An Extensive Look at Taxes: How does endogenous retirement affect optimal taxation?

William B Peterman
Federal Reserve Board of Governors*

May 15, 2013

Abstract

This paper considers the effect on optimal tax policy of including endogenously determined retirement in a life cycle model. Because of two changes to the Frisch labor supply elasticity, endogenous retirement causes the optimal capital tax to increase by more than seventy percent. First, endogenous retirement increases the overall aggregate Frisch labor supply elasticity since agents can change their labor supply, not only on the intensive margin, but also, on the extensive margin. In response, the government lowers the labor tax and increases the capital tax to reduce tax policy distortions. Second, since the choice to retire is more relevant for older individuals, endogenous retirement disproportionately increases an agent's Frisch labor supply elasticity when he is older compared to when he is younger individuals. Ideally, the government would decrease the relative labor income tax on individuals when they are older. If age-dependent taxes are not available then the government mimics such a tax policy by further increasing the capital tax. I find that in my benchmark specification the welfare lost from not accounting for endogenous retirement when determining the optimal tax policy is equivalent to approximately 0.6 percent of total lifetime consumption.

JEL: E24, E62, H21.

Key Words: Optimal Taxation, Capital Taxation, Endogenous Retirement.

*E-mail: william.b.peterman@frb.gov. Views expressed on this site are my own and do not reflect the views of the Federal Reserve System or its staff. For extensive discussion and helpful comments, I thank Glenn Follette, Kevin Novan, Scott Borger, two anonymous referees, as well as seminar participants at the Congressional Budget Office, the Symposium of the Society for Nonlinear Dynamics and Econometrics, and the Conference on Computing in Economics and Finance.

1 Introduction

Previous work demonstrates that in a life cycle model the optimal capital tax is related to the representative cohort's Frisch labor supply elasticity (see Garriga (2001), Atkeson et al. (1999), Erosa and Gervais (2002), Conesa et al. (2009), and Peterman (2013)). Even though decisions on both the intensive and extensive margins affect this elasticity, previous studies that determine the optimal capital tax in a life cycle model typically include retirement exogenously. Motivated by these findings, this paper aims to determine the effect on optimal tax policy of relaxing this assumption and including endogenous retirement.

To assess the quantitative effect of this assumption, I computationally determine the optimal tax policy in two similar life cycle models where the government is required to raise a fixed amount of revenue. In the first model, individuals are forced to retire at an exogenously set age. In the second model, agents endogenously determine when they retire. I find that the optimal capital tax is approximately seventy percent larger in the model with endogenously determined retirement compared to the exogenous model. These results indicate that modeling endogenous retirement is of first order importance when determining the optimal tax policy.

Next, I demonstrate that this relationship between the optimal tax policy and endogenous retirement is robust to two extensions. First, I demonstrate the results are robust to including a more stylized social security program that, like the U.S. social security program, includes both a progressive replacement rate and credits (deductions) for retiring after (before) the normal retirement age. These features are in contrast to the benchmark model which includes a more parsimonious social security program that simply replaces a fraction of an agent's pre-retirement income. Second, I find a similar relationship between endogenous retirement and optimal tax policy when I include idiosyncratic shocks to wages that replicate the wage dynamics estimated in the U.S. economy. Overall, in both cases, including endogenous retirement causes an economically significant increase in the optimal capital tax, although increase is somewhat smaller.

Including endogenous retirement increases the optimal capital tax because it alters the agent's aggregate Frisch labor supply elasticity through two channels.[1] First, by removing the restrictions on labor participation, the agent has two margins (intensive and extensive) by which he can alter his labor supply. In the endogenous model, the aggregate Frisch elasticity is larger compared to the exogenous model where the agent can only alter his labor supply on the intensive margin. Therefore, to mitigate the distortions imposed by the tax code, the government wants to decrease the labor tax and increase the capital tax.[2] Second, since the choice to retire is more relevant for older individuals, including endogenous retirement disproportionately increases the Frisch elasticity of an agent when his is old compared to when he is young. Consequently, the government would like to use age-dependent labor income taxes to decrease the relative tax on the labor income when an agent is old and more responsive. Since I assume that the government cannot use age-dependent taxes, the government mimics such a tax policy by increasing the capital tax.

Empirical studies indicate that decisions on the extensive margin affect both the level of the Frisch labor supply elasticity and the slope of the lifetime Frisch elasticity profile. Therefore, it is natural to examine the relationship between endogenous retirement and the optimal tax policy. Starting with the level, using simulated data, Rogerson and Wallenius (2009) demonstrate that ignoring the extensive margin could significantly lower estimates of the aggregate Frisch elasticity. Moreover, Erosa et al. (2011) find in a calibrated life cycle model that the intensive and extensive margins both account for equal parts of the aggregate labor supply response to a temporary wage change. Finally, Chetty et al. (2011), Peterman (2012), and Fiorito and Zanella (2012) find that incorporating fluctuations on the extensive margin causes empirical estimates of the Frisch elasticity to increase substantially.

With respect to the slope, some additional evidence suggests that older individuals have a higher Frisch elasticity then younger individuals. French (2005), estimating a life cycle model in that allows for a break point in the Frisch elasticity parameter at the age of forty, finds that the Frisch elasticity is over three times larger for the older than to younger

[1] I define the aggregate Frisch elasticity to incorporate changes on both the intensive and extensive margin.

[2] A standard result in public finance is that, if it is necessary to use distortionary taxes, it is optimal to tax inelastically supplied factors at relatively higher rates since this policy will minimize the distortions to the economy.

individuals.[3] This upward sloping Frisch labor supply elasticity profile is consistent with retirement decisions having a disproportionate effect on the elasticity of older individuals.

I choose to limit my focus to endogenous retirement, as opposed to including both endogenous entry and exit from the labor force, for two reasons. First, the empirical evidence is consistent with the extensive margin having a larger impact on the labor supply elasticity of older individuals.[4] Second, Jacobs and Bovenberg (2009) demonstrate that including the decision of when to stop education and enter the workforce could further enhance the motive for a positive capital tax. The authors analyze the trade off between a labor and capital tax in a two-period life cycle model where agents acquire education in the first period and work in the second period. They find that the optimal capital tax is generally positive since a capital tax reduces the necessary labor income tax, which in turn reduces the distortions on the benefit to education. Since previous work indicates that including an endogenous entry decision may further increase the optimal capital tax, the results in this paper can be viewed as a lower bound of the total effect on optimal taxation of assuming labor force participation is exogenous.

This paper is organized as follows: Section 2 introduces the model, and presents the competitive equilibrium. Section 3 describes the functional forms and calibration parameters. Section 4 sets up the computational experiment and section 5 reports the results of the computational experiment. Section 6 demonstrates the robustness of these results and section 7 concludes.

[3] One counter example is Clark and Summers (1981) which finds evidence that the extensive margin causes teenage individuals to be more responsive to the business cycle. My study is primarily focused on individuals after they have finished school and entered the workforce, which excludes many of these teenagers.

[4] It is possible that individuals could make more than just one entry and exit decision over their working lifetime. However, Erosa et al. (2011) document that from the ages of twenty two through fifty only a small fraction of individuals work less than one hundred hours. The small fraction not working over the normal working life implies that generally once a head of household starts working he tends to continue until he retires.

2 Model

In this section, I describe the computational model and stationary competitive equilibrium which I use to determine the effect of endogenous retirement on optimal tax policy. In this baseline model I include a parsimonious social security program and no idiosyncratic wage shocks. Section 6 examines the effect of endogenous retirement on optimal taxation in a model that is extended to include a more stylized social security program that encompasses many of the features in the U.S.'s system and also includes idiosyncratic wage shocks calibrated to match wage dynamics in the U.S. economy.

2.1 Demographics

In the model, time is assumed to be discrete and there are J overlapping generations. Ψ_j is the probability of an agent living to age $j+1$ conditional on being alive at age j. All agents who live to an age of J die the next period. If an agent dies with assets, the assets are confiscated by the government and distributed equally to all the living agents as transfers (Tr_t).[5] An agent retires at the age j_r. In the exogenous model, agents are forced to retire at the age j_{exog} ($j_r = j_{\text{exog}}$). In the endogenous model, agents choose j_r.

In each period a continuum of new agents is born. The population of new agents born each period grows at rate n. Given the population growth rate and conditional survival probabilities, the time invariant cohort shares, $\{\mu_j\}_{j=1}^{J}$, are given by

$$\mu_j = \frac{\Psi_{j-1}}{1+n}\mu_{j-1}, \text{for } i = 2, ..., J, \tag{1}$$

where μ_1 is normalized such that

$$\sum_{j=1}^{J} \mu_j = 1 \tag{2}$$

[5]Closing the model this way affects the optimal tax on capital in both models. Peterman (2013) demonstrates that the government not being able to distinguish between these transfers and ordinary capital causes the optimal tax on capital to increase.

2.2 Individual

Prior to retirement, an individual is endowed with one unit of productive time per period which he splits between providing labor services and leisure. After retirement he uses all his time for leisure. An agent maximizes his expected stream of lifetime utility at each age,

$$u(c_j, h_j) + \sum_{s=1}^{J-j-1} \beta^s \prod_{q=1}^{s} \Psi_q u(c_{s+1}, h_{s+1}), \tag{3}$$

where c_j is the consumption of an agent at age j and h_j is the percent of the time endowment that is spent providing labor services. The agent faces a fixed utility cost to working which implies the disutility from working discontinuously increases when an agent goes from zero to positive hours worked.[6] The discount factor conditional on surviving is β.

An agent receives labor income of $h_j \varepsilon_j w_t$ where ε_j is the agent's age-specific human capital. This labor income is split between consumption and saving using a risk free asset. An agent's stock of assets is denoted by a_j and he receives a pre-tax net return of r_t on the assets per period.

2.3 Firm

Firms are perfectly competitive with constant returns to scale production technology. Aggregate technology is represented by a Cobb-Douglas production function. The aggregate resource constraint is

$$C_t + K_{t+1} - (1-\delta)K_t + G_t \leq K_t^{\alpha} N_t^{1-\alpha}, \tag{4}$$

where K_t, C_t, N_t, G_t, α, and δ represent the aggregate capital stock, aggregate consumption, aggregate labor (measured in efficiency units), government consumption, the capital share,

[6] An alternative formulation that would induce agents to make decisions on the extensive margin is to include a non-linear mapping between hours and productivity (for example see Rogerson and Wallenius (2009)). Although both modeling options create an active extensive margin, I found that solving for a steady state when using a fixed cost was more stable with respect to initial guesses.

and the depreciation rate for physical capital, respectively.

2.4 Government Policy

The government partakes in two activities. First, the government runs a reduced form pay-as-you-go social security program. Second, the government consumes in an unproductive sector. The social security program is included so that, similar to the U.S. economy, agents do not fund all of their post-retirement consumption from savings. Peterman (2013) demonstrates, in a similar model, that if a social security program is not included then the large incentive to save cause the model to predict unrealistic life cycle profiles. Therefore, I include a reduced form program that aligns the general savings incentives in the model and the that.

In this reduced from social security program the government pays an annual benefit, $SS_t(s_t, j_r)$, to all retired individuals. These benefits are a function of both the retirement age, j_r, and the individual's annual labor income s_t. Specifically, the benefits are set such that retired agents receive an exogenously determined fraction, b_t, of their average annual income (s_t.) The social security program is financed by taxing labor income at a flat rate, $\tau_{ss,t}$. The payroll tax rate $\tau_{ss,t}$ is set to ensure the social security system has a balanced budget each period. The social security system is not considered part of the tax policy that the government optimizes.

In the exogenous model agents retire at the age j_{exog} which is considered the "normal retirement age". Similar to the general spirit of the US social security program, in the endogenous model if an agent chooses to retire within a nine year window around the normal retirement age ($j_r \in (j_{\text{nr}} - 4, j_{\text{nr}} + 4)$) the total lifetime benefits from the social security program are actuarially equivalent.[7] For example, if an agent chooses to retire a year earlier (later) than j_{nr} then their annual benefit is reduced (increased) in order to reflect that they receive one additional (less) payment over their retired lifetime. These

[7]See Meyer and Reichenstein (2010) for a discussion about how accurate the U.S. program is at being actuarially equivalent.

adjustments are determined so that in expectation, regardless of when an agent decides to retire, the sum of his total lifetime benefits from the social security program are equal. Since the adjustments are determined such that they are actuarially equivalent, the social security program does not directly affect the retirement age.

The second activity that the government partakes in is consuming an exogenously determined amount, G_t, in an unproductive sector.[8] The government uses two fiscal instruments to finance this consumption. First, the government taxes capital income, $y_k \equiv r_t(a + Tr_t)$, according to a capital income tax schedule $T^K[y_k]$. Second, the government taxes part of each individual's labor income. Half of the pre-tax labor income is accounted for by the employer's contributions to social security, which is not taxable under current U.S. tax law. Therefore, the taxable labor income, which is taxed according to a labor income tax schedule $T^l[y_l]$, is $y_l \equiv w_t \varepsilon_j h_j (1 - .5\tau_{ss})$. I impose three restrictions on the labor and capital income tax policies. First, I assume anonymity of the tax code so the rates cannot be personalized, nor can they be age-dependent. Second, both of the taxes are functions only of the individual's relevant taxable income in the current period. Finally, as is typical in a Ramsey problem, I assume the government cannot use lump sum taxation.

2.5 Definition of Stationary Competitive Equilibrium

In this section I define the competitive equilibria for the computational model. Given a social security benefits formula $B_{ss} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, government expenditures G, and a sequence of population shares $\{\mu_j\}_{j=1}^J$, a stationary competitive equilibrium is a sequence of agent allocations, $\{c_j, a_{j+1}, h_j, j_r\}$, a production plan for the firm (N, K), a government labor tax function $T^l : \mathbb{R}_+ \to \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \to \mathbb{R}_+$, a social security tax rate τ_{ss}, a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, a formula to calculate the lifetime average annual earnings $AE : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}+$, prices (w, r), and transfers Tr such that:

[8] A formulation that induces the same optimal tax policy is if G_t enters the agent's utility function in an additively separable manner.

1. Given prices, policies, transfers, and benefits the agent maximizes equation 3 subject to

$$c_j + a_{j+1} = w\varepsilon_j h_j - \tau_{ss} w s_j h_j + (1+r)(a_j + Tr) - T^l[w\varepsilon_j h_j(1 - .5\tau_{ss})] - T^k[r(a_j + Tr)] \text{ for } j < j_r,$$

$$s_{j+1} = AE(w\varepsilon_j h_j, s_j, j) \text{ for } j < j_r,$$

$$c_j + a_{j+1} = B_{ss}(s_j) + (1+r)(a_j + Tr) - T^k[r(a_j + Tr_t)], \text{ for } j \geq j_r$$

$$h_j = 0, \text{ for } j \geq j_r$$

$$c_j \geq 0, 0 \leq h_j \leq 1, a_j \geq 0, \text{ and } a_1 = 0.$$

2. Prices w and r satisfy

$$r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \text{ and } w = (1-\alpha)\left(\frac{K}{N}\right)^{\alpha}$$

3. The social security policies satisfy

$$\sum_{j=1}^{j_r-1} \varepsilon_j w \mu_j h_j \tau_{ss} = \sum_{j=j_r}^{J} \mu_j B(s_j).$$

4. Transfers are given by

$$Tr = \sum_{j=1}^{J} \mu_j (1 - \Psi_j) a_{j+1}.$$

5. Government budget balance

$$G = \sum_{j=1}^{J} \mu_j T^k[r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j T^l[w\varepsilon_j h_j(1 - .5\tau_{ss})].$$

6. Market clearing

$$K = \sum_{j=1}^{J} \mu_j a_j, \; N = \sum_{j=1}^{J} \mu_j \varepsilon_j h_j \text{ and}$$

$$\sum_{j=1}^{J} \mu_j c_j + \sum_{j=1}^{J} \mu_j a_{j+1} + G = K^{\alpha} N^{1-\alpha} + (1-\delta)K.$$

3 Calibration and Functional Forms

To determine the optimal tax policy it is necessary to choose functional forms and calibrate the model's parameters. I base the model on Conesa et al. (2009) and Peterman (2013). As is typical with a calibration exercise, determining parameter values involves a two-step process. The first step is choosing parameter values for which there are direct estimates in the data. Second, the remaining parameters are determined so that under the baseline-fitted U.S. tax policy, certain targets in the model match the values observed in the U.S. economy.[9] I calibrate all of this second group of parameters in a model with endogenously determined retirement, which implies that the parameter values are the same in both the exogenous and endogenous models. The parameter values are listed in Table 1.

3.1 Demographics

In the model, agents are born at a real world age of 20 that corresponds to a model age of 1. The current population in the U.S. faces a normal retirement age of between 65 and 66.[10] Since I am calibrating the model to the current U.S. economy, I choose the normal retirement age to be 66. In the exogenous model agents are forced to retire at an age of 66. In the endogenous model, the fixed cost to working parameter (χ_2) is calibrated such that individuals choose to retire at age 66.[11] If an individual survives until the age of 100, he dies the next period. I set the conditional survival probabilities in accordance with the estimates in Bell and Miller (2002). I assume a population growth rate of 1.1 percent.

[9] Since these are general equilibrium models, changing one parameter will alter all the values in the model that are used as targets. However, I present the parameter with the target that has the most direct correspondence.

[10] This population includes individuals who are currently retired.

[11] I found that the results of the computational experiment were similar when I used a normal retirement age of 65.

Table 1: **Calibration Parameters**

Parameter	**Value**	**Source**
Demographics:		
Normal Retirement Age: j_{exog}	66	By Assumption
Max Age: J	100	By Assumption
Surv. Prob: Ψ_j		Bell and Miller (2002)
Pop. Growth: n	1.1%	Data
Firm Parameters:		
α	.36	Data
δ	8.33%	$\frac{I}{Y} = 25.5\%$
A	1	Normalization
Calibration Parameters:		
Conditional Discount: β	0.999	$\frac{K}{Y} = 2.7$
Risk aversion: σ_1	2	Conesa et al. (2009)
Frisch Elasticity: σ_2	0.5	Intensive Frisch= $\frac{1}{2}$
Disutility to Labor: χ_1	51.1	Avg. $h_j = \frac{1}{3}$
Fixed Cost to Working: χ_2	0.34	Normal Retirement = 66
Government Parameters:		
Υ_0	.258	Gouveia and Strauss (1994)
Υ_1	.768	Gouveia and Strauss (1994)
G	0.179	17% of Y
b	0.4	Rettenmaier and Saving (2006)

3.2 Preferences

Agents have time-separable preferences over consumption and labor services, and conditional on survival, they discount their future utility by β. I determine β such that the capital-to-output ratio matches U.S. data of 2.7.[12] I use a utility function that is separable and homothetic in both consumption and labor. I choose to use this type of utility function since both Garriga (2001) and Peterman (2013) demonstrate that violating these assumptions will lead to a large optimal tax on capital and I do not want this motive to be confounded with a motive from endogenous retirement.

The homothetic and separable utility function I use is $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} - \chi_2$. I determine χ_2 such that under the baseline-fitted U.S. tax policy, individuals choose to retire at the normal retirement age (66) in the endogenous model. Following Conesa et al. (2009), I set $\sigma_1 = 2$, which controls the relative risk aversion.[13] The parameter σ_2 controls the Frisch elasticity on the intensive margin (intensive Frisch elasticity). The intensive Frisch elasticity is different from the aggregate Frisch elasticity in that it only incorporates changes in hours on the intensive margin while the aggregate Frisch elasticity incorporates changes in hours on both the intensive and extensive margins. The intensive Frisch elasticity is equal to σ_2, while the aggregate Frisch elasticity is implicitly determined within the model. Past micro-econometric studies estimate that the intensive Frisch elasticity is between 0 and 0.5.[14] However, more recent research has shown that these estimates may be biased downward. Reasons for this bias include utilizing weak instruments, not accounting for borrowing constraints, disregarding the life cycle effect of endogenous age-specific human capital, omitting correlated variables such as wage uncertainty, and not accounting for labor market frictions.[15] Therefore, I set σ_2 such that the intensive Frisch elasticity is at the upper

[12] This is the ratio of fixed assets and consumer durable goods, less government fixed assets, to GDP (Conesa et al. (2009)).

[13] Even though Conesa et al. (2009) use a different utility specification, their specification has a parameter that corresponds to σ_1.

[14] For examples see Altonji (1986), MaCurdy (1981), and Domeij and Floden (2006).

[15] Some of these studies include Imai and Keane (2004), Domeij and Floden (2006), Pistaferri (2003), Chetty (2009), and Contreras and Sinclair (2008).

bound of the range (0.5).[16] I calibrate $\{\varepsilon_j\}_{j=0}^{j_r-1}$ using a smoothed version of the estimates in Kaplan (2012) for the part of the labor profiles that deterministically evolves over the lifetime.[17]

3.3 Firm

I assume the aggregate production function is Cobb–Douglas. The capital share parameter, α, is set at .36. The depreciation rate is set to target the observed investment output ratio of 25.5 percent.

3.4 Government Policies and Tax Functions

When determining the second set of parameters such that the targets in the models match the values in the data, it is necessary to use a baseline tax function that mimics the U.S. tax code. I use the estimates of the U.S. tax code from Gouveia and Strauss (1994), which I refer to as the baseline-fitted U.S. tax policy. The authors match the U.S. tax code to the data using a three parameter functional form,

$$T(y;\Upsilon_0,\Upsilon_1,\Upsilon_2) = \Upsilon_0(y-(y^{-\Upsilon_1}+\Upsilon_2)^{-\frac{1}{\Upsilon_1}}), \tag{5}$$

where y represents total labor or capital income. The average tax rate is principally controlled by Υ_0, and Υ_1 governs the progressivity of the tax policy. The third parameter, Υ_2, is set so that taxes satisfy the budget constraint. Gouveia and Strauss (1994) estimate that $\Upsilon_0 = .258$ and $\Upsilon_1 = .768$. The authors do not fit separate tax functions for labor and capital income. Accordingly, under the baseline-fitted U.S. tax policy, the government cannot distinguish labor and capital income. I calibrate government consumption, G, so that it equals

[16] Various studies have demonstrated that adding an endogenous decision about the extensive margin means that the representative agent's Frisch elasticity will be larger (for examples see Chetty (2009), Chetty et al. (2011), Rogerson and Wallenius (2009), and Erosa et al. (2011)).

[17] The profile is smoothed using a quadratic function in age, and extended to cover ages 20 through 70.

17 percent of output under the baseline-fitted U.S. tax policy, as observed in the U.S. data.[18] More specifically, Υ_2 is determined as the value that causes taxes to be equal to 17 percent of GDP. When searching for the optimal tax policy, I restrict my attention to tax policies that induce revenue neutral changes. These policies imply that government consumption is equal under the baseline-fitted U.S. tax policy and the optimal tax policy.

In addition to consuming, the government runs a balanced-budget social security program. Social security benefits are set in the simplified model such that the social security benefits equal a fixed percent (b) of the agent's average annual labor income (s_j). I set the replacement rate, b, at 40 percent according to Rettenmaier and Saving (2006).[19] In the U.S. social security system average annual labor income is calculated as the individual's average income over their 35 highest earning years. For tractability reasons, I simplify the way that an agent's average annual earnings is calculated so it is not necessary to include the entire history of earnings as a state variable. Following French (2005) and Kitao (2012), I update the state variable s_j during the first 35 years after the individual enters the economy as, $s_{j+1} = \frac{s_j(j-1)+y_{l,j}}{j}$, where $y_{l,j}$ is the labor earnings of an individual in that period. After 35 years, I assume that s_{j+1} is only updated if the new earnings exceed the current value, $s_{j+1} = \max\{s_j, \frac{s_j(j-1)+y_{l,j}}{j}\}$. Moreover, in the endogenous model where agents can choose when to retire, if agents retires prior to (after) the normal retirement age then their annual benefit is reduced (increased). These reductions (increases) in the social security benefit for early retirement are calculated such that at the time of actual retirement the net present value of the total stream of their benefits will be equal to those that they would have expected to receive if they had retired at the normal retirement age. The payroll tax, τ_{ss}, is determined so that the social security system is balanced each period.

[18]To determine the appropriate value for calibration, I focus on government expenditures less defense consumption.

[19]In addition, the Social Security Administration estimates that the replacement ratio for the median individual is 40 percent (see table VI.F10 in the 2006 Social Security Trustees Report; available at www.ssa.gov/OACT/TR/TR06/).

4 Computational Experiment

The computational experiment is designed to determine the tax policy that maximizes a given social welfare function. I choose a social welfare function (SWF) that corresponds to a Rawlsian veil of ignorance (Rawls (1971)). Since living agents face no earnings uncertainty, the social welfare is equivalent to maximizing the expected lifetime utility of a newborn,

$$SWF(T^h, T^k) = u(c_1, h_1) + \sum_{j=1}^{J} \beta^j [\prod_{q=0}^{j} \Psi_q] u(c_{j+1}, h_{j+1}), \tag{6}$$

where T^h is the tax policy on labor income and T^k is the tax policy on capital income.

5 Results

I begin by examining the fit of my model compared to the data. Figure 1 plots the life cycle profiles predicted by the endogenous model under the baseline fitted U.S. tax policy and the profiles from the data.[20] I begin by focusing on the percent of the time endowment spent working (the upper left panel) and the compensation for the time spent working (the upper right panel). The labor and labor earnings profiles from the data are constructed from the 1967 - 1999 waves of the Panel Survey of Income Dynamics (PSID). In the model, on average, agents begin by working 37 percent of the total labor endowment at age 20; this fraction generally decreases gradually as the households age, reaching roughly 25 percent of the total labor endowment by age 65. The shape of the model generated profile roughly matches the shape of the average profile from the data. However, for young households between the age of twenty to twenty five, the model over predicts the amount of time spent working compared to the data. In the model agents cannot borrow against future earnings. Therefore, the positive wealth effect that labor supply has on household consumption more than offsets the disincentive to work even though wages are relatively low at the beginning

[20]Earnings, consumption, and savings from the model are converted to real 2012 dollars using the ratio of earnings in the model to the data.

of the life cycle (for further discussion of this prediction, see Heathcote et al. (2010)). Moving to labor earnings, the profiles from the model and data are both hump-shaped with a peak around forty years old.[21]

The lower left panel compares the model predicted consumption profile and the per-capita expenditures on food in the PSID. I find that both profiles are hump-shaped; however, consumption on food tends to peak earlier in the data than total consumption peaks in the model. Additionally, comparing the growth in consumption from age 20 to the peak, the model predicts more growth than is observed in the data. One possible reason for these differences is that the PSID data are limited to expenditures on food while the model generated consumption represents all consumption.

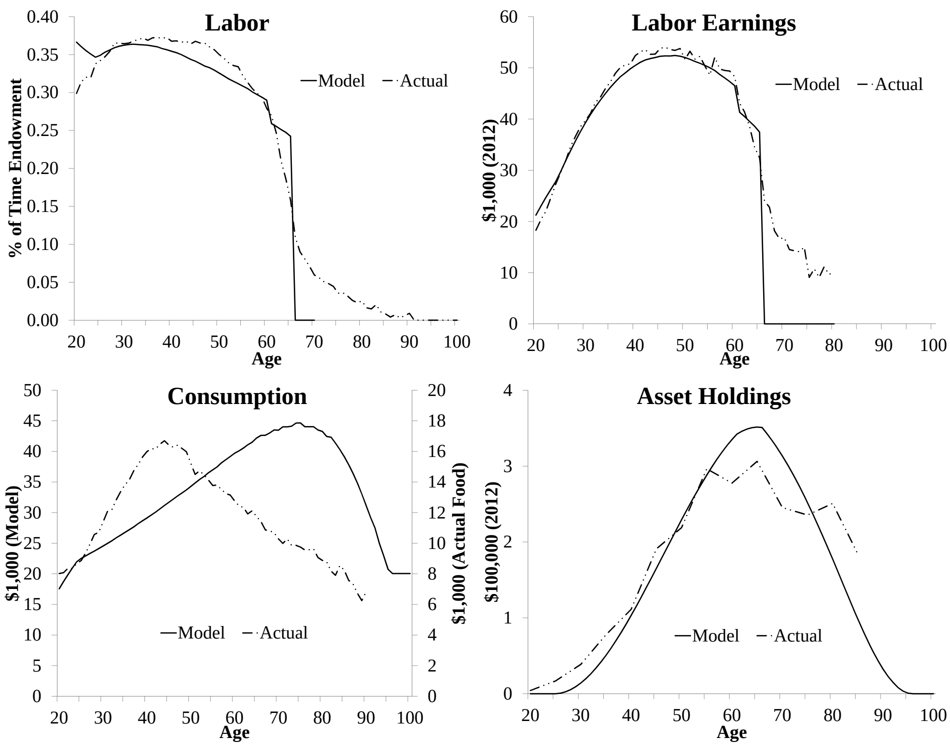
Finally, the lower right panel examines savings in the model and the median of total wealth in the 2007 Survey of Consumer Finances (SCF) for individuals between the ages of 20 and 80.[22] I find that the two profiles are similar; both are hump-shaped, peaking around \$300,000 at the age of 60. One discrepancy between the profiles is that the model predicts that agents will deplete their savings more quickly than is observed in the data. One possible explanation for this discrepancy is that the model does not include any motive for individuals bequest assets to younger generations. Despite these differences, overall the model predicted profiles seem to do a reasonably good job fitting the data.

Turning to the optimal tax policies, I find that flat taxes as opposed to progressive taxes are optimal. Therefore, for notational convenience, I represent the optimal capital and labor tax as τ_k, and τ_h, respectively. Table 2 lists the optimal tax policies in both models. The optimal tax policy in the exogenous model is a 17.9 percent capital tax ($\tau_k = 17.9\%$) and a 23.5 percent labor tax ($\tau_h = 23.5\%$). I find that including endogenously determined retirement has a large impact on the optimal tax policy. The optimal capital tax increases,

[21] I focus my attention on the earnings for the head of the household between ages 20 and 80 throughout the whole sample. I found that the data for individuals older than 80 were extremely volatile.

[22] The wealth profile is smoothed using five year age averages. The data for individuals after age 80 was not included because there were few observations in the sample leading the estimates to be extremely volatile. In the model there is no within cohort heterogeneity. Therefore, in order to prevent the upper tail of the wealth distribution from skewing the statistic for comparison, I choose to focus on the median level of wealth as opposed to the average.

Figure 1: **Life Cycle Profiles**



Note: The model values are converted to actual dollar amounts by comparing the average earnings in the model and the data. The actual labor, earnings, and consumption profiles are from the PSID, while the capital profile is from the SCF.

by more than seventy percent, to 30.6 percent when endogenous retirement is included.

Next I determine the welfare effect of ignoring endogenous retirement when solving for the optimal tax policy. Specifically, I determine how much welfare is lost in the endogenous model if the tax policy is switched from the optimal, which calls for a larger tax on capital, to the optimal tax policy from the exogenous model, which calls for a smaller tax on capital. I find that in the endogenous model, ignoring endogenous retirement when solving for the optimal tax policy causes welfare to decrease by an amount equivalent to approximately 0.6 percent of total lifetime consumption. This result implies that there are non-trivial effects to welfare of determining the optimal tax policy in a misspecified model.

Table 2: **Optimal Tax Policies**

Tax Rate	**Exogenous**	**Endogenous**
τ_k	17.9%	30.6%
τ_h	23.5%	20.4%
$\frac{\tau_k}{\tau_h}$	0.76	1.5

Two differences in the aggregate Frisch labor supply elasticity are responsible for the disparate optimal tax policies in the exogenous and endogenous models.[23] First, including an endogenous retirement decision means that the agents can change their labor supply on an additional margin. This additional margin results in an increase in the overall level of the aggregate Frisch labor supply elasticity. To minimize the distortion that the tax policy induces, the government relies more heavily on a capital income tax as opposed to a labor income tax. Second, since the choice to retire is more relevant for individuals when they are older, endogenous retirement disproportionately increases the Frisch labor supply elasticity when agents are older compared to when they are younger. In response to this change in the slope of the aggregate Frisch elasticity profile, the government wants to reduce the

[23] As noted, the aggregate Frisch elasticity is determined within the model and incorporates labor changes on both the intensive and extensive margin. In contrast, the intensive Frisch elasticity only incorporates changes in hours on the intensive margin and is equal to σ_2.

relative tax rate on the labor income earned when agents are old. Since the government cannot condition labor income taxes on age, a larger capital tax is used to mimic this type of age-dependent tax policy.

Next, I quantify the relative effect of the change in the level and the slope of the aggregate Frisch elasticity profile from including retirement endogenously on the optimal tax policy. I start by quantifying the effect of the increase in the level. I find that the aggregate Frisch elasticity increases from 0.5 in the exogenous model to 1.07 in the endogenous model.[24] To determine the effect of this increase, I determine the optimal tax policy in an altered exogenous model in which the Frisch elasticity is increased to 1.07. I find that the optimal capital tax increases from 17.9 percent to 25.4 percent in the altered exogenous model.[25] This result indicates that the increase in the level of the aggregate Frisch elasticity is responsible for a 7.5 percentage point rise in the optimal capital tax.

To quantify the effect of the change in the slope of the Frisch elasticity profile, I solve for the optimal tax policy in the endogenous model where I allow the government to tax labor income at a different rate when agents are within their retirement window. I find that when I allow for this age-dependent tax, the optimal capital tax drops 7.2 percentage points, from 30.6 percent to 23.4 percent.[26] Overall, the similar changes in the optimal capital tax in the two experiments indicates that the slope channel and the level channel are both responsible for approximately half of the total difference between the optimal tax policies in the exogenous and endogenous models.

Finally, I focus on the differences in the underlying economies with and without en-

[24]I calculate the aggregate Frisch elasticity in the endogenous model by regressing the percent change in hours on wages and marginal utility while restricting my model such that the constant equals zero. In order to confirm that this is an accurate measurement of the Frisch elasticity in the model, I estimate the regression on ages in which the agent does not consider retirement. I find that the estimate of the Frisch elasticity for these younger agents is very close to the parameter value which determine the intensive Frisch elasticity (0.549 versus 0.5). The slightly larger estimate of the Frisch elasticity compared to the parameter value is because in the endogenous model agents internalize both the effect of the change in the wage on earnings this period and on the average annual earnings used to calculate the social security benefit. These additional benefits in the endogenous cause the agent to be more responsive to changes in the wages.

[25]Note, I recalibrate the model such that it matches the other targets.

[26]The sum of the two changes does not add up to the difference between the optimal tax rates in endogenous and exogenous models because of general equilibrium effects and also when both channels are included their interaction slightly diminishes the overall effect.

Table 3: **Aggregate Economic Variables**

	Exogenous			Endogenous		
Aggregate	**Baseline**	**Optimal**	***%Δ from Baseline to Optimal***	**Baseline**	**Optimal**	***%Δ from Baseline to Optimal***
Y	1.06	1.08	1.9%	1.06	1.06	0.7%
K	2.85	2.98	4.4%	2.85	2.8	-1.8%
N	0.6	0.61	0.6%	0.6	0.62	2.1%
Avg Hours	0.33	0.34	0.7%	0.33	0.33	0.2%
w	1.12	1.13	1.4%	1.12	1.1	-1.4%
r	0.05	0.05	-6.3%	0.05	0.05	6.8%
tr	0.04	0.05	3.4%	0.04	0.04	-2.3%
Value	-136.87	-136.42	0.3%	-136.87	-135.89	0.7%
CEV			0.5%			1%
Average Tax Rate	**Baseline**	**Optimal**		**Baseline**	**Optimal**	
Capital	21%	17.9%		21%	30.6%	
Labor	21%	23.5%		21%	20.4%	
Ratio		0.76			1.5	
Marginal Tax Rate	**Baseline**	**Optimal**		**Baseline**	**Optimal**	
Capital	24.5%	17.9%		24.5%	30.6%	
Labor	24.5%	23.5%		24.5%	20.4%	
Ratio		0.76			1.5	

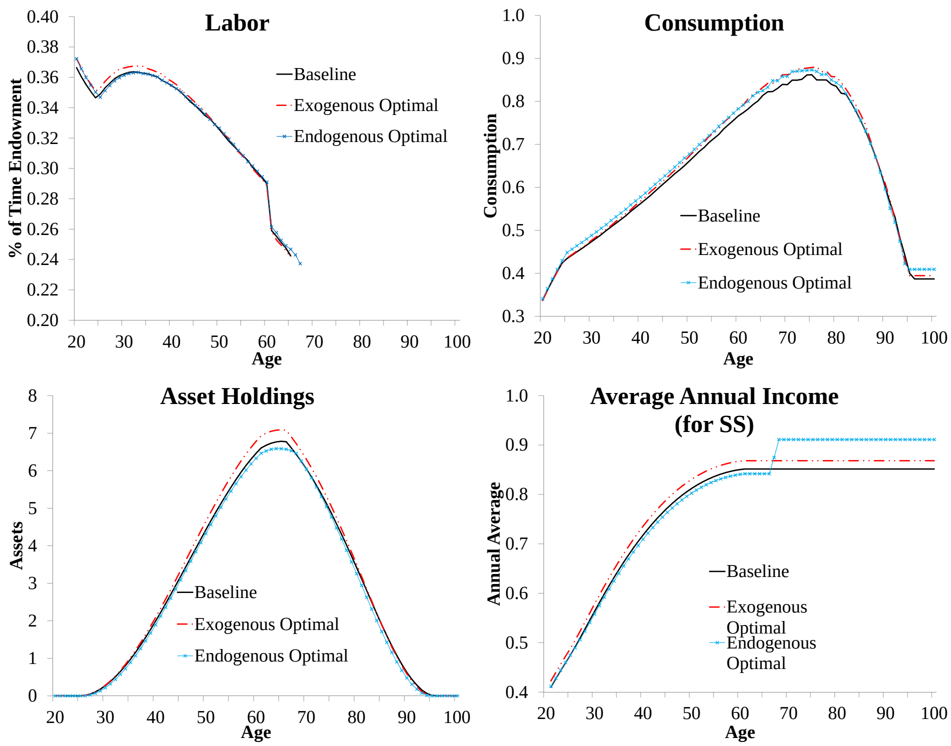
dogenous retirement. Table 3 details the aggregate economic variables; the first, second and third columns describe the model with exogenous retirement and the fourth, fifth, and sixth columns detail the model with endogenous retirement. The first and fourth columns describe the aggregates under the baseline-fitted US tax policy. Since the two models have the same calibration parameters and targets, they are identical under the baseline-fitted U.S. tax policy. The second and fifth columns describe the exogenous and endogenous models when I impose each model's optimal tax policy, respectively. Finally, the third and sixth columns describe the percent change in the aggregates that occurs when I switch from the baseline-fitted U.S. tax policy to the optimal tax policy. Figure 2 plots the life cycle profiles for labor, consumption, and savings. The black lines are the profiles under the baseline-fitted U.S. tax policy. As previously mentioned, by construction the endogenous and exogenous models are identical under the baseline-fitted U.S. tax policy. The red lines are the profiles in the exogenous model under the optimal tax policy. The blue lines are the profiles in the endogenous model under the optimal tax policy.

In the exogenous model, adopting the optimal tax policy causes the average tax rate on capital to fall and the average tax rate on labor to increase.[27] In response to changing from the baseline to optimal tax policy, capital increases and labor stays relatively constant. The changes in these aggregates causes output to increase, the pre-tax wage rate to increase, and the pre-tax return to capital to fall. In order to assess the impact on welfare I calculate the consumption equivalent variation (CEV), which is the uniform percentage increase in consumption required to make an individual indifferent between the baseline and optimal tax policy. I find that the CEV is 0.5 percent of lifetime consumption in the exogenous model.

In the endogenous model, adopting the optimal tax policy causes a large increase in the tax on capital income and a decrease in the tax on labor income. In response to changing from the baseline-fitted U.S. tax policy to the optimal tax policy, aggregate capital decreases

[27] Note that the baseline-fitted U.S. tax policy is progressive and taxes both labor and capital income at the same rate. Therefore, the average tax rate on capital reported for the baseline-fitted U.S. tax policy is the average income tax rate paid by all individuals and the average tax rate on labor is the average income tax rate paid by working individuals.

Figure 2: **Life Cycle Profiles**



Note: These plots are life cycle profiles in both models under the baseline-fitted U.S. tax policy and the optimal tax policies. Since the calibration parameters and targets are the same, the economies in both models are identical under the baseline-fitted U.S. tax policy.

and aggregate labor increases. The changes in capital and labor cause overall output to increase almost one percent. The increase in labor and decrease in capital leads to an increase in the pre-tax return to capital and decrease in the pre-tax wage. Overall, I find that the welfare gains from adopting the optimal tax policy are twice as large in the endogenous model compared to the exogenous model (1.0%).

Next, I examine the effect of adopting the optimal tax policy in the exogenous model in the life cycle profiles in Figure 2. Adopting the optimal tax policy in the exogenous model causes a decrease in the tax on capital, which decreases the implicit tax on younger labor income. The decrease translates into younger agents working more under the optimal tax policy. The changes in the consumption and savings profiles are governed by the change in the after tax return to capital. Adopting the optimal tax policy has two counteracting effects on the marginal after tax return to capital. First, the pre-tax return to capital decreases. Second, the tax on capital income decreases, which increases the after-tax return. The first effect is consistent for all individuals. Since the baseline-fitted U.S. tax policy is progressive and the optimal tax policy is flat, the second effect is larger for individuals with higher income. Overall, the second effect dominates for most individuals and the marginal after-tax return to capital increases, causing the consumption profile to be steeper. The change in the consumption profile is more pronounced for middle-aged individuals with higher incomes since they experience the largest decrease in the tax on capital. Additionally, these higher income individuals respond to the larger drop in the tax on capital by holding more assets.

Since adopting the optimal tax policy in the endogenous model increases the tax on capital, the changes in the life cycle profiles are different compared to the changes in the exogenous model. The larger tax on capital implicitly taxes young labor income at a higher rate. Therefore, in response to adopting the optimal tax policy in the endogenous model, individuals shift hours worked from earlier to later years. In the endogenous model, adopting the optimal tax policy causes agents to change their labor supply not only on the intensive margin but also on the extensive margin: individuals choose to retire later. The higher tax on capital translates into a lower after-tax return to capital, which causes a slight flattening of the consumption profile. Due to the lower return to capital, agents choose to hold less savings under the optimal tax policy. However, since they choose to postpone retirement until after the normal retirement age, they receive a larger social security benefit than in the exogenous model. Therefore, the lower level of savings does not translate into less consumption after retirement.

6 Robustness of Results

In this section, I check the robustness of the results with respect to two of the modeling assumptions in the benchmark results. First, I examine if the relationship between endogenous retirement and optimal taxation changes when I include a more rigorous social security program that incorporates some additional features of the U.S. program. Second, instead of assuming that there is no within cohort heterogeneity, I test if the relationship changes when I include idiosyncratic shocks to agents' productivity. I sequentially add each of these features in the model so that I can individually assess their impact on the results.

6.1 Rigorous Social Security

There are two major differences between the simplified social security program in my benchmark model and the U.S. social security system. First, the actual social security system is progressive since the replacement rate is inversely related to past earnings. The marginal replacement rate has three cut-off points (also known as "bend points"). Thus, the benefits are calculated such that they replace 90% of earnings up to the first bend point, then replace 32% of subsequent earnings up to the second bend point, and 15% of subsequent earnings until the third bend point. After the third bend point benefits do not increase even if an individuals average earnings increase. In contrast, in the benchmark specification, social security benefits are assumed to be a constant fraction of an agent's average earnings over his lifetime. The second difference is how early and late retirement credits are calculated. Although the adjustments for early or late retirement in the U.S. social security program were determined to be actuarially fair on average, they are set at standard rates across all individuals. The adjustments under the current law are such that benefits are reduced $\frac{5}{9}$ of one percent for each month retired prior to the normal retirement age, up to 36 months, and $\frac{5}{12}$ of one percent per month for moths prior to 36 months. Additionally, the delayed retirement credit is set at 8 percent per annum. In contrast, in the benchmark model the adjustments to benefits for early and late retirement are calculated to be actuarially fair.

In order to see how these features affect the optimal tax policy, I set up a more rigorous social security system that incorporates both the progressivity of the benefit formula and the actual adjustments to benefits for early and late retirement in the U.S. Social Security program. I follow Huggett and Parra (2010) and calibrate the bend points of the progressive social security system such that they occur at 0.21, 1.29, and 2.42 times average earnings in the economy. Additionally, I recalibrate the model with this new program to match all the targets in Table 1.[28]

Tables 4 and 2 demonstrate that the optimal capital tax is lower in the models that incorporates the more rigorous social security program compared to the models with the more parsimonious social security program. In the models with the rigorous social security program, the marginal replacement rate is both more progressive and lower compared to the benchmark model. The social planner responds to the lower social security benefit by choosing a lower capital tax. This lower capital tax implicitly increases the annual after-tax return to savings which which helps to replace some of the lower post retirement income from smaller social security benefits (for more details see Peterman (2013)). Despite the lower optimal capital taxes, Table 4 demonstrates that including endogenous retirement still causes a large increase in the optimal capital tax compared to the model with exogenous retirement. The optimal tax on capital is approximately 45 percent larger or five and a half

[28]A list of the new calibration values is available upon request. One difference in this model is that regardless of the value set for χ_2, agents would not retire at 66 because the adjustments are not actuarially fair. Therefore, I calibrated the model so that agents chose to retire at 67.

percentage points larger when retirement is included endogenously. Therefore, including a more rigorous social security program does not materially change the relationship between endogenous retirement and optimal taxation.

Table 4: **Tax Policies (Rigorous Social Security)**

Tax Rate	**Exogenous**	**Endogenous**
τ_k	11.9%	17.3%
τ_h	24.5%	23.7%
$\frac{\tau_k}{\tau_h}$	0.49	0.73

6.2 Idiosyncratic Productivity Shocks

Next, I determine whether including a permanent, persistent, and transitory idiosyncratic shocks to labor income affects the results. I model the labor productivity process, $\omega_t h_t$, such that it follows the specification in Huggett and Parra (2010):

$$\log \omega_t = \varepsilon_j + \alpha + z_t + \nu_t. \tag{7}$$

In this specification ε_j governs the average age-profile of wages (which is included in the simple model), $\alpha \sim NID(0, \sigma_\alpha^2)$ is an individual-specific fixed effect (ability) that is observed at birth and stays fixed for an agent over the life cycle, $\nu_t \sim NID(0, \sigma_\nu^2)$ is a transitory shock to productivity received every period, and z_t is a persistent shock, also received each period, which follows a first-order autoregressive process:

$$z_t = \rho z_{t-1} + \theta_t \text{ with } \theta_t \sim NID(0, \sigma_\theta^2) \text{ and } z_1 = 0. \tag{8}$$

Using PSID data on the joint distribution of wages, hours and consumption, Kaplan (2012) estimates an incomplete markets life-cycle model with endogenous labor supply to match the evolution of both the first and second moments over the life cycle. In particular, the parameters are estimated for the cross-sectional variance-covariance patterns of hours, consumption and wages at different ages over the life cycle.[29].[30] Both the persistent and transitory idiosyncratic shocks to an individual's productivity are distributed normally with a mean of zero. The remaining parameters are also set in accordance with the estimates in Kaplan (2012): $\rho = 0.958$, $\sigma_\alpha^2 = 0.065$, $\sigma_\nu^2 = 0.017$ and $\sigma_\theta^2 = 0.081$. I discretize all three of the shocks in order to solve the model. I use two states to represent the transitory and permanent shocks and five states to represent the persistent shock. For expositional convenience, I refer to the two different states of the permanent shock as high and low ability types.

I calibrate the model which includes both the rigorous social security program and the idiosyncratic shocks to labor productivity using the same targets described in Table 1. A list of the full set of calibration parameters in this more rigorous model are included Table 5.

I begin by examining how well this rigorous model matches the U.S. economy. Figure 3 shows the fraction of retired households between ages 60 and 70, and compares it against estimates from the Panel Survey of Income Dynamics (PSID).[31] The model does not match the data with respect to three aspects. First, in the model, no agent is allowed to retire prior to age 62.[32] This restriction is at odds with the PSID data where households are allowed to retire at any age. Second, the model overstates the fraction of retirees for older agents since all agents are forced to retire by age 70 in the model. Third, the model understates the fraction of younger retirees relative to the data. This is because in the data and per the U.S. social security program, households reporting retirement and collecting social security benefits continue to be eligible to participate in labor market activities, whereas in

[29]For details on estimation of this process, see Appendix E in Kaplan (2012). Heathcote et al. (2010) estimated a similar process – see footnote in Huggett and Parra (2010)

[30]In order to stay consistent with the benchmark model, I use the same deterministic age profile for productivity (ε_j) from Kaplan (2012).

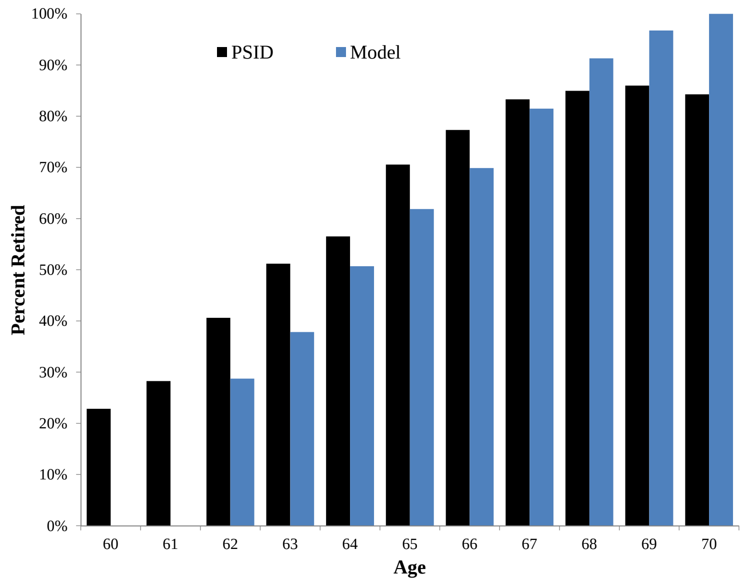
[31]I limit the analysis to male head of households in the 1967 - 1999 surveys. In the PSID individuals report the age at which they retire. I consider an individual retired if they are older than this age.

[32]Although agents are allowed to choose zero labor supply they cannot start claiming social security benefits prior to 62. I find that no agent chooses to work zero hours in the model prior to retirement.

Table 5: **Calibration Parameters**

Parameter	Value	Source
Demographics:		
Normal Retirement Age: j_{nr}	66	By Assumption
Max Age: J	100	By Assumption
Surv. Prob: Ψ_j		Bell and Miller (2002)
Pop. Growth: n	1.1%	Conesa et al. (2009)
Firm Parameters:		
α	.36	Data
δ	8.33%	$\frac{I}{Y} = 25.5\%$
A	1	Normalization
Preference Parameters:		
Conditional Discount: β	0.996	$\frac{K}{Y} = 2.7$
Risk aversion: γ	2	Conesa et al. (2009)
Frisch Elasticity: σ	0.5	Intensive Frisch= $\frac{1}{2}$
Disutility to Labor: χ_1	41.9	Avg. $h_j = \frac{1}{3}$
Fixed Cost to Working: χ_2	0.725	70% retire by J_{nr}
Productivity Parameters:		
Persistent Shock: σ^2_ν	0.017	Kaplan (2012)
Persistent: ρ	0.958	Kaplan (2012)
Permanent Shock: σ^2_α	0.065	Kaplan (2012)
Transitory Shock: σ^2_θ	0.081	Kaplan (2012)
Government Parameters:		
Υ_0	.258	Gouveia and Strauss (1994)
Υ_1	.768	Gouveia and Strauss (1994)
ϕ	17%	Conesa et al. (2009)
Social Security:		
κ_{1a}	6.7%	U.S. SS Program
κ_{1b}	5%	U.S. SS Program
κ_2	8%	U.S. SS Program
τ_{r1}	90%	U.S. SS Program
τ_{r2}	32%	U.S. SS Program
τ_{r3}	15%	U.S. SS Program
b_1	.21 x Avg Earnings	Huggett and Parra (2010)
b_2	1.29 x Avg Earnings	Huggett and Parra (2010)
b_3	2.42 x Avg Earnings	Huggett and Parra (2010)
τ_{ss}	10.0%	Market Clearing

Figure 3: **Fraction of Population Retired**



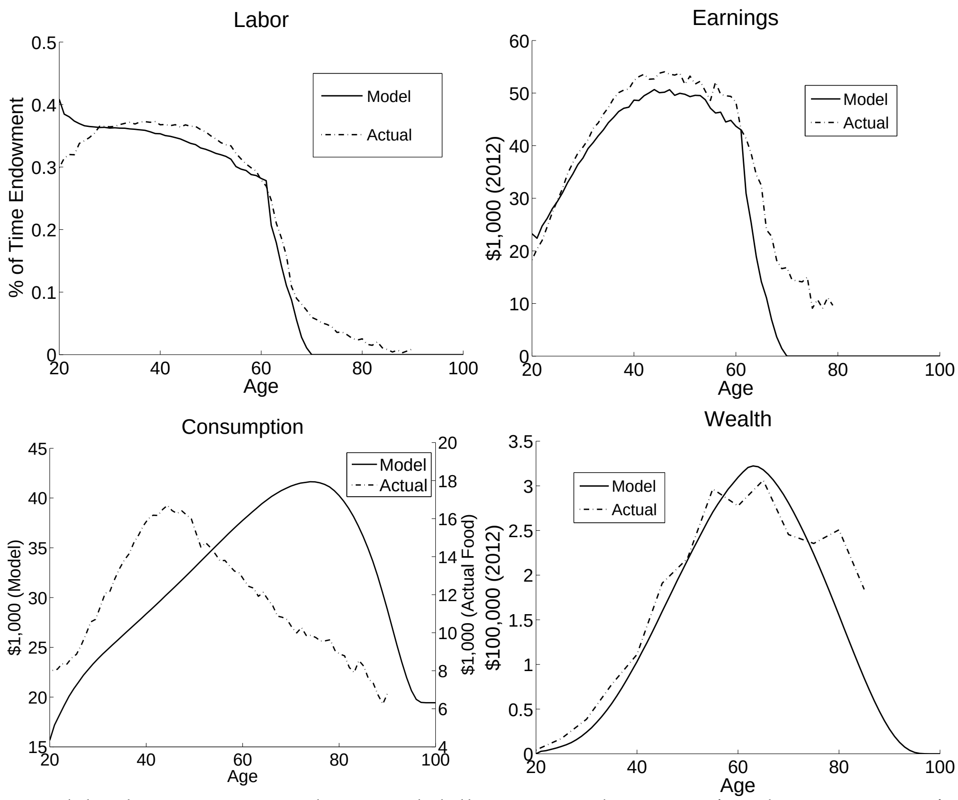
the model the decision to retire is associated with a complete and permanent exit from the labor market.

Next, I examine how well the model matches the observed life cycle profiles. Figure 4 plots the average labor, earnings, consumption, and savings profiles predicted by the more rigorous model and the profiles measured in the data.[33] This more rigorous model does a similar job matching the average profiles as the more parsimonious model in section 5. However, the more rigorous model allows me to examine the implications of the model for heterogenous individuals. Figure 5 plots the average life cycle profiles across all households and across the different skill types (high and low). The model generated average profiles of consumption and savings in Panels A and B follow expected patterns. Not surprisingly, the average consumption and savings of the high skill workers – who on average receive higher wages – are superposed to those of the low skill types.

Turning to panel C, in the model, the high skill workers spend less time in the labor market than their low skill counterparts. To compare this model prediction against empir-

[33]I still use the median level of wealth in the data as a comparison for savings.

Figure 4: **Life Cycle Profiles**



Note: The model values are converted to actual dollar amounts by comparing the average earnings in the model and the data. The actual labor, earnings, and consumption profiles are from the PSID, while the wealth profile is from the SCF.

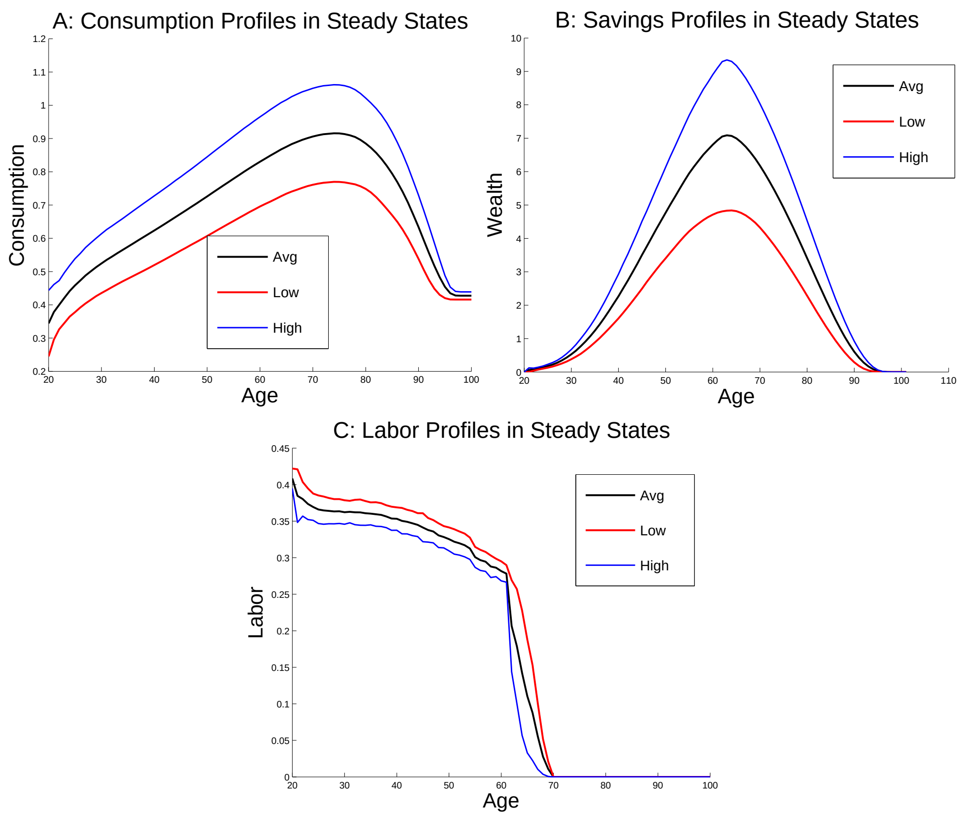
ical estimates, I use the PSID data to construct the age profile of labor hours for heads of households in the top and bottom deciles of the wage distribution. To assign each worker to a given decile of the distribution, I proceed in two steps. First, I use the reported labor earnings and labor hours across time to compute the average wage for the given household head in any given year of the working life cycle. Second, I weight the computed wage observations by hours worked in a given period to arrive with the working lifetime average. Figure 6 shows the estimated profiles for the 1st, 3rd, 7th and 10th deciles of the wage distribution. As in the model, my empirical estimates suggest that workers in the bottom deciles of the wage distribution spend more time working than the individuals with wages in the top tier of the distribution.[34]

Turning to the optimal tax policy, I find that in this model which includes within cohort heterogeneity, the optimal tax policy is no longer simply a flat labor and capital tax. Instead, the optimal labor tax is a progressive tax policy. Figure 7 plots the average and marginal labor income tax rates for the baseline tax policy and the optimal tax policies in the exogenous and endogenous models. In both models, the optimal labor income tax policy is essentially a flat tax on labor income with a small deduction, resulting in a progressive tax policy. In the exogenous model the optimal tax policy is essentially a labor tax of 32.8% with a fixed deduction of $11,700 and a flat capital tax of 10.2%. In the model with endogenous retirement the optimal tax policy is essentially a labor tax of 32.2% with a fixed deduction of $12,250 and a flat capital tax of 13.7%. Even though the government relies even less on a capital tax in this specification which includes within cohort heterogeneity, including endogenous retirement still cause an economically significant increase in the optimal capital tax.

Finally, I turn to the welfare effects of ignoring endogenous retirement. Specifically, I determine the welfare effect in the endogenous model of switching from the optimal tax, which includes a larger capital tax, to the optimal tax policy I solved for in the exogenous

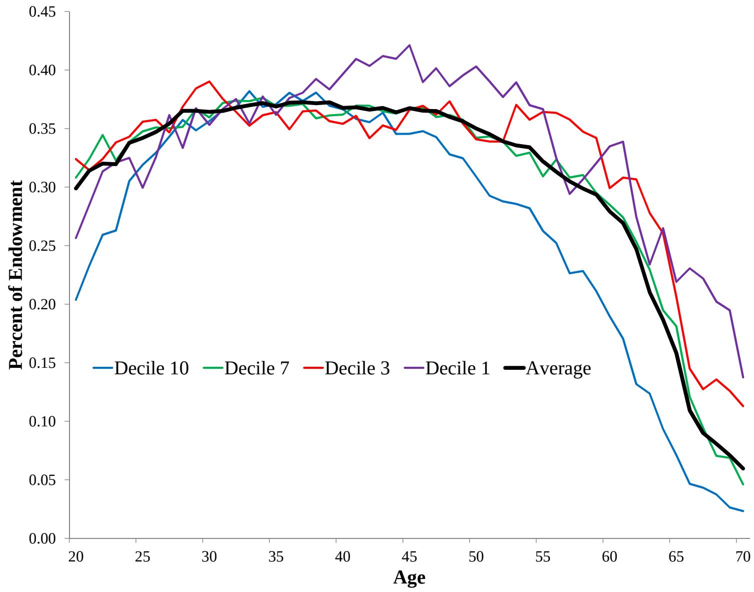
[34]Using the same data source but assigning workers into educational groups instead of wage deciles, Erosa et al. (2011) estimate that the annual labor supply of workers with only a high school diploma is lower than the labor supply of workers with a college degree.

Figure 5: **Steady State Life Cycle Profiles in Model with Rigorous Social Security**



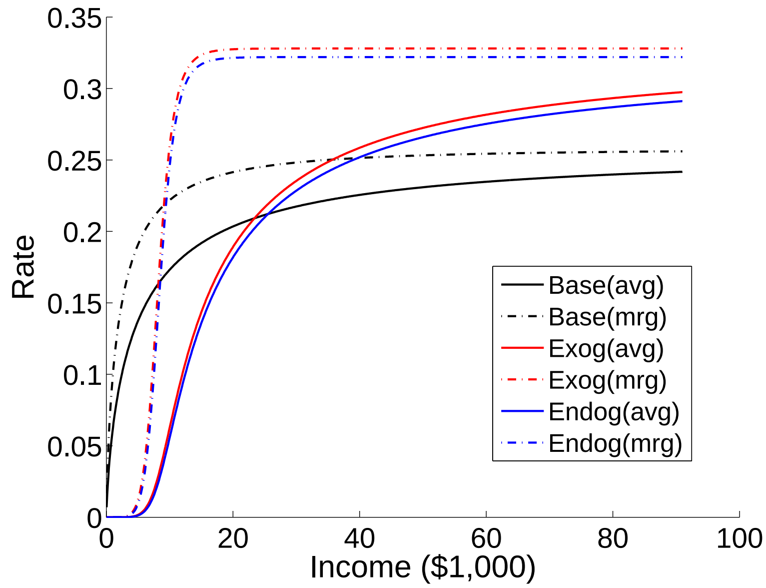
Note: These plots are the average values by age and ability type in all three models.

Figure 6: **Labor Profiles over the Lifecycle**



model, which includes a lower tax on capital. I find that the welfare effects of adopting the optimal tax policy determined in the other model are not as large in this more rigorous model. Overall, I find that adopting the optimal tax policy determined in the other model causes the total welfare in the endogenous model to decrease by the equivalent of approximately one-tenth of a percent of total lifetime consumption. The welfare consequences are not as large in this more rigorous model for two reasons. First, the overall importance of the tax on capital is smaller since the government relies less heavily on a capital tax to fund its expenditures. Second, in this more rigorous model, the idiosyncratic wage shocks leads to heterogenous effects of implementing the suboptimal tax policy. For example, adopting the lower than optimal tax on capital in the endogenous model causes welfare increases for the agents with higher permanent ability who tend to hold more savings. In contrast, the lower ability agents suffer average welfare losses due to the lower tax on capital and higher tax on labor; these losses are equivalent to two-tenths of a percent of their total lifetime consumption. These offsetting changes to welfare lead the suboptimal tax policy to have less of an average effect on total welfare.

Figure 7: **Tax Policies**



7 Conclusion

In this paper I computationally solve for the optimal capital and labor tax rates in separate life cycle models with exogenously and endogenously determined retirement. I find that including endogenous retirement causes a large increase in the optimal tax on capital. In the simple model with exogenous retirement the optimal tax policy is a 23.5 percent tax on labor income and a 17.9 percent tax on capital. In the simple model with endogenous retirement, the optimal tax policy is a 20.4 percent tax on labor income and a 30.6 percent tax on capital income. Relaxing the simplifying assumption that retirement is exogenously determined causes a seventy percent increase in the optimal tax on capital. Furthermore, I find that the welfare cost of adopting the lower optimal tax on capital from the exogenous model in the model with endogenous retirement, which calls for a higher tax on capital, is equivalent to 0.6 percent of total lifetime consumption. These result indicate that the simplifying assumption of exogenously determined retirement has notable consequences when solving for optimal tax policy. I find that the effect of endogenous retirement is somewhat smaller in the more rigorous model mainly owing to the inclusion of a social security program that mimics the program in the U.S., which has a lower average replacement rate.

With the smaller social security program, the government relies less heavily on a tax on capital so the welfare consequences of adopting the wrong tax on capital are smaller.

Including endogenously determined retirement causes the optimal tax on capital to increase because it affects the aggregate Frisch elasticity through two channels. First, it increases the aggregate responsiveness of hours to changes in the after-tax wage. Therefore, the government would prefer to rely more heavily on a capital tax as opposed to a labor tax in order to minimize the distortions induced by the tax policy. Second, including a retirement decision causes individuals to be relatively more responsive to changes in the after-tax wage when they are old. The government would like to condition labor income taxes on age to tax agents when they are older and more responsive at a relatively lower rate. Since age-dependent labor income taxes are unavailable, the government uses a higher tax on capital to mimic such a tax policy.

When modeling the economy, economists are constantly trying to balance realism and tractability. In this paper I demonstrate that the simplifying assumption of exogenously determined retirement has a sizable impact on optimal tax policy. Therefore, future work that examines optimal taxation in a life cycle model should incorporate endogenous retirement. Additionally, most of the previous work analyzing optimal taxation assumes that the social security program is outside of the control of the government. Given the large impacts of endogenous retirement on optimal tax policy it seems that there may be welfare gains from optimizing both the tax policy and the social security program together as opposed to optimizing them in isolation. For example, the results in this paper suggest that it might be optimal to restructure the social security program such that it provides further incentives for individuals to postpone retirement until the end of the retirement window.

References

Altonji, Joseph G., "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," *The Journal of Political Economy*, 1986, *94* (3), S176–S215.

Atkeson, Andrew, Varadarajan Chari, and Patrick Kehoe, "Taxing Capital Income: A Bad Idea," *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 1999, *23* (3), 3–17.

Bell, Felicitie and Michael Miller, "Life Tables for the United States Social Security Area 1900 - 2100," Actuarial Study 120, Office of the Chief Actuary, Social Security Administration 2002.

Chetty, Raj, "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," Working Paper 15616, NBER December 2009.

_ , Adam Guren, Day Manoli, and Andrea Weber, "Does Indivisible Labor Explain the Difference between micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities," NBER Working Paper, Harvard January 2011.

Clark, Kim B. and Lawrence H. Summers, "Demographic Differences in Cyclical Employment Variation," *Journal of Human Resources*, 1981, *16* (1), 61–79.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger, "Taxing Capital? Not a Bad Idea After All!," *American Economic Review*, 2009, *99* (1), 25–38.

Contreras, Juan and Sven Sinclair, "Labor Supply Response in Macroeconomic Models: Assessing the Empirical Validity of the Intertemporal Labor Supply Response from a Stochastic Overlapping Generations Model with Incomplete Markets," MPRA Paper 10533, University Library of Munich, Germany September 2008.

Domeij, David and Martin Floden, "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased," *Review of Economic Dynamics*, 2006, *9* (2), 242 – 262.

Erosa, Andrés and Martin Gervais, "Optimal Taxation in Life Cycle Economies," *Journal of Economic Theory*, 2002, *105*, 338–369.

_ , Luisa Fuster, and Gueorgui Kambourov, "Towards a micro-founded theory of aggregate labor supply," Working Papers 2011-13, Instituto Madrileï¿½o de Estudios Avanzados (IMDEA) Ciencias Sociales July 2011.

Fiorito, Riccardo and Giulio Zanella, "The Anatomy of the Aggregate Labor Supply Elasticity," *Review of Economic Dynamics*, April 2012, *15* (2), 171–187.

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

Garriga, Carlos, "Optimal Fiscal Policy in Overlapping Generations Models," Working Papers in Economics 66, Universitat de Barcelona. Espai de Recerca en Economia January 2001.

Gouveia, Miguel and Robert Strauss, "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis," *National Tax Journal*, June 1994, *47* (2), 317–339.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 2010, *118* (4), pp. 681–722.

Huggett, Mark and Juan Carlos Parra, "How Well Does the U.S. Social Insurance System Provide Social Insurance?," *Journal of Political Economy*, 2010, *118* (1), pp. 76–112.

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

Jacobs, Bas and A. Lans Bovenberg, "Human Capital and Optimal Positive Taxation of Capital Income," *International Tax and Public Finance*, 2009, pp. 1–28. 10.1007/s10797-009-9120-5.

Kaplan, Greg, "Inequality and the life cycle," *Quantitative Economics*, 2012, *3* (3), 471–525.

Kitao, Sagiri, "Sustainable Social Security: Four Options," Working Paper, Hunter College January 2012.

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

Meyer, William and William Reichenstein, "Social Security: When to Start Benefits and How to Minimize Longevity Risk," *Journal of Financial Planning*, 2010, *23* (3), 49–59.

Nakajima, Makoto, "Optimal Capital Income Taxation with Housing," Working Paper 10-11, Federal Reserve Bank of Philadelphia April 2010.

Peterman, William B., "Reconciling micro and macro estimates of the Frisch labor supply elasticity," Finance and Economics Discussion Series 2012-75, Board of Governors of the Federal Reserve System (U.S.) October 2012.

_ , "Determining the motives for a positive optimal tax on capital," *Journal of Economic Dynamics and Control*, 2013, *37* (1), 265–295.

Pistaferri, Luigi, "Anticipated and Unanticipated Wage Changes, Wage Risk, and Intertemporal Labor Supply," *Journal of Labor Economics*, 2003, *21* (3), 729–754.

Rawls, John, *A Theory of Justice*, Cambridge, Massachusetts: Harvard University Press, 1971.

Rettenmaier, Andrew J. and Thomas R. Saving, “How Generous Are Social Security and Medicare?,” Working Paper 290, National Center for Policy Analysis October 2006.

Rogerson, Richard and Johanna Wallenius, “Micro and Macro Elasticities in a Life Cycle Model with Taxes,” *Journal of Economic Theory*, 2009, *144* (6), 2277 – 2292. Dynamic General Equilibrium.