Lifetime Labor Supply and Human Capital Investments

Rodolfo E. Manuelli* Ananth Seshadri† Yongseok Shin‡

March 6, 2011

Preliminary and Incomplete

Abstract

We develop a model to study the quantity and the quality of labor supply: People in our model not only choose when to enter and retire from the labor force (quantity of labor), but also make human capital investment decisions in school and on the job (quality of labor). We provide a theoretical analysis of retirement and human capital investment. In our quantitative analysis, we first show that the economic impact of tax and retirement policies is substantial, with the model explaining the majority of the difference in schooling and retirement between Europe and the United States. Second, we find that a reform of tax and retirement policies in Europe will result in a large increase in effective labor supply and output. This effect materializes primarily through the human capital channel, with the career length (quantity of labor) being relatively inelastic to the policy changes. Our findings suggest that models that abstract from human capital investment decisions will underestimate the true economic impact of shocks and policies.

*Washington University in St. Louis and Federal Reserve Bank of St. Louis

†University of Wisconsin–Madison

‡Washington University in St. Louis and Federal Reserve Bank of St. Louis; yshin@wustl.edu.

1 Introduction

Following Prescott's (2005) Nobel lecture, an active debate on microeconomic vs. macroeconomic labor supply elasticities re-emerged. This debate, in our view, has placed too much emphasis on the quantity of labor supply measured by hours or years worked. The real issue underlying the debate is how taxes and retirement policies incur deadweight losses in the economy. In this context, what matters is the impact of these policies on the supply of quality-adjusted labor, i.e., effective labor. One important limitation of existing studies is their assumption that the quantity of labor is converted into effective labor either at par value or according to an exogenously-given learning-by-doing profile.

In the model we study here, individuals not only choose years of work, i.e., career length (quantity of labor), but also the human capital investment in and out of school, i.e., years of schooling, human capital acquisition while in school, and human capital accumulation on the job (quality of labor).

In our theoretical analysis, we characterize the impact of policies and shocks on human capital investment and retirement decisions, and contrast our results with those of the standard models that abstract from human capital decisions.

In our quantitative analysis, we obtain two main results. First, the impact of taxes and retirement policies on effective labor supply is substantial. For example, these policies explain 40 percent of the gap in retirement ages between France and the US, and all of their difference in schooling. Second, a reform of the tax and retirement policies in Europe will have a large positive effect on aggregate output, mostly through the human capital channel. The quantity of labor, i.e., career length, is relatively inelastic to such policy changes: The retirement age increases, but so does the age at which workers leave school and join the labor force. In this sense, we reconcile the views of Prescott (2005), Ljungqvist and Sargent (2006), and Prescott, Rogerson, and Wallenius (2009), who emphasize the large impact of tax policies on aggregate output, with the seemingly-opposing views of Ljungqvist and Sargent (2010), who emphasize that rigidities built into nonemployment benefits may throw worker onto a corner and limit their career length responses to policy changes.

Through our analysis, we find that the retirement and human capital investment decisions reinforce each other, and that the full impact of tax and retirement policies can be captured only in a model that allows individuals to adjust both the quantity and the quality of their labor supply. For example, most studies on the fiscal sustainability of retirement policies abstract from the endogenous response of workers' human capital to policy changes (Gruber and Wise, 2007). In light of our results, we conjecture that these studies may be grossly underestimating the full economic impact of the proposed reform of tax and retirement policies.

The key elements of our model are labor-leisure choices at the extensive margin, human capital investment in and out of school, and retirement policies that introduce rigidities in workers' labor supply decisions. The first two are what Keane and Rogerson (2010) consider to be important for useful models of labor supply, and the last is what Ljungqvist and Sargent (2011) emphasize in their review article.

We now describe in more detail our model, and then summarize our theoretical and quantitative analysis.

Our model incorporates into the standard continuous-time life-cycle model–which is essential to understand retirement decisions–a dynastic preference structure that guarantees a tractable characterization of the long-run equilibrium. We assume that individuals choose years of schooling and human capital investment in school, and that, when in the labor force, they decide how to allocate their time between market work and investment in their human capital (on-the-job training). In our model, leisure, which gives extra utility, is indivisible (i.e., labor supply decisions at the extensive margin only). Retirement is defined as the continuous period of leisure consumption that lasts till the individual's life expires: In other words, retirement is endogenously-chosen back-loaded leisure consumption. Finally, we model retirement benefits as a non-linear function of lifetime income and retirement age that encapsulates the most salient features of real-world policies.

We use the model to explore, both theoretically and quantitatively, the impact of changes in retirement policies, demographic parameters (e.g., population growth rate and life expectancy), and total factor productivity. A given change affects several margins simultaneously. For example, higher taxes discourage individuals from acquiring human capital in and out of school, and, at the same time, induce workers to retire earlier.

In our theoretical analysis, we first show that, in the long run, reductions in the progressivity of taxes and transfers increase workers' retirement age. We then explore the effects of unanticipated shocks. First, we show that negative wealth shocks, i.e., an unanticipated drop in the value of one's financial wealth, increase the retirement age. Second, we find that a permanent decrease in the wage rate (a negative TFP shock) induces older workers to retire earlier, but has an ambiguous impact on younger workers' retirement that depends on the relative strength of income and substitution effects. In addition, for young workers, such a shock flattens their age-earnings profile: Upon impact, their effective labor supply increases as they allocate more time to market work and less to human capital investment, but over time their human capital and hence effective labor supply is lower than what they would have been in the absence of the shock. Finally, we study the effect of an unanticipated drop in the stock of human capital (e.g., a loss associated with reallocation in the presence of firm- or sector-specific human capital). We show that older workers supply less effective labor and retire earlier. Again, this shock has an ambiguous impact on the retirement of younger workers.

For our quantitative analysis, we calibrate the model to the data on output per worker, demographic variables, and tax and retirement policies for the US and several European countries, as well as Japan and Mexico. We show that our model can closely replicate the observed cross-country differences in schooling and retirement.

In our first quantitative exercise, we analyze the economic impact of the other countries' adopting the US retirement benefit system, which, compared to the prevailing regimes in the European countries, is less generous and less redistributive. We find that the increase in output per worker is very large, tantamount to additional 1.5-percent annual growth over 40 years for countries such as Denmark, France, and Spain. This response is driven almost exclusively by increases in human capital (quality of labor), since we find that the retirement age and years of schooling both increase by about two years, leaving career lengths essentially unchanged. This finding of ours contributes to the recent debate on social security reforms in Europe. We find, in agreement with Prescott's (2005) view, that a reform that lowers taxes can have a large positive economic impact on, for example, aggregate output. At the same time, consistent with Ljungqvist and Sargent (2010), our model suggests that labor supply measured in terms of career length may not respond by much. However, the causes of this inelastic response in our model are different from theirs: institutional rigidities and corner solutions in Ljungqvist and Sargent's model vs. the co-movement of schooling and retirement ages and the elastic adjustment along the human capital investment channel in ours.

In the second set of exercises, we use the model to evaluate the economic impact of certain policy and demographic changes in the US. We find that raising the normal retirement age from 65 to 67 increases output per worker by seven percent in the long run. Next, eliminating the built-in progressivity in the benefit function while holding constant the average replacement ratio has a negligible impact. On the other hand, we find that eliminating the social security benefits and taxes altogether has a dramatic impact: Output per worker increases by 45 percent in the long run. As for demographic changes, increasing life expectancy by two years and halving the population growth rate respectively reduce output per worker in the long run by four and seven percent. This is driven by the higher taxes needed to pay for the larger fraction of retirees in the population (i.e., higher dependency ratio). In all these exercises, the impact on career lengths is negligible, and the effects of policy and demographic changes materialize through the human capital channel.

To sum up, we find that human capital investment in school and on the job is an important margin along which people respond to shocks and policies. In fact, even when the macroeconomic labor supply elasticity—measured by career length—is close to zero, tax and retirement policies have substantial effects on the effective labor supply through the human capital channel. Our model thus captures the full impact of shocks and policies, which has hitherto eluded the labor supply debate centered only on the quantity measures of labor supply.

2 Model

We adapt the basic model from Manuelli and Seshadri (2009) to allow for endogenous retirement. Our model incorporates into the standard continuous-time life-cycle model a dynastic preference structure. We assume that individuals choose years of schooling and human capital investment in school, and that, when in the labor force, they decide how to allocate their time between market work and investment in their human capital. Leisure, which gives extra utility, is indivisible, and retirement is endogenously-chosen back-loaded leisure consumption. Finally, we model retirement benefits as a non-linear function of lifetime income and retirement age.

Life Cycle Individual life span is deterministic and runs from age 0 to T. From age 0 till I, an individual is attached to his parent, who makes all decisions for him. He then becomes independent at age I. At age B, he gives birth to e^f children, who will remain attached to him until he turns $B+I$ (and they turn I). We assume $I <= B$.

Choices An individual starts making decisions at age $a = I$. Between age 6 and T, everyone is endowed with one unit of time at any given instant. Time can be spent on leisure, $\ell(a)$, investment in human capital, $n(a)$, and market work, $1-n(a)-\ell(a)$. Leisure is a binary variable: $\ell(a) \in \{0,1\}$. Both $n(a)$ and $1-n(a)-\ell(a)$ must belong to the closed interval $[0,1]$. The individual has to choose consumption, $c(a)$, and goods input to human capital production, $x(a)$, from age I to T.

In addition, he makes decision for his children while they are attached. Using subscript k for children, he chooses their consumption, $c_k(a)$, between their age 0 and I. At their age 6, he chooses goods input to early childhood human capital investment, x_E. Between their age 6 and I, he chooses their time allocation and goods input to human capital: $\{n_k(a), \ell_k(a), x_k(a)\}$. Finally, when his children become independent, he gives each child a "bequest" of q_k.

Preference An independent adult orders allocations in terms of his own consumption and leisure, and also all his descendants' utility. The parent's altruism may be "imperfect" as in Barro and Becker (1989). It is convenient to decompose the individual's utility function into three components:

$$\int_I^T e^{-\rho(a-I)} U(c(a), \ell(a)) da + e^{-\alpha_0+\alpha_1 f} \int_0^I e^{-\rho(a+B-I)} U(c_k(a), \ell_k(a)) da + e^{-\alpha_0+\alpha_1 f} e^{-\rho B} V_k.$$

The first term is utility from his own consumption and leisure, where ρ is the subjective discount rate. The second term is the utility he derives from his children while they are attached, where α_0 and α_1 control the degree of altruism. The case with $\alpha_0 = 0$ and $\alpha_1 = 1$ corresponds to the standard dynastic preference, while positive α_0 and α_1 less than 1 imply imperfect altruism. The term is the continuation value of the children at their independence. This (endogenous) continuation value V_k of each child will be defined below, when we recursively formulate the individual's problem.

We use the following specification for the flow utility:

$$U(c, \ell) = \frac{(c(1+\zeta\ell))^{1-\theta}}{1-\theta},$$

with the parameters being such that:

$$\frac{(c(1+\zeta))^{1-\theta}}{1-\theta} > \frac{c^{1-\theta}}{1-\theta}.$$

This specification can accommodate both the view that consumption drops at retirement because of the substitutability between home and market goods (French, 2005; Laitner and Silverman,

2008) and the view that in order to supply labor to the market it is necessary to purchase some goods such as transportation and meals outside the home (Aguiar and Hurst, 2009). See the appendix for the latter specification.

Human Capital Production We adopt Ben-Porath's (1967) formulation of the human capital production technology, augmenting it with an early childhood period. To be precise:

$$\dot{h}(a) = z_h(n(a)h(a))^{\gamma_1}x(a)^{\gamma_2} - \delta_h h(a), \qquad a \in [6, T] \tag{1}$$

$$h_k(6) = h_B x_E^{\nu}, \qquad 0 < \gamma_i, \nu < 1 \tag{2}$$

Equation (1) corresponds to the standard human capital production developed by Ben-Porath (1967). The technology to produce a child's human capital at age 6, $h_k(6)$ or h_E, is given by (2), where h_B is the stock of human capital at birth. This specification captures the idea that nutrition and health care are important determinants of early levels of human capital, and those inputs are basically market goods.

Schooling and Retirement Given the life cycle, individuals will choose to front-load their human capital production, and $n(a)$ (weakly) decreases with age. In fact, there can be an age interval $[6, 6+s]$ in which individuals are at a corner: i.e., $n(a) = 1$. We interpret this as schooling, with s being the length of schooling. Note that we use the same human capital production function (1) in and out of school.

At the same time, individuals in our model will choose to back-load their leisure consumption, with $\ell(a) = 1$ for $a \in (R, T]$ and $\ell(a) = 0$ for $a \in [6, R]$, with $6 < R \leq T$. We interpret such an R as the retirement age. As we will show in the appendix, back-loaded leisure is the consequence of human capital depreciation and the equilibrium interest rate being greater than or equal to subjective discount rate.[1]

Taxes and Transfers The government taxes labor income at rate τ. In our formulation, $\tau \equiv \tau_I + \tau_S$, where τ_S is the social security tax rate that is not necessarily equal to the statutory rate. Capital income is taxed at rate τ_I only.

Out of the income tax revenue, the government pays for retirement benefits, non-retirement transfers, and non-transfer expenditures. We denote with $b(a, R)$ the retirement benefit flow at age a for an individual whose chosen retirement age is R.

Individuals' Problem We describe different phases of the individual income process net of taxes and goods input to human capital production. In particular, take an individual with $s \geq 0$ years of schooling—i.e., $n(a) = 1$ for $a \in [6, 6+s)$, $0 <= n(a) < 1$ for $a \in [6+s, T]$—and retirement

[1] Age-dependent retirement benefits also induce the back-loading of leisure, although they are not necessary.

age R—i.e., $\ell(a) = 1$ for $a \in (R, T]$ and $\ell(a) = 0$ for $a \in [6, R]$. Let w be the rental rate of human capital. His net (of taxes and goods input) income at age a is as follows.

$$y(a,R) = \begin{cases} 0 & 0 \le a < 6 \\ -x_E & a = 6 \\ -x(a) & 6 \le a < 6+s \\ (1-\tau)[wh(a)(1-n(a)) - px(a)] & 6+s \le a \le R \\ b(a,R) & R < a \le T \end{cases}$$

The interpretation is straightforward. Until age 6, individuals are not endowed with time, and we summarize the expenditures on early childhood human capital investment as a lump-sum payment at age 6, x_E. From age 6 to $6+s$, net income is the negative of the value of market goods used in schooling, $-x(a)$. The active labor market period runs from age $6+s$ to R. During this period, the net income is $(1-\tau)[wh(a)(1-n(a)) - px(a)]$, where p is the tax-adjusted price of goods used for on-the-job training. To be more specific, we allow for the possibility that a fraction ξ of training expenditures are tax-deductible. Thus, given the income tax rate τ, p is given by $p = (1-\tau\xi)/(1-\tau)$. Throughout our analysis, however, we assume $\xi = 0$; i.e., none of the expenditures on goods used to produce human capital on the job is tax-deductible. This assumption implies that the tax treatment of goods input to human capital production is the same in and out of school. Finally, $b(a,R)$ is retirement benefits, which we assume is not taxed.

We assume that individuals can freely borrow and lend at the after-tax real interest rate $r \equiv (1-\tau_I)\hat{r}$.

Now, the budget constraint is given by

$$\begin{aligned} &\int_I^T e^{-r(a-I)} c(a)da + e^f \int_0^I e^{-r(a+B-I)} c_k(a)da + e^f e^{-rB} q_k \\ &\quad \le \int_I^T e^{-r(a-I)} y(a,R)da + e^f \int_6^I e^{-r(a+B-I)} y_k(a,R_k)da + q + u, \end{aligned} \tag{3}$$

where q is the bequest that the individual starts his independent life with, and u is the non-retirement transfer he receives from the government. The variables with subscript k are his children's. Note in particular that the parent not only pays for his children's consumption and goods input but also receives their labor income until they become independent.

Now, the maximized value of an independent individual at age I who starts with h units of human capital, bequest q, and non-retirement transfer from the government u is:

$$\begin{aligned} V(h,q,u) = \max &\int_I^T e^{-\rho(a-I)} U(c(a), \ell(a))da + \\ e^{-\alpha_0+\alpha_1 f} &\int_0^I e^{-\rho(a+B-I)} U(c_k(a), \ell_k(a))da + e^{-\alpha_0+\alpha_1 f} e^{-\rho B} V_k(h_k(I), q_k, u_k). \end{aligned}$$

The maximization is over $\{c(a), x(a), \ell(a), n(a)\}_{a=I}^{T}$, $\{c_k(a)\}_0^I$, $\{x_k(a), \ell_k(a), n_k(a)\}_{a=6}^{I}$, x_E, and q_k, and is subject to the budget constraint (3) and human capital production technology (1, 2).

3 Theoretical Analysis

We first show that the utility maximization can be transformed into an income maximization problem. We then analyze how retirement decisions are made. Next, we characterize how an individual's human capital investment and retirement respond to unexpected permanent shocks and policy changes.

3.1 Solving the Individual Problems

Given our model specification and especially the assumption of perfect financial markets, we can transform the individual's problem into a version of income maximization problems. With our dynastic preference, to obtain this standard separation result, one must not restrict bequest. An intuitive (and heuristic) argument that shows the correspondence between utility maximization and income maximization is as follows. Suppose that parents—who make human capital accumulation decisions for their children until they turn I—do not choose investments in human capital to maximize the present value of their children's income. In this case, the parent could increase the utility of each child by adopting the income-maximizing human capital policy and adjusting the bequest to pay for this change. It follows that the cost to the parent is the same and his children are made better off.

We proceed in two steps. The first is to simplify and solve the individual's problem for a given retirement age R. Here we obtain an indirect value function in R. In the second step, we solve for the retirement age that maximizes the indirect value function.

For a given retirement age R, which we will solve for in the second step, we proceed to study optimal consumption and human capital accumulation. Consumption expenditures before and in retirement are given by:

$$c(a) = c(I)e^{(r-\rho)(a-I)/\theta}, \ a \in [I, R] \tag{4}$$

$$c(a) = c(I)e^{(r-\rho)(a-I)/\theta}(1+\zeta)^{1/\theta-1}, \ a \in (R, T] \tag{5}$$

To guarantee $\lim_{a\downarrow R} c(R) > c(a)$, we assume $\theta > 1$. In addition, the attached children's consumption is

$$c_k(a) = c(I)e^{\frac{(r-\rho)(a+B-I)-\alpha_0-(1-\alpha_1)f}{\theta}}, \ a \in [0, I) \tag{6}$$

It is convenient to simplify the problem by looking at the sum of the utility that directly accrues to the parent between his age I and T and the utility derived from his attached children. Simple calculations using (4,5,6) simplify this sum into the right-hand side.

$$\int_I^T e^{-\rho(a-I)}U(c(a),\ell(a))da + e^{-\alpha_0+\alpha_1 f}\int_0^I e^{-\rho(a+B-I)}U(c_k(a),\ell_k(a))da = \frac{c(I)^{1-\theta}}{1-\theta}G(I,R)$$

Here, we have introduced the following new notations:

$$\begin{aligned}
&G(I,R) \equiv \widetilde{G}(I,R) + e^{\upsilon(r)I}e^{f-rB}e^{-\mu/\theta}\Delta(I)\\
&\widetilde{G}(I,R) \equiv e^{\upsilon(r)I}\left[(\Delta(R)-\Delta(I)) + (1+\zeta)^{\frac{1-\theta}{\theta}}(\Delta(T)-\Delta(R))\right]\\
&\Delta(x) \equiv \int_0^x e^{-\upsilon(r)a}da = \frac{1-e^{-\upsilon(r)x}}{\upsilon(r)}\\
&\upsilon(r) \equiv \frac{\rho-(1-\theta)r}{\theta} > 0\\
&\mu \equiv \alpha_0 + (1-\alpha_1)f + \rho B - rB
\end{aligned}$$

The μ term captures the difference between the current interest rate r and the effective discount rate for the utility of different generations $(\alpha_0 + (1-\alpha_1)f)/B + \rho$.

The consumption expenditures for the parent, inclusive of the consumption of the attached children, similarly simplify into the right-hand side using (4,5,6):

$$\int_I^T e^{-r(a-I)}c(a)da + e^f\int_0^I e^{-r(a+B-I)}c_k(a)da = c(I)G(I,R)$$

With this notation, the utility of a parent of generation t at age I is:

$$V(h_t(I),q_t,u_t;R_t) = \max_{c_t(I),h_{t+1}(I),q_{t+1}} \frac{c_t(I)^{1-\theta}}{1-\theta}G(I,R_t) + e^{f-rB}e^{-\mu}V(h_{t+1}(I),q_{t+1},u_{t+1};R_{t+1}), \tag{7}$$

with subscript t indexing generations. The budget constraint is simply

$$c_t(I)G(I,R_t) + e^{f-rB}q_{t+1} \leq W(I,T,R_t) + e^{f-rB+rI}W(0,I,R_{t+1}) + q_t + u_t \tag{8}$$

where

$$W(a_1,a_2,R) = \max_{x(a),n(a)}\int_{a_1}^{a_2} e^{-r(a-a_1)}y(a,R)da,$$

is the maximized—over the allocation of time and goods inputs—present discounted value of net labor income for given retirement age R.

Note that the income maximization must also take into account the retirement benefits $b(a,R)$, which may depend on labor income history. To maintain the correspondence between the utility and the income maximization, we assume that the benefits are affine in the present discounted value of net labor income. In particular, through out our analysis, the benefits are assumed to be:

$$b(a,R) = \begin{cases} 0 & a < \max\{R,R_n\} \\ b_m + b_y W(0,R,R) & a \geq \max\{R,R_n\} \end{cases}$$

Here, b_m is a constant and $W(0,R,R)$ is the maximized present discounted value of net labor income. In addition, R_n is the normal retirement age set by law. Those who retire before R_n will have to wait until they turn R_n to receive benefits. On the other hand, those who retire past R_n forfeit the benefit between R_n and their retirement without actuarial compensation.

To determine the level of human capital for a given retirement age, it now suffices to maximize the present discounted value of net income. Note that the optimal choice of human capital is independent of the discount rate: Individuals maximize the present discounted value of net labor income and hence the relevant discount rate is the interest rate.

It is trivial to show that an increases in R—holding others constant—results in higher levels of schooling and human capital. The intuition for this result is simple: As the rate of utilization of human capital increases, its return becomes higher and individuals respond by increasing their human capital investment. The details of the optimal human capital choices are in the appendix.

We now solve for R. The individual's problem (7,8) can be re-written into a sequence problem, with t indexing generations. The objective function is:

$$\sum_{t=0}^{\infty} e^{(f-rB)t} e^{-\mu t} \frac{c_t(I)^{1-\theta}}{1-\theta} G(I,R_t), \tag{9}$$

and the budget constraint is obtained by forward iteration:

$$\sum_{t=0}^{\infty} e^{(f-rB)t} c_t(I) G(I,R_t) \leq \\ \sum_{t=0}^{\infty} e^{(f-rB)t} [W(I,T,R_t) + e^{f-rB+rI} W(0,I,R_{t+1})] + \sum_{t=0}^{\infty} e^{(f-rB)t} u_t + q_0. \tag{10}$$

Assuming that $W(\cdot,\cdot,R)$ is differentiable with respect to R, although it is not everywhere because of the policy-induced kinks in the retirement benefits, the first-order conditions are:

$$e^{-\mu t} c_t(I)^{-\theta} = \Phi,$$

$$e^{-\mu t} \frac{c_t(I)^{1-\theta}}{1-\theta} G_R(I,R_t) - \Phi c_t(I) G_R(I,R_t) = -\Phi W_R(I,T,R_t),$$

where Φ is the Lagrange multiplier on (10), and G_R and W_R denote the partial derivative with respect to R. The conditions further simplify into:

$$W_R(I,T,R_t) = \frac{\theta}{\theta-1} G_R(I,R_t) c_t(I). \tag{11}$$

The second-order conditions require that the left-hand side cuts the right-hand side from above.

Since the retirement benefits are discontinuous at R_n, there are two cases to consider. For $R \geq R_n$, with the notation $D(R) = (e^{-rR} - e^{-rT})/r$,

$$W_R(I,T,R_t) = e^{-r(R_t-I)} \left[(1 + b_y D(R_t))(1-\tau) w h(R_t) - \left(b_m + b_y \int_0^{R_t} e^{-ra} y(a,R_t) da \right) \right],$$

while for $R < R_n$ we obtain

$$W_R(I,T,R_t) = e^{-r(R_t-I)} (1 + b_y D(R_n))(1-\tau) w h(R_t).$$

3.2 Properties of the Solution

In this section we characterize some properties of the solution to the individual problem. In particular, we describe the response of effective labor supply to unanticipated permanent changes in the economic environment.

Throughout our analysis, we restrict our attention to the implications of the model when $\mu = 0$, which corresponds to the steady state. To see this, consider the case in which the interest rate exceeds the generational discount rate, i.e., $\mu < 0$. One consequence is that consumption increases from generation to generation, shifting the right-hand side of (11) upward. The second-order condition implies that the retirement age, R_t, must decrease over generations indexed by t.[2] Intuitively, this is driven by income effects: Earlier retirement allows individuals to enjoy more leisure. The case with $\mu > 0$ can be worked out in a straightforward manner.

3.2.1 Steady State

In the steady state ($\mu = 0$), from the budget constraint 8, we obtain

$$c(I)\widetilde{G}(I,R) = W(I,T,R) + e^{f-rB+rI}W(0,I,R) + (1 - e^{f-rB})q + u.$$

Now, the relevant first-order condition for the choice of retirement age is (except at the kink):

$$W_R(I,T,R) = \frac{\theta}{\theta - 1}\frac{\widetilde{G}_R(I,R)}{\widetilde{G}(I,R)}[W(I,T,R) + e^{f-rB+rI}W(0,I,R) + (1 - e^{f-rB})q + u]. \qquad (12)$$

Hereafter, we denote by $\mathcal{G}(R)$ the right-hand side of equation (12). Since $\widetilde{G}_R(I,R) > 0$ and the function $W_R(I,T,R)$ must intersect $\mathcal{G}(R)$ from above, (12) implies that the marginal benefit of working exceeds the marginal monetary cost at the equilibrium retirement age—if they were equal, the condition would be $W_R(I,T,R) = 0$). The reason is obvious: The opportunity cost of working is the additional utility associated with leisure, which is captured by $\mathcal{G}(R)$. In other words, the first-order condition (12) equates the marginal benefit of additional work (more income) and the marginal cost of forgone utility from leisure converted into units of income.

If $R > R_n$, the forgone retirement benefits are subtracted from the marginal benefit of additional work, and (12) becomes:

$$e^{-r(R-I)}\left[(1 + b_y D(R))(1-\tau)wh(R) - \left(b_m + b_y\int_0^R e^{-ra}y(a,R)da\right)\right] = \mathcal{G}(R) \qquad (13)$$

3.2.2 Long-Run Responses

Changing the Progressivity of Retirement Benefits The above first-order conditions can be used to study how the level and progressivity of the retirement system affect the retirement

[2] This model delivers the result that there is a downward trend in retirement age when interest rates exceed their long-run values—which happens during the transition to the steady state.

age. An increase in the constant portion b_m—holding other things equal—decreases the optimal retirement age, because it shifts down the left-hand side of equation (13) and moves $\mathcal{G}(R)$ up.

To isolate the impact of the progressivity of the retirement benefit regime, we consider a change in (b_m, b_y) that keeps the benefit level constant at the initial retirement age R. Let the benefit level be

$$\bar{b} = b_m + b_y \int_0^R e^{-ra} y(a, R) da.$$

The condition (13) is now

$$e^{-r(R-I)} \left[(1 + b_y D(R))(1 - \tau) w h(R) - \bar{b} \right] = \mathcal{G}(R),$$

and it follows that an increase in b_y accompanied by a decrease in b_m to maintain $\bar{b}$ shifts the left-hand side upward and leaves the right-hand side unchanged. As a result, the retirement age increases. The intuition is clear: Returns to work increase as the system reduces the pure tax-and-redistribute component of the retirement benefits. We conclude that it is not only the level of benefits but also their progressivity that influences retirement.

Changing the Composition of Government Expenditures A change in the composition of government spending has an impact on retirement. Assume that the government, keeping tax rates constant, raises (non-transfer) government consumption and commensurately reduces non-retirement transfers u. In this case, the right-hand side of (12) shifts down and the left-hand side remains unchanged, resulting in an increase in the retirement age.

3.2.3 Short-Run Responses

We now analyze the response of effective labor supply to three types of unanticipated shocks: a one-time drop in the value of accumulated assets, a one-time drop in the stock of human capital, and a permanent decrease in the rental rate of human capital.

We find it convenient to assume that such shocks hit a worker of age $a' > I + B$. That is, his children have already left, and he will not adjust the bequest in response to shocks. In this case, the response captures the standard life-cycle effects.[3] In addition, for simplicity, we consider the case in which individuals choose $R < R_n$ in the absence of shocks, and remain in this range after the shock.[4]

[3]The results are qualitatively similar for ages under $I + B$, but the expressions are algebraically more cumbersome and income effects must include the adjustment of bequests.

[4]If this is not the case, again, the qualitative results remain similar but the expressions need to be adjusted.

Negative Wealth Shock The relevant version of (12) at age a' for a worker who chooses to retire before R_n is

$$W_R(a',T,R) = \frac{\theta}{\theta-1}\frac{\widetilde{G}_R(a',R)}{\widetilde{G}(a',R)}[(1+b_y D(R_n))W(a',T,R)$$
$$+e^{ra'}D(R_n)(b_m + b_y W(0,a',\bar{R})) + A(a')], \tag{14}$$

where $A(a')$ is the value of assets accumulated up until age $a' < R$, and $\bar{R}$ is the pre-shock retirement age choice. As before, the second-order condition requires that the left-hand side of (14) crosses the right-hand side from above.

Consider the effect of an unanticipated drop in the value of financial wealth $A(a')$. This shock shifts the right-hand side of (14) down. Since $W_R(a',T,R)$ is not affected by the shock, the retirement age increases: i.e., $\partial R/\partial A < 0$.

In the appendix we show that $\partial^2 R/\partial A \partial a' > 0$ when evaluated at the pre-shock level of assets. Thus, when faced with an unexpected loss of financial wealth, older workers raise their planned retirement age by less than do younger ones.

Permanent Negative Wage Shock What is the effect of an unanticipated permanent decrease in the rental rate of human capital, w, for a worker a' years old? Abstracting from taxes and retirement benefits here, (27) in the appendix implies that

$$\frac{\partial h^e(a)}{\partial w} = \frac{\gamma_2}{1-\gamma}\frac{C_h}{w}\left(\frac{w}{p}\right)^{\frac{\gamma_2}{1-\gamma}}\pi(a',a),$$

where $h^e(a)$ is the supply of effective labor—defined as $h^e(a) \equiv h(a)(1-n(a))$—by a worker of age $a \geq a'$, and

$$\pi(a',a) \equiv (r+\delta_h)\int_{a'}^{a} e^{-\delta_h(a-u)}m(u)^{\frac{\gamma}{1-\gamma}}du - \gamma_1 m(a)^{\frac{1}{1-\gamma}},$$
$$m(a) \equiv 1 - e^{-(r+\delta_h)(R-a)},$$
$$C_h \equiv \left(\frac{z_h}{r+\delta_h}\gamma_1^{\gamma_1}\gamma_2^{\gamma_2}\right)^{\frac{1}{1-\gamma}}.$$

The sign of the impact on effective labor depends on the sign of $\pi(a',a)$. To understand the impact on retirement, let us first fix R and $a' < R$. Simple calculations show that for each a' there exists $\varphi(a') \in (a',R)$ such that

$$\pi(a',a) = \begin{cases} < 0 & \text{if } a' \leq a < \varphi(a') \\ = 0 & \text{if } a = \varphi(a') \\ > 0 & \text{if } \varphi(a') < a \leq R \end{cases}$$

This result shows that, immediately following a permanent decrease in the wage rate, effective labor supply increases. The size of the increase depends on the worker's age: young workers increase their effective labor supply more than old workers.[5]

Over time, the effect of the change in the human capital rental rate reverses itself. For a given age a' at the time of the shock, there exists another age, $\varphi(a') > a'$, such that the post-shock effective labor supply at that age coincides with what the effective labor supply at that age would have been in the absence of the shock. Finally, for $a > \varphi(a')$, the post-shock effective supply of labor, $h^e(a)$, drops below what it would have been without the shock.

The intuition for this result is as follows. When faced with a permanent decrease in the return to human capital, workers choose to reduce investment in human capital. In the short run, this results in more human capital supplied to the market as he changes his allocation of time away from human capital production. Eventually, the lower levels of investment in human capital is such that the stock of human capital drops below what it would have been without the shock, bringing down effective labor supply. As shown in the appendix, the qualitative response of earnings mirrors that of effective labor. For a fixed retirement age, the endogenous adjustment in human capital investment implies that a drop in the human capital rental rate flattens the age-earnings profile.

What is the impact on the retirement age? In general, it is not possible to determine the effect on retirement because there are two forces at play. While a lower w reduces the marginal benefit of working, $W_R(a', T, R)$, the negative income effect shifts down the right-hand side of (14), and they push the retirement age in opposite directions. However, it is possible to show that for a' close to R (i.e., for older workers), the magnitude of the right-hand side's shift is arbitrarily small. Thus, for older workers, the model predicts that a decrease in w results in earlier retirement. The effect on the retirement plan of younger workers is ambiguous.

Negative Shock to Human Capital Finally, consider the effect of an unanticipated exogenous drop in the stock of human capital at age a'. This could be caused by , for example, reallocation of workers across sectors or firms when human capital is at least partly sector- or firm-specific.

The effect of this one-time shock on effective labor supply diminishes over time. To be precise, for a fixed retirement age, the effect of an exogenous change in $h(a')$, is given by

$$\frac{\partial h^e(a)}{\partial h(a')} = e^{-\delta_h(a-a')}, \quad a' \leq a < R.$$

Unlike the case of the wage shock, the younger the worker is, the smaller is the long-run impact of the human capital loss, for the given retirement age. The intuition is clear. A one-time shock to the stock of human capital does not alter the returns to human capital in the future. Since younger workers have a longer horizon, they accumulate more human capital than older workers.

As in the case of the wage shock, income and substitution effects push the retirement age in opposite directions. A sufficient condition for the substitution effect to dominate, and hence, for

[5] Trivially, if $a' = R$ there is no change in effective labor supply.

the worker to retire earlier than planned is

$$\frac{r+\delta_h}{e^{(r+\delta_h)(R-a')}-1} \geq \frac{\theta}{\theta-1}\frac{\widetilde{G}_R(a',R)}{\widetilde{G}(a',R)}.$$

It is easy to check that this condition is satisfied for olderk workers (i.e., a' cose to R).

To summarize, we find that the impact of shocks on retirement depends on the standard tension between income and substitution effects as well as on the specifics of the determinants of the age-earnings profile. Unlike in the majority of models of lifetime labor supply, the age-earnings profile in our model is endogenously determined and hence is not independent of the nature of shocks. In particular, in the case of wage shocks, our age-earnings profile responds in a non-monotonic fashion.

4 Quantitative Analysis

For our quantitative analysis, we first close the economy and define the equilibrium of the model. The preference and technology parameters are calibrated to the relevant data for the US.

We then ask whether the differences in tax and retirement policies across countries can explain their differences in schooling and retirement behavior.

We consider three counterfactual exercises. First, we impose the retirement policies of the US onto selected European countries. Second, we hit the US with a series of policy and demographic shocks. Finally, we ask whether our model can explain the labor market outcomes in the US circa 1900. In all three exercises, our focus is on the impact on schooling, retirement, and overall effective labor supply along the life cycle.

4.1 Equilibrium

We close the model by specifying the goods production side of the economy and the government budget constraint. To define the steady-state equilibrium, we develop notations for the age distribution in the population. Let $\eta = f/B$ be the growth rate of population. The fraction of the population of age a is

$$\phi(a) = \frac{\eta e^{-\eta a}}{1-e^{-\eta T}}.$$

We assume that the aggregate goods production function is Cobb-Douglas in physical capital and effective labor, with total factor productivity z. Let κ be the ratio of physical capital to effective labor, and δ_k be the depreciation rate of physical capital. Using this notation, output per worker is

$$y = zF(\kappa,1)\frac{\int_{6+s}^{R} h^e(a)\phi(a)da}{\int_{6+s}^{R}\phi(a)da}, \tag{15}$$

and output per capita is

$$y_C = zF(\kappa, 1) \int_{6+s}^{R} h^e(a)\phi(a)da. \tag{16}$$

The steady-state equilibrium requires

$$\int_0^T [c(a) + x(a)]\phi(a)da + g = [zF(\kappa, 1) - (\delta_k + \eta)\kappa] \int_{6+s}^{R} h^e(a)\phi(a)da, \tag{17}$$

where g is per-capita government consumption. This is the goods market clearing condition in per-capita terms.

We assume competitive factor markets, and hence firms equate the marginal product to factor prices:

$$r = (1 - \tau_I)(zF_K(\kappa, 1) - \delta_k), \tag{18}$$

$$w = zF_H(\kappa, 1), \tag{19}$$

where F_K and F_H denote respectively the marginal product of physical capital and human capital.

We assume that the retirement benefits are fully financed by the social security taxes on labor income, with τ_S-revenues being equal to retirement benefit payments. Furthermore, we assume that the government runs a balanced budget in terms of non-retirement revenues and expenditures:

$$g + u = \tau_I[zF(\kappa, 1) - \delta_k\kappa] \int_{6+s}^{R} h^e(a)\phi(a)da.$$

The steady-state competitive equilibrium can be defined in the standard way, and is omitted here.

4.2 Calibration

We calibrate the model to match the relevant data of the US. Although it is a joint calibration of all parameters, we first discuss the determination of preference and technology parameters, followed by that of the policy parameters.

Preference and Technology Parameters We pick $\theta = 1.07$ to match the average effective retirement age (64.6) in the US in 2000. Given this θ, $(1 + \zeta)^{1/\theta - 1} = 0.8$ generates a 20-percent drop in consumption at retirement, as can be seen in equations (4) and (5).

Directly from the data, we set the life span $T = 77.7$ and the population growth rate $\eta = 0.01$, with the age of giving birth B chosen to be 25.

We set the discount rate $\rho = 0.04$, and choose $\alpha_0 = 0.24$ and $\alpha_1 = 0.65$ following Manuelli and Seshadri (2009).[6] From these, we obtain a steady-state after-tax interest rate r of 5.31 percent per year.

[6] Manuelli and Seshadri pick these values for α_0 and α_1 to match the fertility rate, which is an endogenous variable in their paper, and lifetime intergenerational transfers as a percent of GDP (4.5 percent).

We choose $\alpha = 0.33$ to be the capital share parameter of the Cobb-Douglas production function F. For the model to match a capital-output ratio of 2.5, the depreciation rate of physical capital δ_k is 0.071.

For human capital production, we set $\nu = 0.55$, $\gamma_1 = 0.63$, $\gamma_2 = 0.30$, $z_h = 0.35$, and $\delta_h = 0.018$. There is a moment that is most informative about each of the five parameters: in the same order, pre-primary expenditure per pupil relative to GDP per capita (0.14); earnings at age 50 relative to earnings at 25 (2.17); schooling expenditures as a percent of GDP (3.77 percent); average years of schooling among adult males (12.4); earnings at the average retirement age of 64.6 relative to earnings at 55 (0.8).[7]

Our theory implies that, for h_B, z_h, and w, it is only the ratio $h_B^{1-\gamma} z_h^{\nu-1} w^{\nu(1-\gamma_1)-\gamma_2}$ that matters for the moments of interest. We choose $w = 0.38$ and $h_B = 6.3$ arbitrarily, and calibrate z_h as above. Simple derivations show

$$w = z^{1/(1-\alpha)}(1-\alpha)\left(\frac{\alpha(1-\tau_I)}{r+\delta_k(1-\tau_I)}\right)^{\alpha/(1-\alpha)},$$

which implies that $z = 0.50$, given r, δ_k, τ_I, and w.

Retirement Benefits and Taxes To calibrate b_m and b_y, we use the OECD data on replacement rates at different levels of labor income. We consider individuals who earn 0.5 and 1.5 times the average labor income. Let $\bar{\rho}_j, j \in \{0.5, 1.5\}$ be the replacement rates in the data. Given our formulation of the benefit function, the following should hold:

$$b_m + b_y \times 0.5 \times Y(R) = \bar{\rho}_{0.5} \times 0.5 \times (1-\alpha)y,$$

$$b_m + b_y \times 1.5 \times Y(R) = \bar{\rho}_{1.5} \times 1.5 \times (1-\alpha)y,$$

where y on the right-hand side is the output per worker shown in (15) and $1-\alpha$ is the labor income share. The left-hand sides are the retirement benefit flows that a hypothetical retiree would receive if his labor income were respectively 0.5 and 1.5 times the average labor income. We solve for b_m and b_y that satisfy the two equalities. In the data $\bar{\rho}_{0.5} = 0.5033$ and $\rho_{\bar{1}.5} = 0.3411$, and we need $b_m = 35$ and $b_y = 0.093$.

Given our assumption of the balanced-budget retirement regime, we choose the social security tax rate τ_S so that we have just enough revenues to pay for the benefits. We end up with $\tau_S = 0.076$.

We use the estimated total tax wedge τ also reported by OECD. For the US, $\tau = 0.301$, which implies $\tau_I \equiv \tau - \tau_S = 0.225$.

[7] The data sources for these five moments are in Manuelli and Seshadri (2009, p. 787).

4.3 Results

4.3.1 Schooling and Retirement across Countries

To evaluate the explanatory power of our model, we ask whether our model can explain the differences in schooling and retirement across selected OECD countries. In addition to the US, we consider the data from Denmark (DNK), France (FRA), Japan (JPN), Spain (ESP), and Mexico (MEX).[8]

We assume that all preference and technology parameters are the same across countries, except for the TFP of the aggregate production function z.

It is only the demographic and policy variables that are country-specific in our model. To be more specific, firstly, countries have different life spans (T) and population growth rates (η) as in the data. Secondly, we use the OECD data on normal retirement ages (R_n) and replacement ratios ($\bar{\rho}$), with the latter pinning down country-specific b_m and b_y. The retirement benefits, via budget balance, determine τ_S. Combined with the OECD estimates on total tax wedge τ, we also obtain τ_I for each country. We use non-transfer government expenditure per capita (g) from the Penn World Tables Version 6.3 data. Finally, we calibrate the TFP (z) of each country to its output per worker relative to the US. In the top panel of table 1 we report the country-specific moments. In the bottom panel, we also report the average years of schooling among adult males and the average effective retirement age in the OECD data. Note that these two moments are not used for country-specific calibration.

2000 Data	USA	DNK	FRA	JPN	ESP	MEX
T	77.7	76.5	78.8	80.7	81.1	75.7
η	0.0097	0.0021	0.0049	0.0000	0.0077	0.0112
R_n	65	65	60	63	65	65
$\bar{\rho}_{0.5}$	0.5033	1.2397	0.6175	0.4713	0.8118	0.5527
$\bar{\rho}_{1.5}$	0.3411	0.6747	0.4846	0.2945	0.8118	0.3452
τ	0.3010	0.4120	0.4930	0.2950	0.3780	0.1510
G/Y	0.0834	0.1729	0.1652	0.1325	0.1228	0.1282
Output per worker	1.0000	0.8173	0.7950	0.6940	0.7213	0.3427
Schooling (years)	12.40	10.20	10.43	11.40	10.35	9.20
Retirement age	64.6	63.5	58.7	69.5	61.4	73.0

Table 1: Differences across Countries

For each country, we compute the steady-state equilibrium, and ask how our model predictions on schooling and retirement stand up to the data. In table 2 we report the result, with the OECD data in parentheses.

[8]The data for Germany is quite comparable to Denmark and France. For this reason, the results on Germany are not reported.

	DNK	FRA	JPN	ESP	MEX
Schooling	10.95 (10.20)	10.03 (10.43)	11.20 (11.40)	10.88 (10.35)	9.32 (9.20)
Retirement age	62.74 (63.5)	60.00 (58.7)	63.94 (69.5)	64.78 (61.4)	69.98 (73.0)

Table 2: Model Performance against Data in Parentheses

The model does reasonably well in replicating the patterns in schooling and retirement across countries. For Denmark and France, their higher tax rates (τ) and more generous and redistributive retirement benefits discourage human capital investment (i.e., less schooling) and encourage earlier retirement. In the French case, the retirement age in the model is 60, exactly where the policy-induced kink (R_n) is. For other countries, retirement in the model occurs close to, but not exactly at, the normal retirement age.

However, the model clearly misses the high retirement age in Japan, suggesting that it is not rich enough to account for the retirement behavior of Japanese workers.

More interesting, even though the model under-predicts the retirement age in Mexico, it shows that there is no inconsistency between a lower life expectancy and a higher retirement age: Comparing Mexico and the US, Mexican workers have lower life expectancy, but retires five years later in the model. To understand this, the important difference for this phenomenon is u, the non-retirement transfer: While the total tax wedge τ is twice as high in the US, non-transfer government expenditure relative to GDP is higher in Mexico, implying that the lump-sum redistribution is much smaller in Mexico. As shown in section 3.2.2, this explains why Mexican workers may retire later. Although the lower life expectancy and the substitution effect from the lower (after-tax) wage push retirement in the opposite direction, the income effect from the smaller transfer is found to prevail. We note that one caveat with Mexico is the prevalence of the informal sector in its economy, which raises an issue of coverage of the retirement benefit system.[9]

4.3.2 Imposing the US Retirement Policy

In this section, we evaluate the long-run effects of imposing a US-style retirement benefit system in each of the other five countries. The results are reported in table 3. We hold all other parameters constant, and re-calibrate b_m, b_y, τ_S, τ, and R_n. In particular, we choose b_m and b_y for each country to match the US levels of replacement rates ($\bar{\rho}_{0.5} = 0.5033$ and $\bar{\rho}_{1.5} = 0.3411$). We maintain the assumption that the retirement benefit system pays for itself with τ_S. Thus, we obtain a new τ_S for each country, as shown in the first row of table 3. We hold constant τ_I, which means that τ will have to change by as much as τ_S changes for each country.

We compute the resulting new steady states, and report in table 3 the impact of this policy

[9] While the retirement benefits seem quite similar to the US counterparts, the social security tax rate in Mexico necessary to pay for the benefits is only 0.03, while it is 0.076 in the US. This is because the benefit recipients, those aged $\max R, R_n$ to T, are a much smaller fraction of the population in Mexico.

change on the years of schooling (s), retirement age (R), career lengths, output per worker (y), and average effective human capital per schooling year ($\bar{h}_e/s$). Note that the before and after numbers are all from model predictions: In section 4.3.2 the model does not perfectly replicate the schooling and retirement data.

	DNK	FRA	JPN	ESP	MEX
τ_S (percent)	18.5 → 8.2	18.5 → 9.8	12.0 → 12.3	21.0 → 9.8	3.0 → 3.0
Schooling	10.95 → 13.84	10.03 → 12.98	11.20 → 11.44	10.88 → 13.55	9.32 → 9.82
Retirement age	62.74 → 64.89	60.01 → 62.45	63.94 → 65.24	64.78 → 65.36	69.98 → 70.30
Career length	45.79 → 45.05	43.98 → 43.47	46.74 → 47.80	47.90 → 45.81	54.66 → 54.48
Output per worker	0.82 → 1.52	0.80 → 1.49	0.69 → 0.72	0.72 → 1.30	0.34 → 0.37
$\bar{h}_e/s$	0.84 → 1.25	0.78 → 1.13	0.85 → 0.86	0.82 → 1.17	0.69 → 0.72

Table 3: Long-Run Impact of a US-Style Reform

In the European countries with generous retirement benefits and high taxes, the US-style reform brings down τ_S and hence τ by about ten percentage points. This is a significant cut in the total labor wedge, which increases schooling by two or three years and delays retirement by one or two years. The model also shows that the lower taxes in the long run increase output per worker dramatically. To put this number in perspective, note that convergence to a steady state in this model takes approximately forty years owing to the demographic structure: The impact of the social security reform in Denmark, France, and Spain is equivalent to an additional 1.5-percent growth per year in output per worker over the forty years.[10]

By contrast, a switch to the US-style retirement system has hardly any effects for Japan and Mexico, since their retirement benefit systems are, in some sense, similar to the US system already.

We now ask the relative importance of the two margins of labor supply, i.e., quantity vs. quality, for the large response of output per worker. It is obvious from table 3 that the model predicts only negligible changes in career length or the quantity of labor. This is because the increase in retirement age matches (or, if anything, exceeds) the increase in years of schooling, leaving career lengths unchanged. The big impact on output per worker comes from the changes in human capital or the quality of labor. In addition to the increase in years of schooling, the effective human capital per schooling year in these countries rise from a level that is about 15 percent below the US level to a level that is more than 15 percent higher than the US level. In other words, the individuals not only go to school for more years, but also accumulates more human capital in and out of school. In the model, it is through the human capital investment margin that the large impact of the policy reform materializes. Therefore, if one focuses only on the quantity of labor supply, the economic

[10]The model also predicts that the new steady-state output per worker in these countries will be even higher than that of the US. This is driven by our TFP estimates. Since these economies have much higher tax rates than the US to begin with, the model requires that their levels of TFP be higher in order to match the observed levels of output per worker. When the taxes are reduced, their output per worker exceeds the US level, reflecting the underlying TFP differences.

	Retirement	Schooling	$\bar{h}^e/s$	Output per worker
Baseline	64.60	12.40	1.00	1.00
$T = 77.7 \rightarrow 79.7$	64.99	12.26	0.98	0.96
$\eta \div 2$	63.39	12.04	0.95	0.93
$R_n = 65 \rightarrow 67$	64.88	12.73	1.05	1.07
$\bar{\rho}_{0.5} = \bar{\rho}_{1.5}$	64.92	12.48	1.01	1.01
No social security	65.55	14.10	1.30	1.47
$\Delta\tau = -0.01$	64.78	12.66	1.04	1.06

Table 4: Long-Run Impact of Policy/Demographic Changes

impact of the reform will be grossly underestimated. Furthermore, if one proxies human capital with years of schooling, the impact of the reform on human capital (and eventually on output per worker) will be substantially underestimated.

4.3.3 Policy and Demographic Changes

In this section we consider six experiments on the benchmark US economy. These experiments do have implications for policy (e.g., the exercise of raising the normal retirement age from 65 to 67, which is being gradually rolled out in the US), but our main purpose is to further clarify the inner workings of the model by analyzing the effect of one change at a time.

As for demographic changes, we consider in turn the long-run impact of an increase in the life span (T) and the halving of the population growth rate (η). Firstly, we raise T from 77.7 to 79.7. We re-calibrate b_m and b_y to maintain the replacement rate targets, which causes τ_S and hence τ to adjust by the same magnitude. The results are reported in table 4. We find that schooling and retirement age moves in the opposite directions, albeit not by much. Holding other things equal, an increase in life span necessitates more supply of effective labor, raising the retirement age *and* years of schooling. However, the demographic structure now has a larger fraction of benefit-eligible retirees, which requires higher social security taxes (τ_S) and hence τ. This discourages investment in human capital, as partly reflected on the drop in years of schooling, and drives the decline in output per worker. The exercise of halving the population growth rate is also explained by this very mechanism. Again, the output per worker is driven by the quality of labor supply, not the quantity. In particular, in the case of the life span exercise, the wrong conclusion would be reached if one were to focus only on career lengths as a measure of labor supply.

We now consider changes in tax and retirement policies. First, we raise the normal retirement age by two years, while maintaining the replacement rate targets. Now that the benefit-eligible retirees are a smaller fraction of the population, τ_S can be lowered, also bringing down τ. In the long run, while there is hardly any change in career length, human capital investment in and out of school increases, and output per worker rises by seven percent.

Next, we set $b_m = 0$, thereby eliminating the redistributive component of the retirement benefits. We choose the new replacement rate and b_y such that τ_S and hence τ remain unchanged. Somewhat surprisingly, this change has virtually no effect on output or any other variable.

On the other hand, the elimination of the social security system ($b_m = b_y = \tau_S = 0$) has a dramatic long-run impact on output per worker, which is predicted to increase by 47 percent.[11] This is attained with a small *reduction* in career length and a 13-percent increase in years of schooling. The most important factor is the 30-percent increase in the average effective human capital per schooling year.

Finally, we cut the total tax wedge τ by one percentage point, with commensurate reduction in non-retirement transfer (u). In the long run, output per worker rise by six percent, again driven mostly by the increase in human capital investment.

4.3.4 The US in 1900

As a final exercise, we ask whether our model can explain the trends in schooling and retirement over time. In 1900, more than 60 percent of men aged 65 or older were in the labor force, while the figure in 2000 is 20 percent. At the same time, the average schooling among adult males in the US in 1900 was 5.4 years, while it has grown to 12.4 years by 2000. Given our exercises above, where retirement ages and years of schooling tend to move in the same direction, it is not clear how one can generate a downward trend in labor force participation of older workers and a upward trend in their educational attainment.

We cast our model in the year 1900 as follows. We note that the life span was shorter in 1900. We use the life expectancy at age six, which is age 64.3 in 1900.[12] The population growth rate is 0.02. Consistent with the data, we set $\tau = \tau_I = 0$, $g = u = 0$, and $b_m = b_y = \tau_S = 0$. That is, there is no tax or government expenditure, let alone the social security system. We then choose z for 1900 to match the output per worker relative to the 2000 level, which is 0.19.

With these changes, we compute the steady state of our economy as of 1900, and report the predicted years of schooling and retirement age. The bottom panel of table 5. From the data, the average effective retirement age in 1900 is 64.3, which is the end of one's life. That is, there is no retirement. Our model also predicts that workers do no retire. Remarkably, our model also predicts exactly 5.4 years of schooling, which is the empirical counterpart. There are several force at play. First, the shorter life span, holding other things constant, would imply less schooling and earlier retirement. The lower TFP (z) has both income and substitution effects, with the low returns

[11] Given the speed of convergence in this model, this translates into an additional one-percent per year growth over 40 years.

[12] The life expectancy at birth is merely 52 years. This large discrepancy is driven by the high infant mortality rate in 1900. For our model, we find the expectancy at age six to be more relevant. In the above analysis using the 2000 data, there is virtually no difference between the two measures of life expectancy, because of the near-zero infant mortality rates in the countries we consider.

Moment	Data	Model
T (Life expectancy at six)	64.3	64.3
η	0.02	0.02
Output per worker	0.19	0.19
Schooling (years)	5.4	5.4
Retirement age	64.3 (T)	64.3 (T)

Table 5: The US in 1900

to human capital investment primarily responsible for the drop in schooling and human capital investment. The absence of retirement benefits and non-retirement transfers have a large negative income effect that pushes for later retirement. To understand trends in human capital investment and retirement, we conclude, it is necessary to jointly consider the changes in all relevant factors, e.g., demographic variables, technology, and tax/transfer policies.

5 Future Work

For future research we leave the introduction of heterogeneity and a non-negativity restriction on bequest.

We view heterogeneity as being driven by differences in ability to learn, z_h, which can be calibrated to the distribution of schooling years in the US. We will explain the observed relationship between education and retirement behavior. More important, with heterogeneity, we can have a meaningful discussion about the welfare consequences of tax and retirement policies.

The non-negative bequest restriction is not binding in our exercises. However, we expect that it may bind when we add heterogeneity and also when we study transitions, since policy changes that increase future output may cause parents to borrow from their richer children.

Appendix

In this appendix we present some more detailed results about the equilibrium choice of human capital and a description of preferences that result in the same reduced form utility that we are using

Equilibrium Choice of Schooling and Human Capital In this section we characterize the functions $W_j(R)$. Since the choice of investments in human capital for a given retirement age is not distorted, it is possible to characterize the objects that we are interested in —$W_j(R)$ functions— by using the solution to a standard income maximization problem.

It is straightforward to show that the before tax net present discounted value of income of an a year old individual who is in the labor force, has human capital h, and who plans to retire at age R, $V(h,a,R)$ is given by

$$\tilde{W}(h,a,R) = (1-\tau)w\left[\frac{m(a)h}{r+\delta_h} + (1-\gamma)C_h\left(\frac{w}{p}\right)^{\frac{\gamma_2}{1-\gamma}}\int_a^R e^{-r(u-a)}m(u)^{\frac{1}{1-\gamma}}du\right], \tag{20}$$

where

$$m(a) = 1 - e^{-(r+\delta_h)(R-a)}$$
$$C_h = \left(\frac{z_h}{r+\delta_h}\gamma_1^{\gamma_1}\gamma_2^{\gamma_2}\right)^{\frac{1}{1-\gamma}}$$

Thus, the problem faced by a "hypothetical" individual who makes optimal decision since birth, conditional on the retirement age, is

$$Y(R) = \max -x_E - \int_6^{6+s} e^{-ra}x(a) + e^{-r(6+s)}(1-\tau)\tilde{W}(h(6+s), 6+s, R),$$

subject to

$$h(6) = h_E = h_B x_E^{\nu},$$
$$\dot{h}(a) = z_h h(a)^{\gamma_1}x(a)^{\gamma_2} - \delta_h h(a), \quad \text{for } 6 \le a \le 6+s.$$

The Hamiltonian for this problem (before age $6+s$) is simply

$$H = -x + \hat{q}(z_h h^{\gamma_1}x^{\gamma_2} - \delta_h h).$$

where the costate variable satisfies

$$\frac{d\hat{q}(a)}{da} = r\hat{q}(a) - \hat{q}(a)[\gamma_1 z_h h^{\gamma_1}x^{\gamma_2}h^{-1} - \delta_h].$$

The boundary conditions are

$$\hat{q}(6+s) = \frac{(1-\tau)w}{r+\delta}m(6+s),$$

and that the value of the Hamiltonian at the boundary point (that is, at age $6+s$) must equal the negative of the impact on the continuation value $V(h(6+s), 6+s, R)$. Formally, it is given by

$$-x(6+s) + \hat{q}(6+s)[z_h h(6+s)^{\gamma_1}x(6+s)^{\gamma_2} - \delta_h h(6+s)]$$
$$= -\frac{\partial[e^{-r(6+s)}V(h(6+s), 6+s, R)]}{\partial s}.$$

Since the optimal choice of investment in school quality requires that

$$x(a) = \hat{q}(a)\gamma_2 z_h h(a)^{\gamma_1}x(a)^{\gamma_2}, \ 6 \le a \le 6+s,$$

the choice of schooling must satisfy,

$$\frac{m(6+s)}{r+\delta_h}\Bigg[(1-\gamma_2) \tag{21}$$
$$\left(\gamma_2\frac{(1-\tau)w}{r+\delta_h}m(6+s)\right)^{\frac{\gamma_2}{1-\gamma_2}}(z_h h(6+s)^{\gamma_1})^{\frac{1}{1-\gamma_2}}-\delta_h h(6+s)\Bigg]$$
$$=\frac{h(6+s)}{r+\delta_h}[rm(6+s)+(r+\delta_h)(1-m(6+s))]$$
$$+(1-\gamma)C_h\left(\frac{w}{p}\right)^{\frac{\gamma_2}{1-\gamma}}m(6+s)^{\frac{1}{1-\gamma}}.$$

The previous condition gives a relationship between years of schooling and human capital at the end of the schooling period, $h(6+s)$. It is possible to show (see Manuelli and Seshadri (2009)) that the optimal (from an individual point of view) level of human capital at age a (for $a \leq 6+s$) is given by

$$h(a)=h_E e^{-\delta_h(a-6)}\left[1+\left(h_E^{-(1-\gamma)}\hat{q}_E^{\gamma_2}\gamma_2^{\gamma_2}z_h\right)^{\frac{1}{1-\gamma_2}}\frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r+\delta_h(1-\gamma_1)}\right. \tag{22}$$
$$\left.\left(e^{\frac{\gamma_2 r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)}-1\right)\right]^{\frac{1}{1-\gamma_1}},\ a\in[6,6+s].$$

Thus, in particular, $h(6+s)$, satisfies

$$h(6+s)=h_E e^{-\delta_h s}\left[1+\left(h_E^{-(1-\gamma)}\hat{q}_E^{\gamma_2}\gamma_2^{\gamma_2}z_h\right)^{\frac{1}{1-\gamma_2}}\frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r+\delta_h(1-\gamma_1)}\right. \tag{23}$$
$$\left.\left(e^{\frac{\gamma_2 r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}s}-1\right)\right]^{\frac{1}{1-\gamma_1}}.$$

Given the shadow price of capital $\hat{q}$, the optimal investment in early childhood capital solves

$$\max \hat{q}_E h_E - x_E,$$

subject to

$$h_E \leq h_B x_E^{\nu},$$

which implies that

$$h_E=(\nu\hat{q}_E)^{\frac{\nu}{1-\nu}}h_B^{\frac{1}{1-\nu}}.$$

Since we know that $M(a)=\hat{q}(a)h^{\gamma_1}(a)$ satisfies

$$\dot{M}(a)=M(a)\left[\frac{d\hat{q}(a)}{da}\frac{1}{\hat{q}(a)}+\gamma_1\frac{\dot{h}(a)}{h(a)}\right],$$

simple manipulations show that

$$\frac{d\hat{q}(a)}{da}\frac{1}{\hat{q}(a)} + \gamma_1 \frac{\dot{h}(a)}{h(a)} = r + \delta_h(1-\gamma_1).$$

Thus, the function $M(a)$ satisfies the first order ordinary differential equation

$$\dot{M}(a) = M(a)[r + \delta_h(1-\gamma_1)]$$

whose solution is

$$M(a) = M(6)e^{[r+\delta_h(1-\gamma_1)](a-6)}.$$

Thus,

$$\hat{q}(6+s)h(6+s)^{\gamma_1} = \hat{q}_E h_E^{\gamma_1} e^{(r+\delta_h(1-\gamma_1))s},$$

or

$$(1-\tau)\frac{w}{r+\delta} m(6+s)h(6+s)^{\gamma_1} = (\nu^{\nu} h_B)^{\frac{\gamma_1}{1-\nu}} \hat{q}_E^{\frac{1-\nu(1-\gamma_1)}{1-\nu}} e^{(r+\delta_h(1-\gamma_1))s}. \tag{24}$$

To summarize, the equilibrium values of $(h(6+s), s, \hat{q}_E)$ are the solutions to the system formed by equations (21) and (24) and the equilibrium level of human capital at the end of the schooling period given by

$$\begin{aligned} h(6+s) = (\nu \hat{q}_E)^{\frac{\nu}{1-\nu}} h_B^{\frac{1}{1-\nu}} e^{-\delta_h s} \Bigg[1 \\ + \left((\nu \hat{q}_E)^{-\frac{(1-\gamma)\nu}{1-\nu}} h_B^{-\frac{(1-\gamma)}{1-\nu}} \hat{q}_E^{\gamma_2} \gamma_2^{\gamma_2} z_h \right)^{\frac{1}{1-\gamma_2}} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \\ \left(e^{\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)} s} - 1 \right) \Bigg]^{\frac{1}{1-\gamma_1}}, \end{aligned} \tag{25}$$

The solution is conditional on R. The human capital of a working age individual is given by

$$h(a) = e^{-\delta_h(a-6-s)} h(6+s) + (r+\delta_h) C_h \left(\frac{w}{p}\right)^{\frac{\gamma_2}{1-\gamma}} \int_{6+s}^{a} e^{-\delta_h(a-u)} m(u)^{\frac{\gamma}{1-\gamma}} du, \tag{26}$$

Since effective labor supply is $h^e(a) = h(a)(1-n(a))$, it follows that the amount of human capital supplied to the market by individuals of age a is

$$h^e(a) = \begin{cases} h(a) - \left(\frac{z_h}{r+\delta_h} \gamma_1^{1-\gamma_2} \gamma_2^{\gamma_2} \left(\frac{w}{p}\right)^{\gamma_2} \right)^{\frac{1}{1-\gamma}} m(a)^{\frac{1}{1-\gamma}}, & a \in [6+s, R] \\ 0 & a \notin [6+s, R] \end{cases} \tag{27}$$

Alternative Specification of Preferences As an example of a utility function that yields the specification in the text as a reduced form but that explicitly considers the presence of some component of consumption expenditures that are job related consider a function $\hat{u}(c_\ell, c_w, \ell)$. In this formulation c_ℓ denotes consumption not related to supplying labor, c_w denotes consumption necessary to supply labor, ℓ denotes leisure. Given that total consumption expenditures are given by $c_\ell + c_w = c$, the indirect utility function is simply

$$u(c, l) = \max_{c_\ell} \hat{u}(c_\ell, c - c_\ell, \ell).$$

As an example of a functional form consistent with our specification let

$$\hat{u}(c_\ell, c_w, \ell) = \left(\frac{c_\ell^{\phi(1-\ell)+\ell} c_w^{(1-\phi)(1-\ell)}}{\phi^\phi (1-\phi)^{1-\phi}} \right)^{1-\theta} \frac{1}{1-\theta},$$

the indirect utility function over consumption expenditures corresponds to our specification with

$$1 + \zeta = (\phi^\phi (1-\phi)^{1-\phi})^{-1}.$$

Age Effects of Unexpected Changes in Assets In this section we argue that $\partial^2 R/\partial A \partial a > 0$ when evaluated at the pre-shock level of assets. The relevant first order condition is

$$W_R(a, T, R) = \frac{\theta}{\theta - 1} \frac{\tilde{G}_R(a, R)}{\tilde{G}(a, R)} [(1 + b_y D(R_n)) W(a, T, R) \tag{28}$$

$$+ e^{ra} D(R_n)(b_m + b_y W(0, a, \bar{R})) + A(a)].$$

Totally differentiating with respect to $A(a)$, we obtain that

$$M(a, R) \frac{\partial R}{\partial A} = \frac{\theta}{\theta - 1} \frac{\tilde{G}_R(a, R)}{\tilde{G}(a, R)},$$

where

$$\begin{aligned} M(a, R) \quad = \quad & - \left[(r - \upsilon(r)) + \frac{\tilde{G}_R(a, R)}{\tilde{G}(a, R)} \frac{1}{\theta - 1} (1 + \theta b_y D(R_n)) \right] W_R(a, T, R) + \\ & e^{-r(R-a)} \frac{d}{dR} y(R, R). \end{aligned}$$

Since the second order conditions require $M(a, R) \leq 0$, it follows that, as indicated in the text, $\partial R/\partial A \leq 0$.

Since the solution to our problem is time consistent, the optimal choice of retirement date is independent of age. Thus, provided that the term $A(a)$ in equation (28) is given by the optimal choice of the consumer, it must be the case that $\partial R/\partial a = 0$. Totally differentiating the previous expression with respect to a and imposing that $\partial R/\partial a = 0$ when the expression is evaluated at the optimal (pre-shock) level of assets, we obtain

$$r M(a, R) \frac{\partial R}{\partial A} + M(a, R) \frac{\partial^2 R}{\partial A \partial a} = - \frac{\theta}{\theta - 1} \frac{\tilde{G}_R(a, R)}{\tilde{G}(a, R)^2} e^{-\upsilon(r)a}.$$

Using the previous expression for $\partial R/\partial A$ we find that

$$\underbrace{M(a,R)}_{-}\underbrace{\frac{\partial^2 R}{\partial A \partial a}}_{+} = \underbrace{-\frac{\theta}{\theta-1}\frac{\tilde{G}_R(a,R)}{\tilde{G}(a,R)}[\frac{e^{-\upsilon(r)a}}{\tilde{G}(a,R)}+r]}_{-}.$$