercept	μ_{a_0}	−4.945	−4.907
Leisure: standard deviation of intercept	σ_{a_0}	.664	.584
Human capital productivity, mean	μ_π	.537	.589
Human capital productivity, standard deviation	σ_π	1.007	1.507
Correlation between a_0 and π	ρ	−.804	−.472
Initial human capital level at age 18:			
Intercept	γ_0	2.477	2.382
Coefficient on a_0	γ_{a_0}	.092	.086
Coefficient on π	γ_π	.739	.765
Standard deviation of error term	σ_{H_0}	.004	.104
Additional Social Security Application effects:			
Effect of resource constraint ($\times 10^{-3}$)	b_{62}	.617	.460
Effect of health insurance: constant ($\times 10^{-3}$)	b_{65}	.471	.097
Effect of health insurance: coefficient on t ($\times 10^{-3}$)	b_{65t}	.473	.354
χ^2 statistic[a] (df = 207)		1,793	1,859

Note.—Indirect inference estimates. Estimates use a diagonal weighting matrix. 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 human capital production.

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

estimates in the literature. This is not easily done, as there is a very large range of estimates and none is directly comparable with our number. Some are larger than our 8.7% 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%.[25] A second literature looks at displacement

[25] A classic early paper on this topic is Mincer and Polachek (1974), who estimate a net depreciation rate of around 1.5% per year. Mincer and Ofek (1982) go beyond this to discuss the difference between short- and long-term losses from interruption. In the long run, individuals invest in human capital to offset the initial loss, so Mincer and Ofek's (1982) definition of short-term losses is more closely related to our concept of depreciation. Using panel

from the Displaced Worker Survey and also finds a wide range of estimates—many of which are not directly comparable with ours.[26] A third literature examines the effect of the length of an unemployment spell on the wage at rehire. A recent and convincingly identified paper of this type is Schmieder, von Wachter, and Bender (2016). 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 44. They see an increase in unemployment duration at these two discontinuity points, so they use the kink points as instruments to estimate the effect of the length of unemployment duration on reemployment wages. They find that one extra month of unemployment leads to a decrease in wages of 0.8%, which gives an annual rate close to our estimate of 8.7%. Keane and Wolpin (1997) is a paper similar in style 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 a style similar to our paper in the sense that it is a structural life-cycle model of labor supply and human capital formation. Interestingly, their analysis also reveals a substantial depreciation of human capital ranging from 6% to 11%.

B. 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

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 nonparticipation in the labor market leads to a decline in earnings of 25%. Kunze (2002), Gorlich and de Grip (2009), and Adda, Dustmann, and Stevens (2017) all use German data (Institut für Arbeitsmarkt- und Berufsforschung 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% wage losses 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, Dustmann, and Stevens (2017) find a range of estimates, typically with small numbers but the largest being 6.9%.

[26] While much of this literature is more focused on earnings than wages, some papers look at weekly earnings. Both Ruhm (1991) and Farber (1993) estimate the effect of a displacement on reemployment 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 simply from the incidence of displacement. Li (2013) uses the same data but produces annualized versions so that the effects can be more easily compared with 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 depreciation of 9.4% for installation and repair workers, 7.7% for production workers, and 17.4% for workers in transportation.

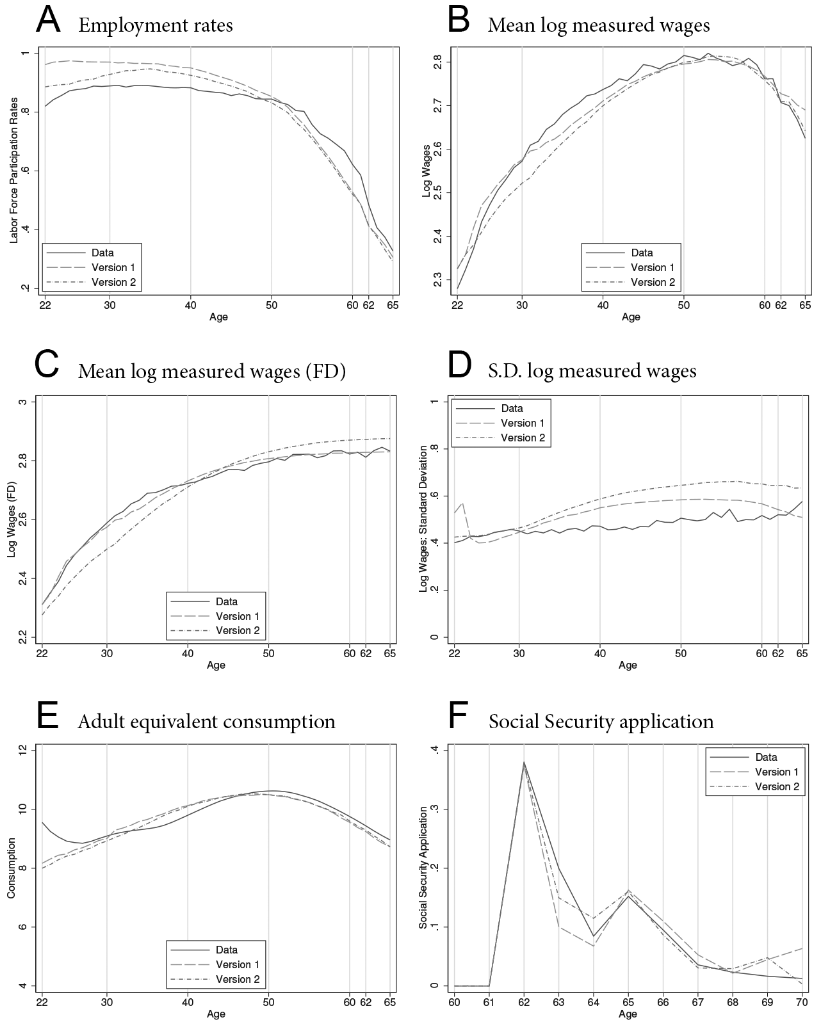

FIG. 4.—Fit of the alternative models with no depreciation at work. The solid line in each panel is identical to what is shown in figure 1. "Version 1" (dashed lines) and "Version 2" (short-dash-dotted lines) refer to two different versions of the estimation results of the "no depreciation at work" model, with estimates presented in table 4. All panels show variables changing over age. Panel *A* presents a dummy variable for employment. Panel *B* shows the raw mean logged wage. Panel *C* shows the age dummies from a regression on age dummies with individual-specific fixed effects using a first-difference estimator. Panel *D* shows the raw standard deviation of wages from the sample in *B*. Panel *E* shows how the adult equivalent consumption evolves over age. Panel *F* plots the age of Social Security application.

is exactly indifferent between working or not, in which case it is infinite. Therefore, we cannot construct the standard Marshallian and Hicksian labor supply elasticities. However, to compare our elasticity with 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% (i.e., 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 call it the analogue to the empirical elasticity (AEE) and define it formally below as it is an approximation of the extensive margin Frisch elasticity.

Let h_t^b represent the employment rate at age t in the baseline model and h_t^t represent 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%. We then define that elasticity to be

$$\text{AEE} \equiv \frac{\log(h_t^t) - \log(h_t^b)}{\log(w_t^t) - \log(w_t^b)}. \tag{26}$$

While not perfect, we find this to be a convenient summary statistic.[27]

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 prime age range 30–50. This basic pattern is similar to plots in Keane and Wasi (2016) that also show U-shaped elasticities.[28] 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 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, Hansen, and Heckman 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 is 0.75 or greater given empirically reasonable levels of nonconvexities or fixed costs.

[27] 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.

[28] Our levels are not directly comparable with theirs.

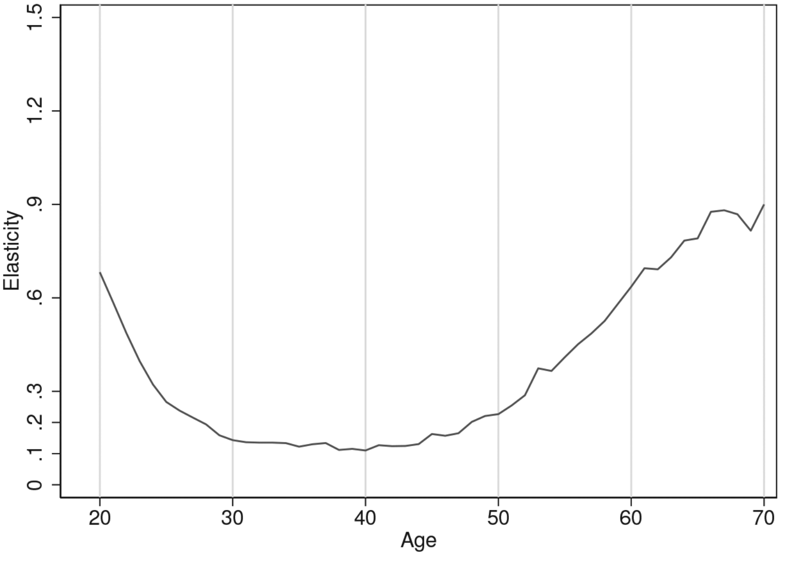

FIG. 5.—Analogue to the empirical elasticity of labor supply. The solid line presents simulations of the average value of $(\log(h_t^t) - \log(h_t^b))/(\log(w_t^t) - \log(w_t^b))$ over the life cycle.

The average of our estimates between ages 55 and 65 is remarkably close to theirs.

VI. 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 that we can fit the data 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). Given the importance of disability insurance in the United States for this group, it is also important that we incorporate disability. We do this by including it 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.

A. *Health and Disability*

Our main addition 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 (SSDI) or Supplemental Security Income (details in app. sec. 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.[29]

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

$$\gamma(a_{i0}, M_{it}, S_{it}, \varepsilon_{it}) = \exp\left(a_{i0} + \sum_{j=1}^{2} a_j \mathbf{1}\{M_{it} = j\} + a_{hS}^0 + a_{hS}^1 t + \varepsilon_{it}\right). \quad (27)$$

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.[30] We normalize $a_{hg}^0 = 0$ and assume that leisure taste changes only for nonhealthy people—that is, we assume that $a_{he}^1 = a_{hg}^1 = 0$ but estimate a_{hb}^1 and a_{hd}^1.

To estimate these five new parameters, $\{a_{he}^0, a_{hb}^0, a_{hd}^0, a_{hb}^1, a_{hd}^1\}$, we include three more sets of moment conditions: the difference in employment rates between workers with excellent health and workers with good health, the employment rate difference between workers with good health and workers with bad health, and the difference between workers with bad health and those with disability, across ages 50–65. The data moments are derived from the CPS March data, and the raw patterns can be seen in figure 6*G*.

[29] 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. Figures B2*a*–B2*c* plot the health transition probabilities at each age; fig. B2*d* plots the distribution of four health statuses. Between ages 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).

[30] 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 fit the interaction of health and labor supply in the data only 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. Health modeled this way also captures its main effect on future wages via human capital accumulation (e.g., Hokayem and Ziliak 2014; Capatina, Keane, and Maruyama 2020).

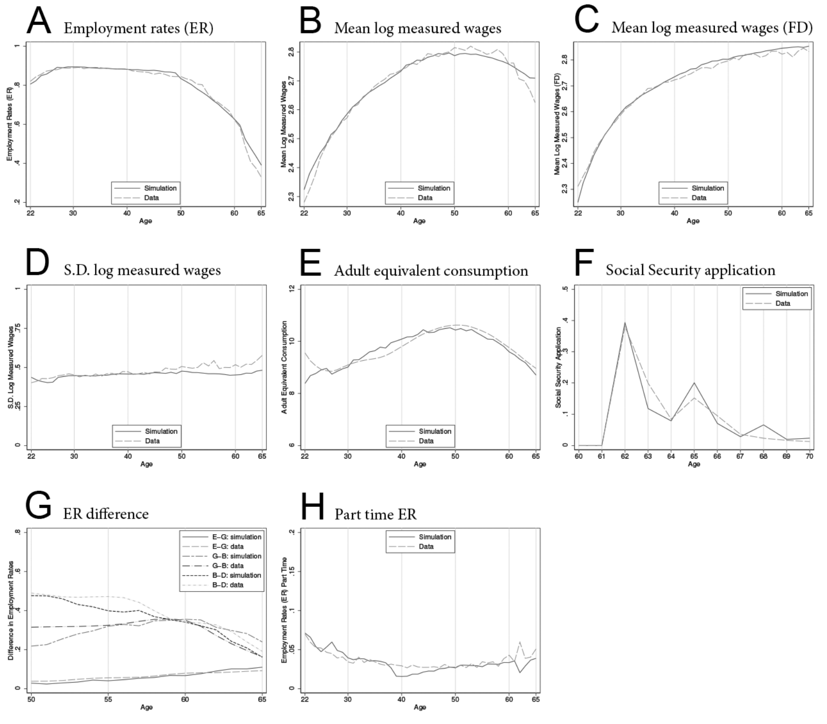

FIG. 6.—Fit of the extended model with health, disability, and part-time option. The dashed line in *A*–*F* is identical to what is shown in figure 1, while the solid line simulates the same statistic over age from the extended model. All panels show variables changing over age. Panel *A* presents a dummy variable for employment. Panel *B* shows the raw mean logged wage. Panel *C* shows the age dummies from a regression on age dummies with individual-specific fixed effects using a first-difference estimator. Panel *D* shows the raw standard deviation of wages from the sample in *B*. Panel *E* shows how the adult equivalent consumption evolves over age. Panel *F* plots the age of Social Security application. In panel *G*, "E–G," "G–B," and "B–D" refer to the difference in employment rates between workers with excellent health (E) and workers with good health (G), the employment rate difference between workers with good health (G) and workers with bad health (B), and the difference between workers with bad health (B) and those with disability (D), respectively; "Data" presents mean values from the CPS, while "Simulation" refers to their analogous summary statistics simulated from the extended model. Panel *H* is analogous to *A*, which shows part-time employment rates from the SIPP (dashed line) and the simulations from the model (solid line).

B. Part-Time Work

We also include the part-time work as a choice in the extended model. At each period, an individual decides to work full-time ($\ell_{it} = 0$), 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.[31] One can see from figure 6*H* that part-time work is uncommon for the sample we study. We model working part-time as spending half of the time in the labor market and the other half at leisure. We let ϱ_t be the parameter that determines utility from part-time work. It varies across ages but not across individuals, and we assume that the utility of leisure associated with part-time work is $\gamma(a_0, M_{it}, \varepsilon_{it})\varrho_t$, where

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

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, $\{a_\varrho^0, a_\varrho^1, a_\varrho^2\}$, 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 $I_{it} \in [0, 1/2]$ and the effective work time is $1/2 - I_{it}$, 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[32]

$$\ell_{it} = \underset{\ell \in \{0,p,1\}}{\arg\max}\left\{\tilde{V}_{t\ell}\left(\tilde{X}_{it}\right) + \gamma_{it}(\mathbf{1}\{\ell = 1\} + \varrho_t \mathbf{1}\{\ell = p\})\right\}, \tag{28}$$

where

$$\begin{aligned}&\left\{\mathcal{C}_{itp}\left(\tilde{X}_{it}\right), \mathcal{I}_{itp}\left(\tilde{X}_{it}\right), \mathcal{SSA}_{itp}\left(\tilde{X}_{it}\right)\right\}\\ &\equiv \underset{c,I,\text{ssa}}{\arg\max}\left\{\psi_{tM_{it}}\frac{c^{1-\eta_c}}{1-\eta_c} + \varpi_t(\text{ssa}) + \beta E\left[V_{t+1}(X_{it+1})|\tilde{X}_{it}, c, \ell_{it} = p, I, \text{ssa}\right]\right\},\end{aligned} \tag{29}$$

and we define

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

The details of solving this model depend on the three values of $\tilde{V}_{t\ell}(\tilde{X}_{it})$ as well as ϱ_t. We discuss the details in appendix A. In the case in which all

[31] 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.

[32] Note that now the $\tilde{X}_{it}$ includes the health status S_{it} defined previously.

three options may be chosen,[33] the model is like an ordered probit with cutoffs

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

where

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

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

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

C. *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 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 and the slight increase in part-time work before retirement.[34]

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,[35] to control for health we simulate a counterfactual in which there was no health change. We eliminate the importance of health for individuals 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 at and after age 50, the second experiment assumes that their taste for leisure

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

[34] This is something Iskhakov and Keane (2021) had trouble matching.

[35] 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

Parameter	Symbol	Estimate	Standard Error
Human capital depreciation	δ	.094	.001
Human capital production function: I factor	α_I	.117	.006
Human capital production function: H factor	α_H	.090	.004
Standard deviation of human capital innovation	σ_ξ	.048	.002
Consumption: constant relative risk aversion	η_c	3.968	.034
Consumption shifter: coefficient on t ($\times 10$)	φ_1	.195	.008
Consumption shifter: coefficient on t^2 ($\times 10^2$)	φ_2	.103	.003
Consumption shifter: coefficient on t^3 ($\times 10^3$)	φ_3	−.033	.001
Consumption shifter: coefficient on married	φ_4	1.704	.066
Leisure: standard deviation of shock	σ_ε	.453	.022
Leisure: spouse not working	a_1	−.259	.011
Leisure: spouse working	a_2	−.694	.045
Leisure: excellent health	a^0_{he}	−.259	.012
Leisure: bad health	a^0_{hb}	.270	.019
Leisure: bad health time trend	a^1_{hb}	.011	.001
Leisure: disabled	a^0_{hd}	2.947	.088
Leisure: disabled time trend	a^1_{hd}	.016	.001
Part-time utility: constant	a^0_ϱ	−1.314	.043
Part-time utility: coefficient on t ($\times 10$)	a^1_ϱ	.251	.014
Part-time utility: coefficient on t^2 ($\times 10^2$)	a^2_ϱ	−.013	.001
Bequest weight	b_1	29,080,438	2,396,900
Parameter heterogeneity:			
Leisure: mean of intercept	μ_{a_0}	−5.867	.069
Leisure: standard deviation of intercept	σ_{a_0}	2.319	.091
Human capital productivity, mean	μ_π	1.826	.042
Human capital productivity, standard deviation	σ_π	.652	.028
Correlation between a_0 and π	ρ	−.548	.023
Initial human capital level at age 18:			
Intercept	γ_0	2.207	.200
Coefficient on a_0	γ_{a_0}	.293	.017
Coefficient on π	γ_π	1.013	.092
Standard deviation of error term	σ_{H_0}	.010	.017
Additional Social Security Application effects:			
Effect of resource constraint ($\times 10^3$)	b_{62}	.272	.023
Effect of health insurance: constant ($\times 10^3$)	b_{65}	.012	.001
Effect of health insurance: coefficient on t ($\times 10^3$)	b_{65t}	.289	.026
χ^2 statistic[a] (df = 291)		913	

NOTE.—Indirect inference estimates. Estimates use a diagonal weighting matrix. 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.

[a] This is the J-statistic. The critical value of the χ^2 distribution is $\chi^2_{(291,0.01)} = 350$.

does not change with age. That is, we assume that the taste for leisure is now

$$\gamma(a_{i0}, M_{it}, S_{it}, \varepsilon_{it}) = \exp\left(a_{i0} + \sum_{j=1}^{2} a_j \mathbf{1}\{M_{it} = j\} + a^0_{hS} + a^1_{hS} \cdot 50 + \varepsilon_{it}\right), \forall\ t > 50. \quad (31)$$

We then re-solve the modified model and simulate the life-cycle profile for each individual using the same estimates from the aforementioned

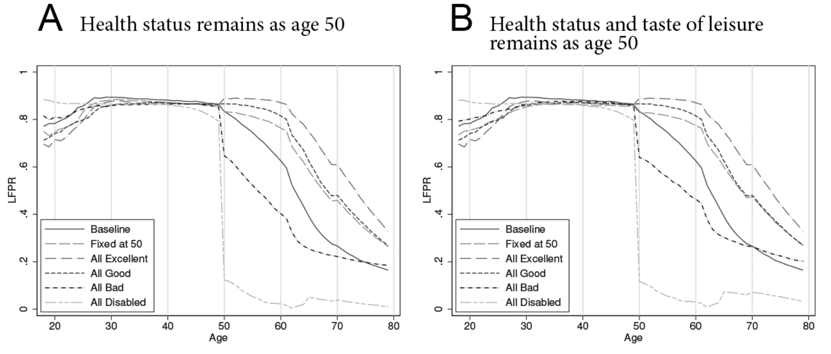

FIG. 7.—Sensitivity to health preferences: employment rates. This figure shows the life-cycle labor supply with alternative counterfactuals about health. "Baseline" in both panels is from our extended model. "Fixed at 50" refers to the experiment where the health status (*A*) or the health status and taste for leisure (*B*) remains as age 50. "All Excellent," "All Good," "All Bad," or "All Disabled" refers to the experiment where the health status (*A*) or the health status and taste for leisure (*B*) becomes excellent/good/bad/disabled permanently at age 50, respectively.

extended model with health and the part-time option.[36] 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 simulation from the extended model is 13.3%. When we assume that the taste of leisure does not vary with age for all health status, the difference in labor supply between the second counterfactual and the simulations from the extended model is only slightly larger, 14.2%. 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 small effect arises from the fact that bad health is relatively uncommon, not from the fact that it does not 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

[36] 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.

this by assuming that a worker's health status becomes excellent (or good/bad/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 by simulating a counterfactual that removes this 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.

VII. Changes in Tax and Social Security

The above 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 nine counterfactual policy experiments that reflect various changes in tax and Social Security rules:

i) Increase taxes proportionally by 50% (e.g., from 10% to 15%).
ii) Eliminate the Social Security earnings test.
iii) Increase the NRA from 65 to 67.
iv) Reduce Social Security benefits by 20%.
v) Make the Social Security benefit less progressive.
vi) Make the Social Security benefit depend on the monthly earning of the last working year.

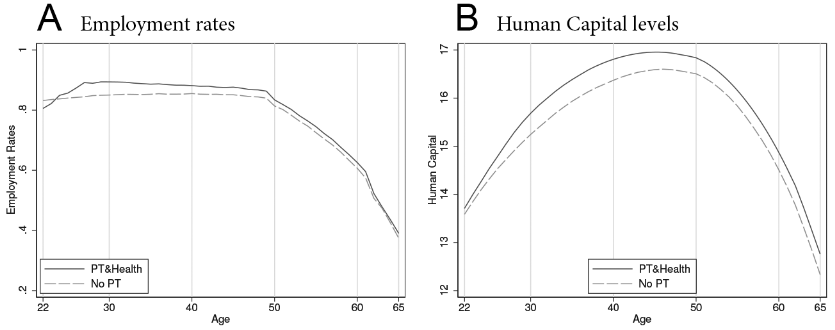

FIG. 8.—Sensitivity to part-time option: turn off the part-time option. Panel *A* presents labor supply over the life cycle, while panel *B* presents human capital levels. "PT&Health" refers to the extended model with health, disability, and a part-time option. "No PT" refers to the experiment that removes the part-time option.

vii) Eliminate Social Security taxes.
viii) Eliminate Social Security benefits.
ix) Eliminate the 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 as accounting for equilibrium effects on wages.

The results of these experiments are summarized in columns 2–10 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.[37]

Column 2 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 provides a modest 0.84% more of the total lifetime labor supply, or four additional months over the life cycle.[38] Most novel is the effect on human capital investment, which increases by 0.76%, leading to a 0.84% increase in the human capital level and a 0.16% increase in the measured log wages. The magnitude of the change in investment and human capital is similar to the change in labor supply. Investment can increase for two reasons: it almost mechanically increases when labor supply increases but also increases directly (conditional on working). We find that the orders of magnitude of the two effects are similar but that the second is the larger of the two.

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 0.94%, and the pretax wage income increases by 1.51%.[39] Annual consumption reduces by 3.2%. Note that these effects are averaged across the life cycle. Figure 9*A* shows how the effects change at different ages.

[37] To make these two models comparable, we keep the SSDI and Supplemental Security Income benefit unchanged in all experiments.

[38] In our model, leisure and consumption are separable. In the simplest static form of that model without human capital, whether the income or substitution effect dominates depends on whether η_c is larger or smaller than one. We estimate η_c to be around four, which is well within the estimates in the literature, so it is not surprising that the income effect dominates the substitution effect.

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

TABLE 6

Effects of Changing Taxes or Social Security Rules

	Baseline (Level) (1)	Tax Increase 50% (%Δ) (2)	No Earnings Test (%Δ) (3)	NRA = 67 (%Δ) (4)	Reduce SSB 20% (%Δ) (5)	Linear PIA (%Δ) (6)	AIME = Last Wage (%Δ) (7)	No SS Taxes (%Δ) (8)	No SS Benefit (%Δ) (9)	No SS System (%Δ) (10)
	A. Extended Model with Health and Part-Time Option									
Labor supply	41.509	.843	.021	.423	.533	.252	−.325	−3.474	5.967	1.234
Labor supply (full-time equivalent)	.662	.921	.114	.505	.743	.365	−.403	−3.248	6.495	1.904
Labor supply (part-time)	.033	1.946	−16.725	−1.547	−7.345	−9.686	−2.379	−12.870	58.955	48.834
Labor supply (full-time)	.645	.894	.633	.559	.953	.625	−.351	−2.998	5.135	.688
Effective labor	.606	.935	.024	.491	.800	.337	−.880	−3.210	6.684	2.084
Wage income	10.376	1.509	.129	.638	.705	.602	−.425	−3.871	6.645	1.250
Average lnw	2.663	.159	.027	.076	.125	−.002	−.030	−.118	−.348	−.678
Human capital	14.105	.836	.048	.254	.234	.071	.053	−2.669	4.034	.969
Investment	.055	.759	.082	.669	.124	.673	4.861	−3.659	4.410	.077
Consumption	9.010	−3.203	.019	−.305	−.991	.300	−.064	4.708	−3.566	1.172
	B. Baseline Model									
Labor supply	40.672	2.172	.319	.724	1.370	.625	−1.141	−6.549	8.793	.372
Effective labor	.612	2.198	.303	.725	1.367	.594	−1.568	−6.516	8.877	.254
Wage income	10.062	3.691	.317	.922	1.866	.620	−1.621	−8.999	11.132	.391
Average lnw	2.563	1.091	.173	.247	.592	−.398	.150	−2.213	2.670	−.208
Human capital	13.761	2.110	.218	.518	1.045	.475	−.606	−5.642	5.893	.735
Investment	.044	1.810	.546	.705	1.409	1.059	4.835	−7.008	7.620	.017
Consumption	9.008	−2.518	.052	−.403	−.486	.406	−.630	3.012	−2.411	.330

Note.—“Level” refers to the annual value averaged over the whole life cycle except for “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 40.672 years from age 18 to 80. “%Δ” refers to the percentage of the difference of the total value between the current experiment and the baseline model relative to the level in the baseline model. For example, in col. 4, the labor supply increases by 0.724% of the labor supply in the baseline model. “Linear PIA” refers to the modified Social Security benefit formula in eq. (C4): $\text{PIA}_{it} = 0.4475 \cdot \text{AIME}_{it}$. “AIME = Last Wage” refers to the modified AIME calculation in eq. (C5): $\text{AIME}_{it+1} = w_t \cdot \mathbf{1}\{l_t < 1\} + \text{AIME}_{it} \cdot \mathbf{1}\{l_t = 1\}$. For “Labor supply (full-time equivalent),” working part-time is counted as 0.5. “Wage income” refers to the pretax wage income.

The manner in which Social Security rules affect labor supply and wages is of central interest to policy makers. The remaining eight experiments are devoted to answering these questions. In the first five, we manipulate the current Social Security rules (cols. 3–7), while in the last three we decompose the distortionary effects of the current Social Security system (cols. 8–10).

First we remove the Social Security earnings test, which is effective between ages 62 and 70. In the second one, we delay NRA by 2 years: the new NRA is age 67 in this counterfactual experiment, while it is age 65 in the baseline model.[40] In the third one, we reduce the Social Security benefit proportionally by 20%. The results are presented in columns 3–5 in table 6 and in figure 9*A*. Removing the Social Security earnings test between ages 62 and 70 has a smaller effect on most variables; delaying the NRA by 2 years has a larger impact, and reducing the generosity of the Social Security benefit has the largest impact on most variables. For instance, in the extended model they increase the labor supply by 0.02%, 0.42%, and 0.53%, respectively. We again find that changes in investment and human capital are of similar orders of magnitude as the change in labor supply (though which is largest varies across the counterfactuals). 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*. In terms of understanding the economics behind the effects, delaying the NRA is particularly interesting as the direct effect does not occur until later in life. We can see that it induces higher human capital investment (both directly and through employment contributing roughly the same amount), resulting in persistently higher human capital levels and therefore higher wages at older ages. As seen in figure 9*A*, the wage difference is small before age 60 but increases substantially after that, reaching 2% 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 0.4%, 27%, and 43%, respectively.

In the next two experiments, we introduce more dramatic changes to the current Social Security rules. In column 6 of table 6, we make the Social Security benefits less progressive by totally flattening the benefit formula in equation (C4),

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

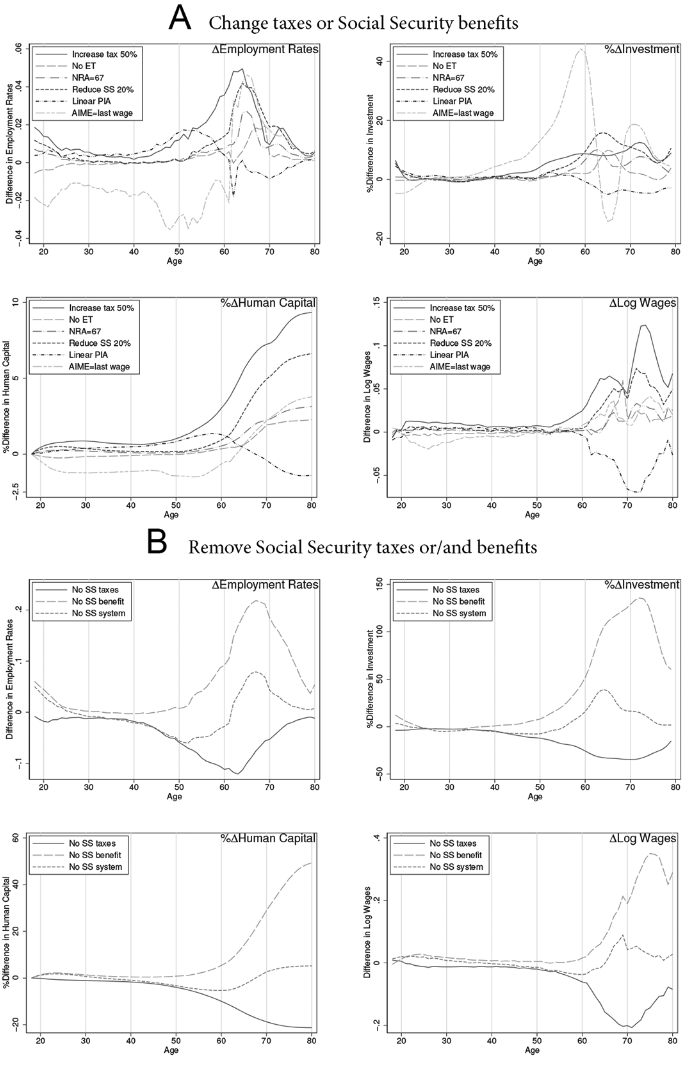

FIG. 9.—Policy experiments: change taxes or Social Security benefits. For each panel, the lines represent the difference between the counterfactual experiment versus the simulated values from the extended model over the life cycle, either in levels (Δ) or percentages (%Δ). Each line comes from a different policy experiment described above. "Increase tax 50%" refers to increasing taxes proportionally by 50%. "No ET" refers to no Social Security earnings test. "NRA = 67" refers to moving the NRA to age 67. "Reduce SS 20%" refers to reducing Social Security benefits by 20%. "Linear PIA" refers to a flat Social Security benefit. "AIME=last wage" refers to changing the formula so that Social Security benefits depend only on the last wage before claiming. "No SS taxes" refers to eliminating Social Security taxes. "No SS benefits" refers to eliminating Social Security benefits. "No SS system" refers to eliminating both taxes and benefits.

$$\text{PIA}_{it} = \lambda \cdot \text{AIME}_{it},$$

where $\lambda = 0.4475$ is chosen such that the Social Security benefit is close to budget neutral in the extended model with health and part-time option.[41] A less progressive Social Security system induces more investment in the human capital and slightly more labor supply (1 month). In column 7, we modify the calculation of AIME—for example, equation (C5),

$$\text{AIME}_{it+1} = w_t \mathbf{1}\{l_t < 1\} + \text{AIME}_{it} \mathbf{1}\{l_t = 1\}$$

such that the AIME depends on the monthly earning of the last working year only. This induces human capital investment to be redistributed more toward Social Security claiming ages, and overall it reduces labor supply by 1.6 months. This case leads to a large increase in human capital investment—much larger than the effect on labor supply. This makes sense as this policy directly rewards human capital accumulation rather than indirectly by affecting labor supply.

In the last three experiments, we decompose the effect of the current US Social Security system into the individual effects of the Social Security taxes and the Social Security benefit. In column 8 of table 6, we keep the Social Security benefit but eliminate the Social Security taxes (the payroll taxes);[42] in column 9, we remove the Social Security benefit completely but keep the Social Security taxes; in column 10, 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 3.5% less labor or 17 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, 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 6.0%, which is higher than the 3.5% 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 the benefit are removed. Column 10 indicates that eliminating the current Social Security system increases average labor supply by 1.2% or 6 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. In all three experiments, the effects on investment and human

[41] The "SS Benefit" in panel A of table D3 is −0.006% in col. 5, close to zero.

[42] The income taxes are still effective.

capital are of a similar order of magnitude as labor supply, except in the last experiment where the change in investment is much smaller.

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 ages 60 and 75):[43] as high as 3%–7% when removing the earnings test, delaying NRA by 2 years, reducing the Social Security benefit, or adjusting the AIME calculation; up to −7% when making Social Security less progressive; up to 35% when removing Social Security benefits; up to −20% when removing Social Security taxes; or up to −9% 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.

VIII. Conclusion

This paper develops and estimates a rich life-cycle model that merges a Ben-Porath-style human capital framework with a neoclassical-style life-cycle model of endogenous labor supply and uses it to examine changes in the taxes and the Social Security system. We use it to study life-cycle labor supply with a particular focus on older individuals, which is typically referred to as retirement. 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 working for older workers. We show that the parsimonious model is able to fit the main features of life-cycle measured wages (with and without fixed effects), labor supply, and 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 reestimate the model without allowing depreciation on the job and show that the model cannot fit the data as well. We also estimate an extension

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

of the model allowing for both health shocks and a 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 preset parameters.

We use the estimated model to simulate the impacts of various policy changes. While previous 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 magnitude of these results is roughly similar to the change in labor supply. 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.

Data Availability

Data and the code to replicate the results in the tables and figures can be found in the Harvard Dataverse, https://doi.org/10.7910/DVN/EHXLE4 (Fan, Seshadri, and Taber 2023).

References

Acemoglu, Daron, and Jörn-Steffen Pischke. 1998. "Why Do Firms Train? Theory and Evidence." *Q.J.E.* 113 (1): 79–119.

———. 1999. "The Structure of Wages and Investment in General Training." *J.P.E.* 107 (3): 539–72.

Adda, Jérôme, Christian Dustmann, and Katrien Stevens. 2017. "The Career Costs of Children." *J.P.E.* 125 (2): 293–337.

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *Q.J.E.* 109 (3): 659–84.

Altonji, Joseph G. 1986. "Intertemporal Substitution in Labor Supply: Evidence from Micro Data." *J.P.E.* 94 (3, pt. 2): S176–S215.

Autor, David H., and Mark G. Duggan. 2006. "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." *J. Econ. Perspectives* 20 (3): 71–96.

Becker, Gary. 1962. "Investment in Human Capital: A Theoretical Analysis." *J.P.E.* 70 (5): 9–49.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *J.P.E.* 75 (4): 352–65.

Blau, David. 2008. "Retirement and Consumption in a Life Cycle Model." *J. Labor Econ.* 26 (1): 35–71.

Blau, David, and Tetyana Shvydko. 2011. "Labor Market Rigidities and the Employment Behavior of Older Workers." *Indus. and Labor Relations Rev.* 64 (3): 464–84.

Blundell, Richard, Monica Costa-Dias, David Goll, and Costas Meghir. 2021. "Wages, Experience and Training of Women over the Lifecycle." *J. Labor Econ.* 39 (S1): S275–S315.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathon Shaw. 2016. "Female Labour Supply, Human Capital and Welfare Reform." *Econometrica* 84 (5): 1705–53.

Browning, Martin, Lars Peter Hansen, and James J. Heckman. 1999. "Micro Data and General Equilibrium Models." In *Handbook of Macroeconomics*, vol. 1A, edited by John B. Taylor and Michael Woodford, 525–602. Amsterdam: North-Holland.

Capatina, Elena, Michael Keane, and Shiko Maruyama. 2020. "Health Shocks and the Evolution of Earnings over the Life-Cycle." Working paper.

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

———. 2013. "Revisiting the Hump-Shaped Wage Profile." Working paper.

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

Currie, Janet, and Brigitte C. Madrian. 1999. "Health, Health Insurance and the Labor Market." In *Handbook of Labor Economics*, vol. 3C, edited by Orley C. Ashenfelter and David Card, 3309–416. Amsterdam: North-Holland.

De Nardi, Mariacristina. 2004. "Wealth Inequality and Intergenerational Links." *Rev. Econ. Studies* 71 (3): 743–68.

Fan, Xiaodong, Ananth Seshadri, and Christopher Taber. 2023. "Replication Codes for 'Estimation of a Life-Cycle Model with Human Capital, Labor Supply, and Retirement.'" Harvard Dataverse, https://doi.org/10.7910/DVN/EHXLE4.

Farber, Henry. 1993. "The Incidence and Costs of Job Loss: 1982–91." *Brookings Papers Econ. Activity (Microeconomics)* (1): 73–132.

Fernández-Villaverde, Jesús, and Dirk Krueger. 2007. "Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data." *Rev. Econ. and Statis.* 89 (3): 552–65.

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

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

French, Eric, and John Bailey Jones. 2011. "The Effects of Health Insurance and Self-Insurance on Retirement Behavior." *Econometrica* 79 (3): 693–732.

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

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

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

Gustman, Alan L., and Thomas L. Steinmeier. 1986. "A Structural Retirement Model." *Econometrica* 54 (3): 555–84.

Haan, Peter, and Victoria Prowse. 2014. "Longevity, Life-Cycle Behavior and Pension Reform." *J. Econometrics* 178:582–601.

Haley, William. 1976. "Estimation of the Earnings Profile from Optimal Human Capital Accumulation." *Econometrica* 44 (6): 1223–88.

Heckman, James J. 1975. "Estimates of a Human Capital Production Function Embedded in a Life-Cycle Model of Labor Supply." In *Household Production and Consumption*, edited by N. Terleckyj, 99–138. New York: Columbia Univ. Press.

———. 1976. "A Life-Cycle Model of Earnings, Learning, and Consumption." *J.P.E.* 84 (4, pt. 2): S11–S44.

Heckman, James J., Lance Lochner, and Ricardo Cossa. 2003. "Learning-by-Doing vs. On-the-Job Training: Using Variation Induced by the EITC to Distinguish between Models of Skill Formation." In *Designing Inclusion: Tools to Raise Low-End Pay and Employment in Private Enterprise*, edited by Edmund S. Phelps, 74–130. Cambridge: Cambridge Univ. Press.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998a. "Explaining Rising Wage Inequality: Explanations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Rev. Econ. Dynamics* 1 (1): 1–58.

———. 1998b. "Tax Policy and Human-Capital Formation." *A.E.R.: Papers and Proc.* 88 (2): 293–97.

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

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In *Handbook of the Economics of Education*, vol. 1, edited by Eric Hanushek and Finis Welch, 307–458. Amsterdam: North-Holland.

Hokayem, Charles, and James P. Ziliak. 2014. "Health, Human Capital, and Life Cycle Labor Supply." *A.E.R.* 104 (5): 127–31.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes. 1995. "Precautionary Saving and Social Insurance." *J.P.E.* 103 (2): 360–99.

Imai, Susumu, and Michael P. Keane. 2004. "Intertemporal Labor Supply and Human Capital Accumulation." *Internat. Econ. Rev.* 45 (2): 601–41.

Iskhakov, Fedor, and Michael Keane. 2021. "Effects of Taxes and Safety Net Pensions on Life-Cycle Labor Supply, Savings and Human Capital: The Case of Australia." *J. Econometrics* 223 (2): 401–32.

Johnson, Richard W., and David Neumark. 1996. "Wage Declines among Older Men." *Rev. Econ. and Statis.* 78 (4): 740–48.

Keane, Michael P., and Nada Wasi. 2016. "Labour Supply: The Roles of Human Capital and the Extensive Margin." *Econ. J.* 126 (592): 578–617.

Keane, Michael P., and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *J.P.E.* 105 (3): 473–522.

Kunze, Astrid. 2002. "The Timings of Careers and Human Capital Depreciation." IZA Working Paper no. 509, Inst. Labor Econ., Bonn, Germany.

Kuruscu, Burhanettin. 2006. "Training and Lifetime Income." *A.E.R.* 96 (3): 832–46.

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

Light, Audrey, and Manuelita Ureta. 1995. "Early-Career Work Experience and Gender Wage Differentials." *J. Labor Econ.* 13 (1): 121–54.

MaCurdy, Thomas E. 1981. "An Empirical Model of Labor Supply in a Life-Cycle Setting." *J.P.E.* 89 (6): 1059–85.

Madrian, Brigitte C., and Lars John Lefgren. 2000. "An Approach to Longitudinally Matching Current Population Survey (CPS) Respondents." *J. Econ. and Soc. Measurement* 26 (1): 31–62.

Manuelli, Rodolfo E., Ananth Seshadri, and Yongseok Shin. 2012. "Lifetime Labor Supply and Human Capital Investment." Working paper.
Mincer, Jacob. 1974. *Schooling, Experience and Earnings.* New York: NBER Press.
Mincer, Jacob, and Haim Ofek. 1982. "Interrupted Work Careers: Depreciation and Restoration of Human Capital." *J. Human Resources* 17 (1): 3–24.
Mincer, Jacob, and Soloman Polachek. 1974. "Family Investments in Human Capital: Earnings of Women." *J.P.E.* 82 (2, pt. 2): S76–S108.
Mulligan, Casey B. 1998. "Substitution over Time: Another Look at Life-Cycle Labor Supply." *NBER Macroeconomics Ann.* 13:75–134.
Neal, Derek. 1999. "The Complexity of Job Mobility among Young Men." *J. Labor Econ.* 17 (2): 237–61.
Prescott, Edward C., Richard Rogerson, and Johanna Wallenius. 2009. "Lifetime Aggregate Labor Supply with Endogenous Workweek Length." *Rev. Econ. Dynamics* 12:23–36.
Rogerson, Richard, and Johanna Wallenius. 2013. "Nonconvexities, Retirement and the Elasticity of Labor Supply." *A.E.R.* 103 (4): 1445–62.
Rosen, Sherwin. 1972. "Learning and Experience in the Labor Market." *J. Human Resources* 7 (3): 362–42.
———. 1976. "A Theory of Life Earnings." *J.P.E.* 84 (4, pt. 2): S45–S67.
Ruhm, Christopher. 1991. "Are Workers Permanently Scarred by Job Displacement?" *A.E.R.* 81 (1): 319–24.
Rupert, Peter, and Giulio Zanella. 2015. "Revisiting Wage, Earnings, and Hours Profiles." *J. Monetary Econ.* 72:114–30.
Rust, John, and Christopher Phelan. 1997. "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets." *Econometrica* 65 (4): 781–831.
Sanders, Carl, and Christopher Taber. 2012. "Life-Cycle Wage Growth and Heterogeneous Human Capital." *Ann. Rev. Econ.* 4 (1): 399–425.
Schmieder, Johannes F., Till von Wachter, and Stefan Bender. 2016. "The Effect of Unemployment Benefits and Nonemployment Durations on Wages." *A.E.R.* 106 (3): 739–77.
Shaw, Kathryn L. 1989. "Life-Cycle Labor Supply with Human Capital Accumulation." *Internat. Econ. Rev.* 30 (2): 431–56.
Taber, Christopher. 2002. "Tax Reform and Human Capital Accumulation: Evidence from an Empirical General Equilibrium Model of Skill Formation." *Advances Econ. Anal. and Policy* 2 (1).
Topel, Robert H., and Michael P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Q.J.E.* 107 (2): 439–79.
Wallenius, Johanna. 2011. "Human Capital Accumulation and the Intertemporal Elasticity of Substitution of Labor: How Large Is the Bias?" *Rev. Econ. Dynamics* 14 (4): 577–91.