Age-Targeted Income Taxation, Labor Supply, and Retirement[*]

Johan Gustafsson[†]

March 1, 2021

Abstract

This paper studies the life-cycle effects of favorable marginal tax treatment of older workers on their optimal life cycle labor supply, retirement timing, and savings. I develop a structural model in continuous time where the life-cycle of a representative agent is divided into three distinct phases: pre-treatment, post-treatment, and retirement. Solutions for consumption/savings, labor supply/leisure, and retirement timing are then obtained by solving the model as a salvage value problem. I then calibrate the model to Swedish earnings data and find that the increased extensive margin labor supply is partially offset by a reduction in hours worked during the pre-treatment period. The total effect is however an increase in life-cycle labor supply and consumption.

Keywords: Retirement age, life cycle, tax heterogeneity, savings, consumption, leisure

JEL Codes: D15, J22, J26

1 Introduction

Age-targeted labor income tax treatment constitutes a small-scale policy reform aimed to induce longer working lives (e.g., Biggs, 2012; Laitner and Silverman, 2012; Alpert and Powell, 2013). By lowering labor income taxes for individuals on the verge of retirement, i.e., by modifying taxes based on age, and increasing the net returns to labor, the policymaker increases the cost of retirement. Such a policy was introduced in Sweden in 2007 as part of a general political ambition to increase overall labor supply. More specifically, the policy implied more generous earned income tax credits (EITC), and a reduced payroll tax, for workers older than 65. The tax reform thereby targeted both the labor supply of and labor demand for older workers.[1]

In this paper, I model the effects of age-targeted labor income tax credits on life-cycle labor supply, consumption/savings and retirement timing. For a Cobb-Douglas specification of preferences, I derive analytical expressions for life-cycle consumption and leisure as well as the condition for optimal retirement age conditional on the age-differentiated tax scheme. I then calibrate the model to match the average evolution of earnings over the life cycle based on evidence provided by Swedish register data. The representative life-cycle profile obtained from the calibration then serves as the baseline profile for numerical simulations of the tax policy.

The Swedish reform of age-targeted EITCs has received limited attention in the literature. Laun (2017) studies the labor supply response of older workers to the Swedish tax reform in a static labor supply model. The results indicate a positive effect on the extensive margin labor supply and suggest that the tax treatment promotes longer working lives and

---

[1]Beyond the Swedish context and in light of the ongoing demographics change toward aging populations, such policy consideration is highly relevant for policymakers in all economies with developed welfare systems given the general public opposition to deterministically increasing eligibility age and lowering benefit levels (Lacomba and Lagos, 2006).

a delayed labor force exit. Since the model is static however, the analysis is reduced to the immediate labor-leisure tradeoff facing older workers. The author discusses the possible dynamic effects introduced by policy uncertainty and anticipation effects of the non-treated, which, if not accounted for, risk overestimating the labor supply response to the tax treatment.

The model presented below recognizes the importance of modelling the labor supply responses on both the intensive and extensive margin when studying people's responses to age-differentiated changes in the payoff to labor. Conventional life-cycle theory predicts that a rational agent will account for changes in the opportunity cost of leisure over the whole working life when deciding their optimal labor/leisure and consumption/savings trajectories. In this context, favorable tax treatment of the elderly implies an increase in the post-retirement leisure cost, which encourages the individual to substitute labor supply from non-treated age to treated age. Ultimately, a static labor supply model fails to account for the life-cycle effects of the change in the opportunity cost of leisure since it is limited to the labor-leisure tradeoff at a particular instant in time.

This paper aims to provide explicit conditions for optimal life-cycle labor supply on both the intensive and extensive margin as a function of the age-differentiated tax scheme.[2] The model is easy to calibrate and able to rationalize the observed life-cycle dynamics of leisure and savings. As such, the paper also contributes to the area of quantitative macroeconomics where households in continuous-time general equilibrium models are often modeled to supply

---

[2]The empirical literature on lifetime labor supply response to tax-benefit schemes contains examples of ambitious model contributions, which have been shown to give a good fit to micro data. French (2005) estimates that an anticipated permanent wage increase of 20 % at age 60 leads to an increase in life-time labor supply by 1,906 hour increase in lifetime labor supply in total, but a 519 hour decrease in labor supply prior to treatment. This finding suggests that while total labor supply increases following the anticipated tax change, this effect is partially offset by increased leisure activity pre-treatment. Keane and Wasi (2016) show that the inclusion of human capital accumulation leads to increased estimates of Hicksian and Frisch elasticities for older individuals relative to younger. This suggests that older workers constitute a target group where marginal changes in the payoff to labor supply is expected to have relatively large effects on the labor-leisure tradeoff.

labor inelastically, and/or retire at a fixed age.

For this purpose, I build on the seminal contribution by Heckman (1974), which concerns the labor-leisure tradeoff of a representative agent in a partial equilibrium life-cycle environment. In his model, the agent decides on consumption and labor supply over the life cycle in a friction-free environment, governed by the dynamics of risk-free savings. The optimal labor supply profile follows from the trade-off between foregone labor income and a subsequent contraction of the income stream when spending more time in leisure activities. Intra-period utility is assumed non-separable in the arguments, thereby capturing the interdependence of consumption and earnings through complementarity in the utility function. The model is consistent with a hump-shaped consumption profile conditional on a hump-shaped labor productivity profile without the inclusion of credit constraints.[3]. A limitation of the model is its dependence on interior solutions for labor supply, i.e., Heckman only formally models life-cycle profiles assuming a non-binding participation constraint.

Gahramanov and Tang (2016) realize that the assumption of an interior labor supply profile in the benchmark Heckman specification is a shortcoming since the possibility of retirement creates a salient corner solution. They modify the Heckman model by restricting leisure to be interior to the time endowment for the working life and equal to the time endowment when retired. They then resort to a numerical solver to compute the optimal consumption and leisure trajectories, where retirement age coincides with the age at which the time endowment constraint binds.[4] While solving explicitly for the retirement age, the solution is conditional on that labor supply smoothly approaches zero. This in turn results in continuous trajectories of the optimal controls over the entire life-cycle domain, an ob-

---

[3]Bütler (2001) adds structure to the model by assuming that preferences admit to a Cobb-Douglas specification.

[4]The resulting life-cycle profiles are compared with the benchmark Heckman model where any infeasible leisure values are simply censored at the time endowment. They find that for some reasonable behavioral parameterization, the unconstrained Heckman model yields notably different solutions for labor supply profiles relative to their augmented model.

servation not necessarily consistent with the effects of age-targeted taxation, which imply discrete changes to the net labor income stream.

Another optimal control approach that has been used to solve for endogenous retirement timing treats the problem as a two-stage optimal control problem (Tomiyama, 1985), where the life cycle is decomposed into two subdomains: working life and retirement. Each subdomain is modeled as separate planning horizons for the control variables, where transversality conditions on the co-state variable and Hamiltonian functions ensure optimal savings behavior and retirement timing. Kuhn et al. (2015) solve for optimal retirement timing and health investments, but abstract from explicit solutions for intensive margin labor supply, Caliendo and Findley (2020) solve for extensive and intensive margin labor supply as separate cases while keeping the other margin fixed. In this paper, both are solved for simultaneously.

Life-cycle labor supply including both the intensive and extensive margin has also been examined in Mao et al. (2014), and revisited in Mao et al. (2019). Mao et al. (2014) analyze the implications of a defined benefit social security scheme, and Mao et al. (2019) consider the possibility of part-time retirement. While their model solves for optimal consumption, leisure, and retirement timing, it treats initial consumption as a parameter, thereby implicitly imputing a value for the intrinsic marginal utility of wealth. The model dynamics are therefore stylized as the optimal control of consumption is not fully solved from the utility maximization problem.

To fully integrate the implications of age-targeted taxation in terms of the model dynamics, I augment the Heckman model with two regime shifts. One regime following tax treatment, which alters the net labor payoff and subsequently the dynamics of assets at an exogenously determined point in time. The second regime concerns the extensive margin labor supply, i.e., the timing of retirement. As a result, the lifecycle can be decomposed into

three distinct subdomains: the working life pre- and post-treatment, and retirement. I solve for the extensive margin labor supply following a free terminal time salvage value principle[5] (see e.g., Hartl and Sethi, 1983; Seierstad and Sydsaeter, 1987). That is, the labor market exit is determined from the optimality condition that the agent traverses into retirement when the marginal utility of extending the working life is equal to the marginal utility loss of less time spent in retirement.

The tax treatment exogenously imposes a discrete change in the asset accumulation function, both indicating the terminus of the first control horizon and the start of the second. I therefore model the traverse between these subdomain 1 and 2 as a fixed terminal time salvage value. The agent's savings behavior is such that the marginal utility of wealth at the end of the pre-treatment subdomain is equal to the marginal utility of wealth at the start of the treatment subdomain. Ultimately, the salvage value transversality conditions constrain the savings behavior over each subdomain of the agent's life such that utility is maximized over the entire life-cycle domain. These conditions are in essence a continuous time analogue to the principles of optimality in discrete time dynamic programming, and an extension of Tomiyama's two-stage control problem into a three-stage control problem. The present paper thereby illustrates how to conduct the analysis of multiple stage optimal control problems via the salvage value principle. Furthermore, is adds to the literature on age-tagged taxation by including both margins of labor supply with closed form solutions for all decision margins of the agent. To the best of my knowledge, previous literature does not provide such complete characterization within a theoretically tractable framework.

I derive optimal control solutions for life-cycle leisure and consumption profiles in a fixed factor price setting[6], while considering both margins of labor supply. The model assumes

---

[5]A free terminal time salvage value refers to a control problem where the terminus of the control horizon is a decision variable. In this context, the terminus relates to the end of the working life. In a fixed terminal time salvage value problem, the terminus is exogenously imposed.

[6]Since the institutional setting concerns Sweden, which is a small and open economy, fixed factor prices

an interdependence, or complementarity, of consumption and leisure to obtain a realistic co-movement between earnings and consumption. For mathematical convenience, I do not consider unemployment spells or other risks associated with the labor market. In addition, retirement is treated as an absorbing state. That is, the decision to terminate the second sub-domain is modeled as an irreversible choice of the representative agent. In future research, it would be interesting to extend the analysis to allow for additional switches following various labor market frictions.

Calibrating the model to the Swedish earnings data suggests that favorable marginal tax treatment of older workers delays retirement and increases life-cycle labor supply. However, the leisure intensity increases during the primary working life as a consequence of leisure becoming relatively less expensive. The net effect is still an increase in total working hours over the whole life cycle. This finding is robust to various reasonable behavioral parameter configurations, as well as a Strulik and Trimborn (2018)-type hyperbolic discounting function. I explore such scenarios numerically. While the main analysis concerns a self-financing agent, I also consider the effect of a defined contribution-type social security scheme in a sensitivity analysis, which introduces a notion of forced savings in a less beneficial transfer technology than a risk-free credit account with a strictly positive return.

The outline of the paper is as follows: the life-cycle model is introduced in Section 2, complete with the specification of preferences and the principles of optimality following the salvage value principle. Section 3 contains the calibration and the numerical analyses. Section 4 concludes the paper.

are considered a reasonable assumption.

6

2 The model

2.1 Model setup

Time is continuous and denoted by t. Consider a representative agent who enters the work-force at $t = 0$, with a maximum life length of $t = T > 0$, defining the start and terminus of the representative life-cycle domain. Further, let $R \in (0, T)$ represent the timing of tax treatment and $F \in [R, T]$ the timing of retirement. The probability of surviving to a certain age $t \leq T$ follows a survival polynomial, $\mathbb{P}(t)$, which the agent incorporates in the decision-problem.

While employed, the individual earns a market-determined efficiency wage, $w(t)$, per unit of labor. Earnings are subject to a labor income tax that varies over time, $\tau_1 \geq 0 \ \forall \ t \in [0, R)$ and $\tau_2 \geq 0 \ \forall \ t \in [R, F)$, with $\tau_1 \geq \tau_2$. The individual derives utility from the consumption of non-durable goods and services, $c(t)$, and from leisure, $h(t)$. Any savings grow by the time-invariant risk-free interest rate, $r \geq 0$, and flow into the individual asset account, $k(t)$. I assume perfect credit markets. Let $k(0) = k_0$ denote initial wealth and $k(T) = k_T$ terminal wealth. The agent is endowed with 1 unit of time at each instant in time, which they divide between labor and leisure. Further let the intrinsic rate of time preferences be denoted by $\theta \geq 0$. To avoid notational clutter, let $\Theta(t) = \mathbb{P}(t)e^{-\theta t}$ represent the effective discount factor, where $\mathbb{P}(t)$ is the survival probability.[7]

Consider an exogenously determined point in time, $t = R \in (0, T)$, the date after which the individual is eligible for labor income tax credits. I assume that there exists an interior point in the post-treatment domain, $F \in (R, T)$, after which the individual makes

---

[7]In models where mortality is endogenous to health investments made by the individual, the performance index should include a baseline utility to ensure that flow-utility is always non-negative. This consideration follows from Hall and Jones (2007). If utility is negative, it would be welfare improving for the agent to end their life, which implies that health capital becomes a bad. However, since mortality is exogenous to the model, I abstract from its inclusion.

the irreversible choice of retiring and consuming out of accumulated savings.[8] The life-cycle domain is thus decomposed into three subdomains: primary working life, retirement eligibility, and full-time retirement. The maximization problem is then augmented to include the choice of retirement age, F. By construction, both controls, $c(t)$ and $h(t)$, are piecewise continuous throughout the subdomains separated by retirement eligibility, R, and full-time retirement, F. $k(t)$ is continuous over the whole lifecycle.

2.2   The life-cycle maximization problem

The representative agent maximizes lifetime utility, subject to its budget and time endowment constraints.

$$\mathcal{W}(t) \triangleq \max_{c(t),h(t),F} \left\{ \int_0^R \Theta(t)\{u(c(t),h(t))\}dt + \right.$$

$$\left. \int_R^F \Theta(t)\{u(c(t),h(t))\}dt + \int_F^T \Theta(t)\{u(c(t),1)\}dt \right\}$$

Subject to:

$$\dot{k} = \begin{cases} [1-h(t)](1-\tau_1)w(t) + rk(t) - c(t) & \text{for } t \in [0,R), \\ [1-h(t)](1-\tau_2)w(t) + rk(t) - c(t) & \text{for } t \in [R,F), \\ rk(t) - c(t) & \text{for } t \in [F,T]; \end{cases}$$

$$k(0) = k_0,$$

$$k(T) = k_T;$$

$$h(t) \begin{cases} \leq 1 & \text{for } t \in [0,F), \\ = 1 & \text{for } t \in [F,T]; \end{cases}$$

The optimization problem over each additive subdomain of the life cycle, constrained by the corresponding asset accumulation function, can be represented by a Hamiltonian function

---

[8]That is, an interior point of the life-cycle domain from which $h^*(t) = 1$. The notion of an interior or exterior optimum is addressed in Section 2.3.

from which optimal controls are derived for each planning horizon by the maximum principle. Let the Lagrangean[9] and the Hamiltonians be defined as follows:

$$\mathcal{L}_1(t) \triangleq \Theta(t)u(c(t), h(t)) + \mu_1(t)\dot{k} - \lambda(t)\{h(t) - 1\}, t \in [0, R),$$
$$\mathcal{H}_2(t) \triangleq \Theta(t)u(c(t), h(t)) + \mu_2(t)\dot{k}, t \in [R, F), \tag{1}$$
$$\mathcal{H}_3(t) \triangleq \Theta(t)u(c(t), 1) + \mu_3(t)\dot{k}, t \in [F, T).$$

The time-endowment constraint is explicitly modeled by a complementary slackness condition for the first subdomain, following the specification in Gahramanov and Tang (2016), and by a free terminal time salvage value principle for the residual of the life cycle. That is, if the leisure constraint binds during the first subdomain, the agent can return to the labor market during the second subdomain and then permanently retire when traversing into the third subdomain.[10]

2.3   Salvage value principle for optimality

Since optimization treats each subdomain as separate planning horizons, I need to consider the optimality conditions for the transition between subdomains and the timing of labor market exit. By construction, any optimal control derived for each consecutive subdomain can then be viewed as an optimal response to the level of accumulated savings at the start of the relevant subdomain. In this section, I present the transversality conditions for optimal transitions over both the exogenous and endogenously defined switching points. The transversality conditions together with the boundary conditions for the asset account form the conditions for the subdomain controls to be optimal over the entirety of the life-cycle domain and not only for the separate planning horizons. I apply the salvage value principle

---

[9]Formally accounting for inequality constraints (mixed constraints) on control variables by a Kuhn-Tucker condition results in a Lagrangean (or *Generalized Hamiltonian*) function. The optimal controls are derived by using the maximum principle (Seierstad and Sydsaeter, 1987).

[10]The life cycle can in such case be thought of as four subdomains. While the existence of an interior solution final retirement is assumed to always exist, the case when the time constraint binds pre-treatment is not.

to the optimal control problem introduced in Section 2.2. Through the resulting transversality conditions on retirement timing and assets saved for each switching point, I link the optimal controls for each subdomain.

Following the salvage value principle outlined in Hartl and Sethi (1983) and Seierstad and Sydsaeter (1987), and recalling that the life-cycle utility is the sum of the utility attached to each life-cycle subdomain, I reformulate the performance index of the agent as follows:

$$\int_0^R \Theta(t)u(c(t), h(t))dt + S_1(k(R), S_2(F, k(F))), \tag{2}$$

where $S_1(k(R), S_2(F, k(F)))$ denotes the salvage value of assets saved for the terminus of the pre-treatment subdomain and $S_2(F, k(F))$ represents the salvage value of assets saved for full-time retirement. The salvage value for the pre-treatment subdomain is given by:

$$S_1 = S_1(k(R), S_2(F, k(F))) = \int_R^F \Theta(t)u(c(t), h(t))dt + S_2(F, k(F)). \tag{3}$$

The life-cycle utility maximization problem can subsequently be solved using backward induction. The optimal consumption over the full-time retirement phase will be conditional on full-time leisure, and a function of assets saved for retirement. Hence, consumption during the full-time retirement phase can be viewed as an optimal response function of wealth at the time of retirement, $t = F$. Similarly, the consumption and leisure controls during $t \in [R, F)$ will be functions of assets held at time $t = R$. The agent realizes these optimal response functions when deciding on the labor supply and savings behavior during the pre-treatment phase. Assuming that the controls are optimal over each subdomain, I proceed by deriving explicit expressions for the transversality conditions linking each sub-domain.

The decision associated with full-time retirement includes two choices: one concerns the level of assets saved for retirement, i.e., a terminal condition on assets for the second sub-

domain, and the other concerns the timing of retirement. The salvage value at the time of full-time retirement can be written as:

$$S_2 = S_2(F, k(F)) = \int_F^T \Theta(t)u(c(t), 1)dt. \tag{4}$$

Equation (4) can in turn be rewritten as a function of the Hamiltonian and the intertemporal budget constraint adhering to the optimal control problem of the third subdomain as specified in equation (1). The resulting expression becomes:

$$S_2 = \int_F^T \{\mathcal{H}_3(t, c(t), \mu_3(t)) - \mu_3(t)\dot{k}_3\}dt = S_2, \tag{5}$$

where $\mu_3(t) = \mu(t)|_{t \in [F,T]}$, and $\dot{k}_3 = \dot{k}|_{t \in [F,T]}$. By integration-by-parts, equation (5) can be expressed as:

$$S_2 = \int_F^T \{\mathcal{H}_3(t, c(t), \mu_3(t)) + \dot{\mu}_3(t)k(t)\}dt + \mu_3(F)k(F) - \mu_3(T)k(T). \tag{6}$$

Using the above expression for the salvage value of assets saved for retirement, I state the optimal control horizon of subdomain 2, which in turn can be viewed as a salvage value of assets saved for the terminus of subdomain 1. Therefore, denote the performance index of subdomain 2 as follows:

$$S_1 = \int_R^F \{\mathcal{H}_2(t, c(t), h(t), \mu_3(t)) + \dot{\mu}_2(t)k(t)\}dt + \mu_2(R)k(R) - \mu_2(F)k(F) + S_2 \tag{7}$$

For the allocation of assets from subdomain 1 to subdomain 2 to be optimal, a marginal change in assets at the timing of retirement has to be 0,

$$\frac{\partial S_1}{\partial k(F)} = -\mu_2(F) + \mu_3(F) = 0. \tag{8}$$

Equation (8) implies that the optimal amount of assets saved for retirement satisfies the

following transversality condition:

$$\mu_2(F) = \lim_{t \to F} \mu(t)|_{t \in [R, F)} \overset{\Delta}{=} \mu_3(F). \tag{9}$$

Since the retirement timing problem essentially can be thought of as a free terminal time salvage value problem, I next derive the transversality condition for optimal retirement timing, conditional on that $k_2(F) = k_3(F)$.

$$\frac{\partial S_1}{\partial F} = \mathcal{H}_2(F) - \mathcal{H}_3(F) \begin{cases} \leq 0 & \text{if } F^* = R \\ = 0 & \text{if } F^* \in (R, T) \\ \geq 0 & \text{if } F^* = T \end{cases} \tag{10}$$

Recall that the agent can withdraw from the labor market before receiving tax treatment, in which case $\lambda(t) > 0$ for $t = t^* \in [0, R)$. If $F^* \in (R, T)$, independent of any possible switching point in previous subdomains, equation (10) simplifies to:

$$\mathcal{H}_2(F) = \mathcal{H}_3(F). \tag{11}$$

The optimal switching time is based on the condition that the marginal utility of expanding the second subdomain is equal to the marginal utility loss of contracting the third subdomain control horizon.

Since the tax treatment is exogenously imposed at time R, and thus not a decision variable of the agent, only the analogue of equation (9) needs to be satisfied. That is:

$$\mu_1(R) = \lim_{t \to R} \mu(t)|_{t \in [0, R)} \overset{\Delta}{=} \mu_2(R), \tag{12}$$

which together with equation (9) implies that the marginal utility of assets is continuous

over the entire life-cycle domain. Ultimately, the behavior of the agent is consistent with conventional preferences for smoothing the marginal utility of wealth over the life cycle.

2.4 Preferences

To study the implication of interdependence of consumption and leisure choices for the effect of tax treatment, I consider a Cobb-Douglas specification of the instantaneous utility function (felicity) as follows:

$$u(c(t), h(t)) = \frac{[c(t)^\phi h(t)^{1-\phi}]^{1-\sigma}}{1-\sigma}, \tag{13}$$

with $\lim_{\sigma \to 1} \frac{[c(t)^\phi h(t)^{1-\phi}]^{1-\sigma}}{1-\sigma} = \phi ln(c(t)) + (1-\phi)ln(h(t))$. Let $\phi \in (0, 1)$ and $\sigma > 0$. $1/\sigma$ is the elasticity of intertemporal substitution with respect to the consumption-leisure composite good, and ϕ is the relative utility weight for consumption. Conditional on an interior solution for labor supply, this functional form is consistent with the often observed co-movement of consumption and earnings.[11] For the consumption hump profile to be robust to the inclusion of retirement, the net of the efficient discount factor and interest rate must be positive during the retirement phase to assure a decline in consumption at old age.

2.5 Optimal controls

As mentioned in Section 2.3, I treat each subdomain of the life cycle as a separate planning problem for which I derive optimal controls of consumption, leisure, and the law of motion governing the savings dynamics. I then implement the transversality conditions from the salvage value principle to link the optimal controls and to solve the life-cycle utility

---

[11]The Cobb-Douglas specification is also known to be consistent with balanced growth, making it popular within the macroeconomics literature. It is also used by Bütler (2001) to illustrate structural solutions for the Heckman (1974) model.

maximization model. Recall the Lagrangean and the Hamiltonians from equations (1):

$$\mathcal{L}_1(t) = \Theta(t)\left\{\frac{[c(t)^\phi h(t)^{1-\phi}]^{1-\sigma}}{1-\sigma}\right\} + \mu_1(t)\{[1-h(t)](1-\tau_1)w(t) + rk(t) - c(t)\}$$
$$-\lambda(t)[h(t) - 1] \tag{14}$$

$$\mathcal{H}_2(t) = \Theta(t)\left\{\frac{[c(t)^\phi h(t)^{1-\phi}]^{1-\sigma}}{1-\sigma}\right\} + \mu_2(t)\{[1-h(t)](1-\tau_2)w(t) + rk(t) - c(t)\} \tag{15}$$

$$\mathcal{H}_3(t) = \Theta(t)\left\{\frac{[c(t)^{\phi(1-\sigma)}}{1-\sigma}\right\} + \mu_3(t)\{rk(t) - c(t)\} \tag{16}$$

I solve for the optimal controls of each subdomain by using the maximum principle. For subdomain 1, i.e., for $t \in [0, R)$, I express the first-order conditions as follows:

$$\frac{d\mathcal{L}_1(t)}{dc(t)} = \phi\Theta(t)c(t)^{\phi(1-\sigma)-1}h(t)^{(1-\phi)(1-\sigma)} - \mu_1(t) = 0 \tag{17}$$

$$\frac{d\mathcal{L}_1(t)}{dh(t)} = (1-\phi)\Theta(t)c(t)^{\phi(1-\sigma)}h(t)^{\phi(\sigma-1)-\sigma} - \mu_1(t)w(t)(1-\tau_1) - \lambda(t) = 0 \tag{18}$$

$$\dot{\mu}_1(t) = -r\mu_1(t). \tag{19}$$

For now, I abstract from multiple labor market exits, i.e., I assume $\lambda(t) = 0 \ \forall t \in [0, R]$. The optimal controls for subdomains 2 and 3 are expressed as follows:

$$\frac{d\mathcal{H}_2(t)}{dc(t)} = \phi\Theta(t)c(t)^{\phi(1-\sigma)-1}h(t)^{(1-\phi)(1-\sigma)} - \mu_2(t) = 0 \tag{20}$$

$$\frac{d\mathcal{H}_2(t)}{dh(t)} = (1-\phi)\Theta(t)c(t)^{\phi(1-\sigma)}h(t)^{\phi(\sigma-1)-\sigma} - \mu_1(t)w(t)(1-\tau_2) = 0 \tag{21}$$

$$\dot{\mu}_2(t) = -r\mu_2(t). \tag{22}$$

$$\frac{d\mathcal{H}_3(t)}{dc(t)} = \phi\Theta(t)c(t)^{\phi(1-\sigma)-1} - \mu_3(t) = 0 \tag{23}$$

$$\dot{\mu}_3(t) = -r\mu_3(t). \tag{24}$$

Since the agent saves in a risk-free asset throughout the entire life-cycle domain, the law of motion governing the marginal utility of wealth remains unchanged throughout the life cycle (see equations [19], [22] and [24]). By implementing the transversality conditions in equations (9) and (12), which ensure continuity of the co-state variable, I can solve for the law of motion over the entire life-cycle domain,

$$\mu(t) = \mu(0)e^{-rt}. \tag{25}$$

The marginal utility of money at time zero, $\mu(0)$, is an unknown constant to be solved for. Solving equation (17), (18), (19), (20), (21), and (23) and then using equation (25) gives the optimal controls of consumption and leisure for each subdomain.

$$c^*(t) = \begin{cases} \left[\dfrac{\mu(0)e^{-rt}}{\phi\Theta(t)\left[\frac{(1-\phi)}{\phi w(t)(1-\tau_1)}\right]^{(1-\phi)(1-\sigma)}} \right]^{-\frac{1}{\sigma}} & \text{for } t \in [0,R), \\[3em] \left[\dfrac{\mu(0)e^{-rt}}{\phi\Theta(t)\left[\frac{(1-\phi)}{\phi w(t)(1-\tau_2)}\right]^{(1-\phi)(1-\sigma)}} \right]^{-\frac{1}{\sigma}} & \text{for } t \in [R,F), \\[3em] \left[\dfrac{\mu(0)e^{-rt}}{\phi\Theta(t)} \right]^{\frac{1}{\phi(1-\sigma)-1}} & \text{for } t \in [F,T], \end{cases} \tag{26}$$

$$h^*(t) = \begin{cases} c^*(t) * \left[\dfrac{(1-\phi)}{\phi(1-\tau_1)w(t)} \right] & \text{for } t \in [0,R), \\[2em] c^*(t) * \left[\dfrac{(1-\phi)}{\phi(1-\tau_2)w(t)} \right] & \text{for } t \in [R,F), \\[2em] 1 & \text{for } t \in [F,T]. \end{cases} \tag{27}$$

So far, the subdomain optimal controls solutions are interdependent intertemporally through the law of motion, but since the retirement timing is endogenous, solving for the optimal terminus of the second subdomain remains. Recall from equation (11) that the transversality condition ensuring optimality in the free terminal time salvage value principle implies that the marginal utility of increasing the second subdomain is equal to the marginal utility loss of contracting the third subdomain. That is, the following equation must be satisfied for the

15

optimal retirement age to be interior to the time domain bounded by tax treatment and the maximum length of the life cycle, $F^* \in (R, T)$:

$$\Theta(F^*) \left\{ \frac{[c(F^*)^\phi h(F^*)^{1-\phi}]^{1-\sigma}}{1-\sigma} \right\} +$$

$$\mu(0)e^{-rF^*} \{[1 - h(F^*)](1 - \tau_2)w(F^*) + rA(F^*) - c(F^*)\} \tag{28}$$

$$= \Theta(F^*) \left\{ \frac{[c(F^*)^{\phi(1-\sigma)}}{1-\sigma} \right\} + \mu(0)e^{-rF^*} \{rA(F^*) - c(F^*)\}.$$

The next step would be to solve for the two unknowns F^* and $\mu(0)$ from the simultaneous equation system constituted by equation (28) and the life-cycle budget constraint. However, in the specification without social security, the dynamics of savings remain unchanged with the exception that the labor supply is positive in the second subdomain and zero in the third. From equation (28) we can observe that the RHS is equal to the LHS when the leisure intensity at the terminus of the second subdomain is equal to the time endowment constraint, i.e. $h(F^*) = 1$. Given the piece-wise continuous property of the optimal control profiles, leisure will therefore gradually trend to zero. If the optimal leisure profile only contains one exit from the labor market as an interior solution to subdomain 2, I can solve for $\mu(0)$ from the optimal control of leisure for the second subdomain expressed in equation (27) by substituting in $h(F^*) = 1$ and rearranging the terms:

$$\mu(0) = \left\{ \frac{(1 - \tau_2)w(F^*)e^{-rF^*}}{(1 - \phi)\Theta(F^*)\left[\frac{e^{-rF^*}}{\phi\Theta(F^*)}\right]^{\frac{\phi(1-\sigma)}{\phi(1-\sigma)-1}}} \right\}^{\phi(1-\sigma)-1}. \tag{29}$$

If on the other hand there exists a period in time, t^*, interior to the first subdomain where the leisure constraint also binds, i.e. $\lambda(t^*) > 0$, the intrinsic marginal utility of wealth satisfies the following two equations:

$$\mu(0) = \left\{ \frac{(1 - \tau_2)w(F^*)e^{-rF^*}}{(1 - \phi)\Theta(F^*)\left[\frac{e^{-rF^*}}{\phi\Theta(F^*)}\right]^{\frac{\phi(1-\sigma)}{\phi(1-\sigma)-1}}} \right\}^{\phi(1-\sigma)-1} = \left\{ \frac{(1 - \tau_1)w(t^*)e^{-rt^*}}{(1 - \phi)\Theta(t^*)\left[\frac{e^{-rt^*}}{\phi\Theta(t^*)}\right]^{\frac{\phi(1-\sigma)}{\phi(1-\sigma)-1}}} \right\}^{\phi(1-\sigma)-1}. \tag{30}$$

Lastly, utilizing the boundary conditions on the asset account, optimal controls must be operated such that the following condition is satisfied if only exiting the labor market once:

$$
k_0 + \int_0^R \{[1 - h^*(t)]w(t)(1 - \tau_1)\}e^{-rt}dt + \int_R^{F^*} \{[1 - h^*(t)]w(t)(1 - \tau_2)\}e^{-rt}dt
$$
$$
= \int_0^T \{c^*(t)e^{-rt}\}dt + k_T e^{-rT}, \tag{31}
$$

or if the agent exits the labor market prior to tax treatment and then re-enters following treatment:

$$
k_0 + \int_0^{t^*} \{[1 - h^*(t)]w(t)(1 - \tau_1)\}e^{-rt}dt + \int_R^{F^*} \{[1 - h^*(t)]w(t)(1 - \tau_2)\}e^{-rt}dt
$$
$$
= \int_0^T \{c^*(t)e^{-rt}\}dt + k_T e^{-rT}. \tag{32}
$$

Equations (31) and (32) must be satisfied for lifetime income to be equal to lifetime consumption. That is, the solutions to retirement timing (and pre-treatment labor market withdrawal as in the case of equation [32]) must satisfy the life-time budget constraint. If the wage profile and the survival probability function have a simple structure, it is possible to restate equation (31) and (32) as tractable expressions. For example, setting $w(t) = w$ and $\mathbb{P}(t) = 1$, assuming one labor market exit interior to $t \in (R, T)$, one can obtain the functional primitives to rewrite equation (31) into the following analytic expression:

$$
k_0 - \frac{w}{r}\left[(1 - \tau_1)[e^{-rR} - 1] + (1 - \tau_2)[e^{-rF^*} - e^{-rR}]\right] +
$$
$$
\frac{\sigma}{\phi\theta}\left\{\left[\frac{\mu(0)}{\phi\left[\frac{(1-\phi)}{\phi(1-\tau_1)w}\right]^{(1-\sigma)(1-\phi)}}\right]^{-\frac{1}{\sigma}}[e^{-\frac{\theta}{\sigma}R} - 1] + \left[\frac{\mu(0)}{\phi\left[\frac{(1-\phi)}{\phi(1-\tau_2)w}\right]^{(1-\sigma)(1-\phi)}}\right]^{-\frac{1}{\sigma}}[e^{-\frac{\theta}{\sigma}F^*} - e^{-\frac{\theta}{\sigma}R}]\right\}
$$
$$
-\frac{\phi(1-\sigma) - 1}{r\phi(\sigma - 1) - \theta}\left[\frac{\mu(0)}{\phi}\right]^{\frac{1}{\phi(1-\sigma)-1}}\left[e^{\frac{\theta-r}{\phi(1-\sigma)-1}T} - e^{\frac{\theta-r}{\phi(1-\sigma)-1}F^*}\right] - k_T e^{-rT} = 0, \tag{33}
$$

17

where

$$\mu(0) = \left\{ \frac{(1-\tau_2)we^{(\theta-r)F^*}}{(1-\phi)\left[\frac{e^{(\theta-r)F^*}}{\phi}\right]^{\frac{\phi(1-\sigma)}{\phi(1-\sigma)-1}}} \right\}^{\phi(1-\sigma)-1}. \tag{34}$$

For the calibration of the model and numerical solution, including more complex assumptions of the functional form of survival and wage, I turn to a numerical solver.

3 Numerical analysis

In this section I begin by calibrating the model to Swedish earnings and survival data to obtain a per-unit of efficient labor wage and a survival polynomial that are able to predict some key features of life-cycle dynamics. It should be noted that the paper makes a general assessment of the intertemporal effects of age-targeted income taxation. The calibration mainly fills the purpose of obtaining a wage profile that captures declining marginal productivity such that retirement becomes an interior solution to the life-cycle domain, as well as an effective discount factor for simulations. The calibrated life-cycle profile is then used to quantify the effects of the tax treatment without being specific to the Swedish institutional setting.

In the subsequent subsection, I perform a number of modeling experiments to complement the main analysis by (1) conducting a sensitivity analysis with regard to the behavioral parameterization and (2) introducing policy uncertainty in two ways. Additional experiments have been relegated to the appendix. These include the introduction of a hyperbolic discounting factor as specified in Strulik and Trimborn (2018) and the consideration of heterogeneity in inherited and bequeathed wealth.

In the final subsection, I augment the model with a mandatory defined contribution social security scheme with a stylized contribution-benefit formula typical of modern public

pension schemes.[12]

3.1 Calibration

Agents are modeled from age 25, which denotes labor market entry, and onward. The maximum lifetime length is 85, which implies that $T = 60$. The interest rate is fixed at $r = 3.5\%$ which is standard in the literature. I assume that the agent neither inherits nor bequeaths any wealth, i.e., $k_0 = k_T = 0$. I calibrate the model for the simulated life-cycle profile to be compatible with some key observable life-cycle facts: (i) Consumption peaks around age 45–55 (around the same age as earnings peak). (ii) The agent works on average around 30–35 hours a week.[13] (iii) The agent enters full-time retirement close to age 65. The utility parameterization corresponds to an individual with modest impatience, $\theta = 2\%$, and a relatively low degree of risk-aversion, $\sigma = 2$, and who given a relative weight of consumption preference of $\phi = 0.34$ on average works 34 hours/week. For the sensitivity analysis, I rely on numerical illustrations that are not necessarily fit to these goals. I then proceed with simulating the effect of the tax treatment on the life-cycle profile.

3.1.1 Survival

As previously disclosed, the subjective discount consists of the intrinsic discount factor which is weighted by survival probability. I fit a polynomial of order six, $\mathbb{P}(\Xi, t)$, to survival data from the Swedish life table[14], where $\Xi = [\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6]$ is a vector of parameters.

$$\mathbb{P}(\Xi, t) = \exp \left\{ \xi_0 + \xi_1 t + \xi_2 t^2 + \xi_3 t^3 + \xi_4 t^4 + \xi_5 t^5 + \xi_5 t^6 \right\} \tag{35}$$

---

[12]The numerical analysis was conducted using SageMath, a computer algebra system with a Python-like syntax. All programming codes are available upon request.

[13]I consider 6 days available for work per week and that the agent sleeps for 7 hours each night, making 18 hours/day available for labor supply and leisure activities. This implies that the agent has 5,008 hours available to distribute between labor supply and leisure activities per year.

[14]Data was obtained from Statistics Sweden (SCB) for the year 2019.

The calibration results in the following specification for survival probability:

$$\mathbb{P}(t) = \exp(0.0022 - 0.0003 * t + 2.5336 * 10^{-6} * t^2 - 3.3851 * 10^{-6} * t^3 - 2.3323 * 10^{-8} * t^4$$

$$+6.4364 * 10^{-9} * t^5 - 1.1423 * 10^{-10} * t^6).$$

3.1.2 Wage profile

I specify a wage function $w(t)$ such that the simulated life-cycle earnings profile of the representative agent conforms with the observed average life-cycle earnings profile for Swedish males.[15] Since earnings are observed on a yearly basis, I fit the observed average life-cycle earnings with a spline, $\Upsilon(t)$, which allows me to treat the vector of discrete earnings observations as continuous data.

I define $w(\Gamma, t)$ as an exponential polynomial of order seven,[16] where $\Gamma = [\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7]$ is a vector of parameters.

$$w(\Gamma, t) = \exp\{\gamma_0 + \gamma_1 t + \gamma_2 t^2 + \gamma_3 t^3 + \gamma_4 t^4 + \gamma_5 t^5 + \gamma_6 t^6 + \gamma_7 t^7\} \tag{36}$$

I obtain the parameters of the wage function numerically through the following minimization process: (1) I fix retirement at $F = 40.5$, which corresponds to an actual retirement age of 65.5.[17] (2) I define a performance index $G(\Gamma)$ which minimizes the sum of squared difference at discrete time intervals between predicted earnings of the model and the observed earnings

---

[15]Earnings data was collected from the ASTRID register database for the years 2002–2015, which includes earnings data for the Swedish taxable population. The reason for studying a male representative agent follows from a more stable average earnings profile. The observed average female earnings profile has a less concave evolution which makes fitting a well-behaved wage polynomial more difficult.

[16]While resorting to higher degree polynomials is a natural consideration to improve flexibility, it is difficult to avoid Runges phenomenon of oscillating tales (Runge, 1901). Furthermore, increasing the order of the polynomial beyond the seventh order specification does not yield any visible improvements to the fit in this paper.

[17]The average retirement age in Sweden is closer to 64.4, which implies that the average Swede retires before receiving treatment. While the model allows for labor market exit before treatment and a return post-treatment, I consider it to be less likely to be representative as transitions into and out of the labor market are typically associated with various fixed costs. Since the model in this paper abstracts from such costs for tractability, I resort to the above consideration.

from the data.

$$G(\Gamma) = \min_{\Gamma} \sum_{t=0}^{F}(w(\Gamma,t)[1 - h(w(\Gamma,t),F,t)] - \Upsilon(t))^2+$$

$$\left[\int_0^F \{[1 - h(w(\Gamma,t),F,t)]w(\Gamma,t)e^{-rt}\}dt - \int_0^T \{c(w(\Gamma,t),t),F,t)e^{-rt}\}dt\right]^2. \tag{37}$$

Note that the minimization problem in equation (37) includes the life-cycle budget constraint, which ensures that the model fits a feasible optimal labor supply profile conditional on that the observed retirement age is optimal. The calibration results in the following wage profile specification:

$$w(t) = \exp(8.4191 + 2.5049 * (t/F) - 2.5375 * (t/F)^2 - 0.0246 * (t/F)^3 - 0.0001 * (t/F)^4$$

$$+0.0014 * (t/F)^5 - 3.1144 * 10^{-5} * (t/F))^6 - 1.3352 * 10^{-6} * (t/F)^7),$$

where the scaling of the time measure prevents the parameters from exploding. The resulting $w(t)$ is illustrated in Figure (1).

Figure 1: Calibrated per-unit of efficient labor wage.

Evidently, the calibrated wage profile is hump-shaped. While one might expect the wage decline to be less extreme at older ages, especially if following the Swedish institutional setting with rigid wages, one can argue that the estimated profile captures not only a labor efficiency profile but also a health decline or other factors that constrain the ability to actually attend work and participate in the labor market. This is to be expected given that the only "health factor" explicitly included in the model is the survival probability, which represents neither ability nor work performance.

3.1.3   Calibrated profile

Figure 2 illustrates the fitted earnings profile and survival probability polynomial. While peak earnings of the simulated profile coincide well with observed earnings, the simulated earnings profile is evidently much less stable over the middle-aged portion of the life cycle, with a less pronounced drop around retirement age. This follows from the fact that a hump-shaped wage profile does not represent the true rigidity of life-cycle wages in Sweden. However, a non-declining wage profile would not predict an interior (or at least not a realistic) solution to the timing of retirement in this modeling framework. If accounting for organizational constraints and/or fixed costs of labor supply, one expects actual labor supply to be more binary than what a high-frequency friction-free framework is able to predict. For example, while a discrete drop in earnings naturally follows from a discrete transition from working 75 % to full time retirement, it is not obvious when the cessation of working life becomes the result of a smooth reduction in labor productivity.[18] It is reasonable to criticize the continuous time framework and in particular the overly convex leisure profile it predicts under the assumption of a hump-shape wage profile (e.g., Carroll and Summers, 1991; Browning and Crossley, 2001). This could potentially be explained by the absence of fixed costs associated with labor supply. Such frictions would possibly predict a more sudden labor market exit. It is therefore reasonable to expect the intensive margin to be highly

---

[18]As will become clear when introducing social security, the output is consistent with both a discrete reduction in hours worked at the timing of retirement, and a discrete drop in consumption.

variable with the assumed payoff structure in this model framework. While one should not ignore the weakness of the continuous time framework, it has the advantage of allowing for the modeling of the extensive margin decision in a tractable way. However, it is reasonable to reserve the interpretations of the simulated leisure profile in terms of average hours since a fully flexible intensive margin is rarely, if ever, consistent with real world labor markets.
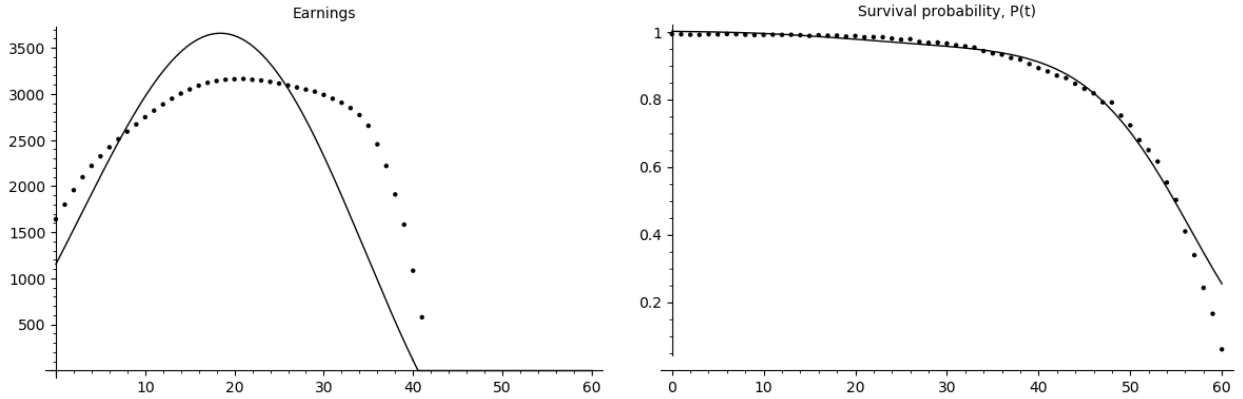

Figure 2: Calibrated earnings and survival probability. Dotted lines represent observations from the earnings and mortality data respectively, and solid lines are the model fit. The value of earnings is in hundreds of Swedish kronor (SEK).

As evident in Figure 3, the timing of peak consumption conforms with the target of age 50–55, with savings peaking just prior to retirement. Labor supply intensity follows the per-unit of efficient labor supply wage profile. The agent spends eight years in debt, becomes solvent, and subsequently accumulates a relatively substantial wealth before entering retirement. This is driven by the net of the effective discount rate and interest rate on savings. When the agent is only moderately impatient, consumption and labor supply expectedly peak when the payoff to labor is the largest.

Since the model only includes non-durable consumption and the net of the interest rate and the intrinsic discount rate is positive, no major loans are undertaken when young. One can argue that the savings profile is more realistic if considering housing as both a real estate good and a liquid asset. When treating debt in the model as net of housing value, it is reasonable that predicted debt is practically eliminated as long as the agent is not too
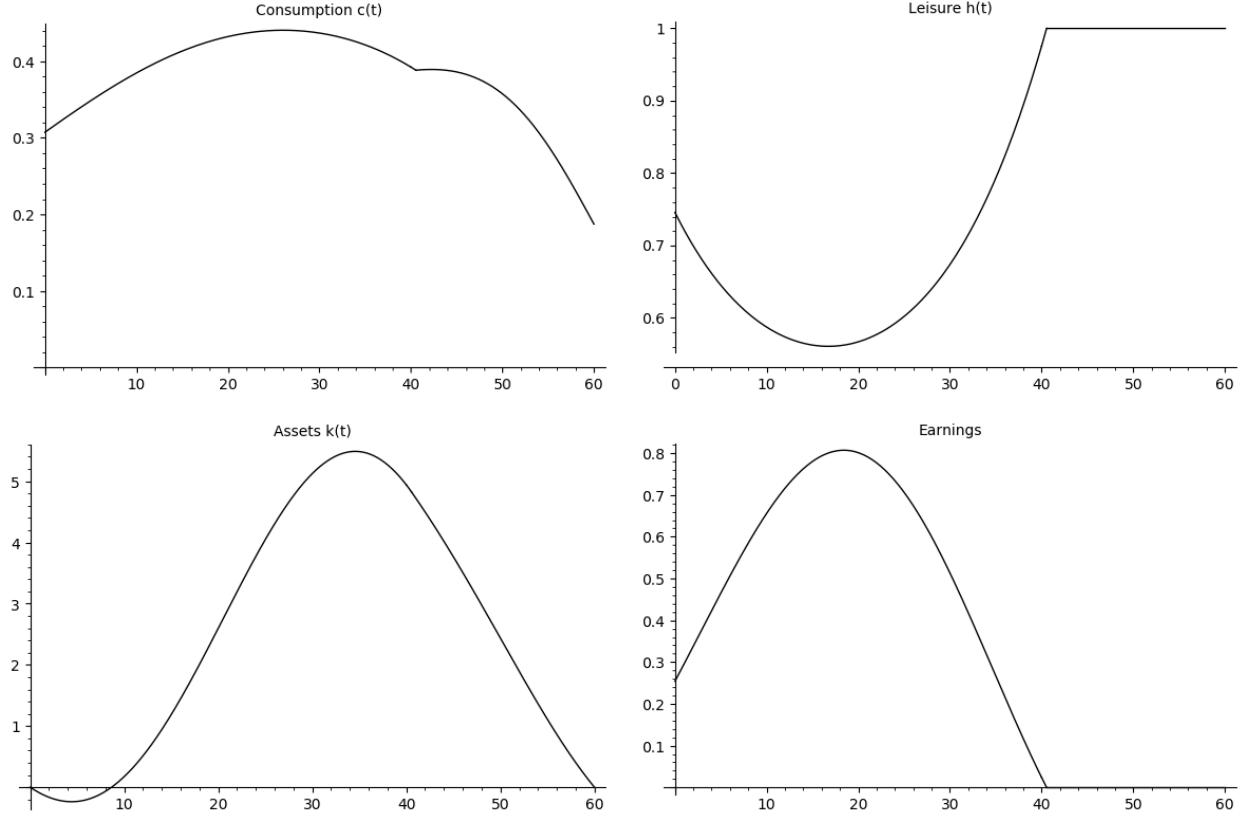
Figure 3: Calibrated life-cycle profile. $r = 0.035, \theta = 2\%, \phi = 0.34, \sigma = 2$. Retirement age $=$ 65.55, age of peak consumption $= 51$, average weekly hours of work $= 34$.

impatient. Since this is not the case for the calibrated profile, the savings profile is intuitive given the structure of preferences and the intertemporal budget constraint.

3.1.4   Effect of tax treatment

I proceed with the main analysis of the effects of the favorable marginal tax treatment on the optimal control profiles of the representative agent. I consider two main scenarios where tax treatment happens at either age 60 or 65, which implies $R = 35$ or $R = 40$. To compute the magnitude of the effect of the age-targeted EITCs on tax revenue, $\mathcal{R}$, I weight the lifetime contributions of the representative agent by survival probability, which in an aggregate context would correspond to the total tax revenue at any instant in time.

24

$$\mathcal{R} = \int_0^R \mathbb{P}(t)[1 - h(t)]w(t)\tau_1 dt + \int_R^F \mathbb{P}(t)[1 - h(t)]w(t)\tau_2 dt \tag{38}$$

Figure (4) illustrates the effects of a 10 percentage point tax treatment at age 65 on consumption, leisure, and savings when considering the profile calibrated to fit observed life-cycle earnings. Since the effects are smoothed over extensive portions of the life cycle, and thus difficult to observe from the figures, I quantify the effects on labor supply and retirement timing in Table 1.
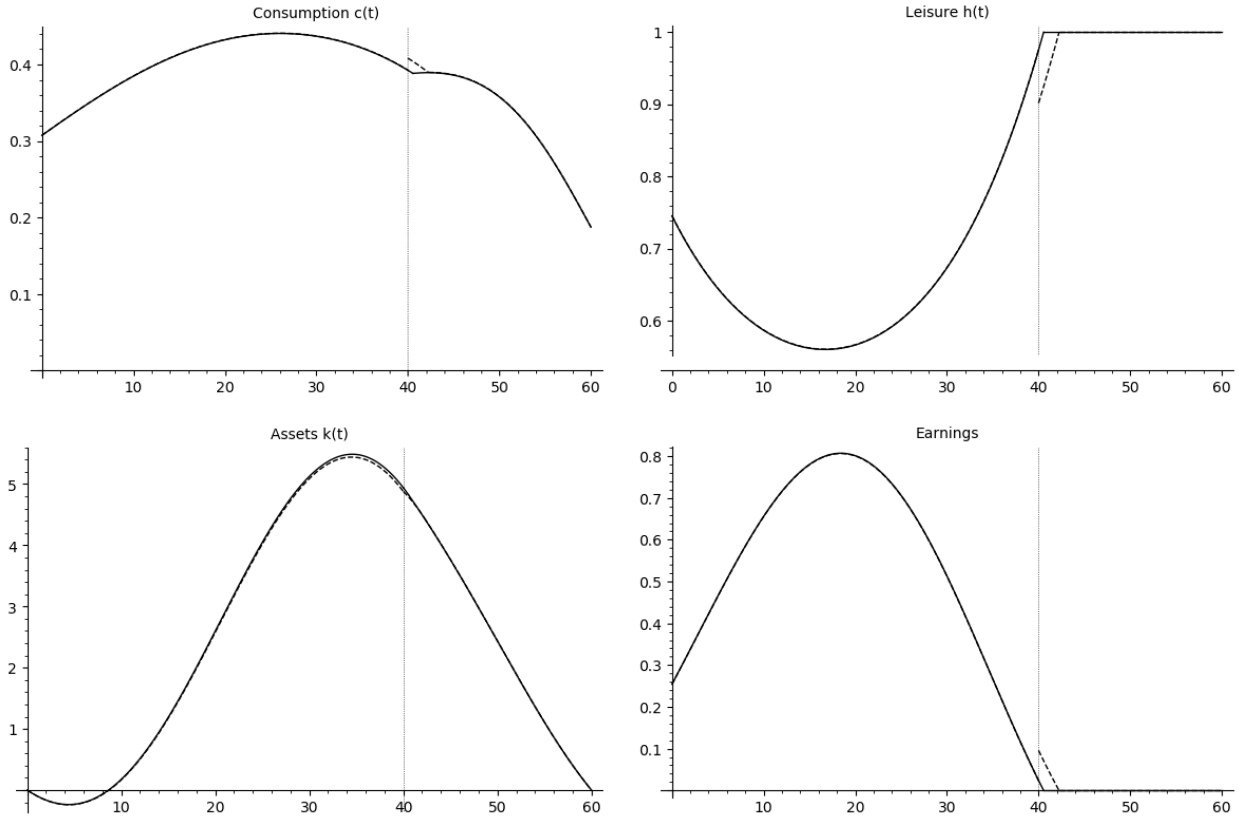

Figure 4: Effect of tax treatment on the life-cycle profile ($\tau_1 = 0.2$, $\tau_2 = 0.1$). The solid lines illustrate the control profile, while the dashed lines illustrate the treated life-cycle profile. $r = 3, 5\%, \theta = 2\%, \phi = 0.34, \sigma = 2$.

Ceteris paribus, the tax treatment induces the agent to delay retirement, but partially due to substitution of labor supply from non-treated to treated age. Given the discrete increase in the net wage following treatment, the agent chooses to increase the hours worked

post-treatment and delays retirement. The agent saves less intensely during their working life as a result of the anticipated increase in net wage following the tax regime shift when eligible for retirement, as they do not need the same capital buffer for late life consumption relative to the control profile. The discrete change of the savings rate is clearly visible at the timing of tax treatment as a slight kink in the asset profile. Consumption expectedly increases over the entire life cycle following the increase in life-cycle net income, with a distinct increase coinciding with the treatment and the increased net wage.

Table 1 presents the magnitude of the effects. These could be interpreted as the aggregate quantities of an OLG economy in a steady state with a constant continuum of identical altruistic agents. Changing the marginal tax at age 65 from 0.20 to 0.1 (a 12.5 % increase in the net wage) leads to a delayed retirement by 1.68 years and increases life-cycle labor supply by 0.59 %, which corresponds to approximately 10 weeks.[19] The labor supply of the pre-treatment phase decreases by 0.13 %, or by 2 weeks, which underlines that the increase in extensive margin labor supply can be partially ascribed an intertemporal substitution of labor supply. Lifetime consumption is computed to increase by 0.86 %. If treatment instead takes place at age 60 instead, the model predicts total labor supply to increase by 2.23 % and labor supply pre-treatment to decrease by 0.65 %. This is expected since treatment now coincides with an age when the agent is more productive. While this is certainly confirmed when considering total labor supply, it is interesting to note that the retirement age increases by only 1.61 years. The effect is seemingly larger on the intensive margin labor supply than on the participation margin.

To provide some comparison for the labor supply response magnitudes, French and Jones (2012) estimate that an anticipated wage increase of 20% at age 60 leads to a 1,906 hour increase in total life-cycle labor supply and a 519 hour decrease in pre-treatment labor supply.

---

[19]1 week = five 8-hour working days

Table 1: Quantitative results. Ret. age = retirement age, LS1 = pre-treatment labor supply, TLS = total life-cycle labor supply, C = rotal consumption, Tax tev. = tax revenues.

Treatment at age 65						
τ_1	τ_2	Ret. age	%Δ LS1	% Δ TLS	% ΔC	% Δ Tax rev.
0.20	0.20	65.55	-	-	-	-
-	0.18	65.90	-0.01	0.07	0.02	0.029
-	0.16	66.25	-0.03	0.16	0.04	0.058
-	0.14	66.59	-0.05	0.28	0.08	0.082
-	0.12	66.91	-0.08	0.43	0.12	0.092
-	0.10	67.23	-0.12	0.59	0.16	0.084
-	0.08	67.53	-0.15	0.77	0.21	0.054
-	0.06	67.83	-0.20	0.97	0.27	-0.002
-	0.04	68.13	-0.24	1.18	0.33	-0.087
-	0.02	68.42	-0.29	1.41	0.40	-0.20
-	0.00	68.68	-0.35	1.64	0.48	-0.355
Treatment at age 60						
τ_1	τ_2	Ret. age	%Δ LS1	% Δ TLS	% ΔC	% Δ Tax rev.
0.20	0.20	65.55	-	-	-	-
-	0.18	65.90	-0.12	0.44	0.17	-0.08
-	0.16	66.23	-0.24	0.88	0.34	-0.24
-	0.14	66.55	-0.38	1.33	0.52	-0.48
-	0.12	66.86	-0.51	1.78	0.70	-0.81
-	0.10	67.16	-0.65	2.23	0.89	-1.21
-	0.08	67.46	-0.80	2.69	1.09	-1.70
-	0.06	67.74	-0.95	3.14	1.29	-2.28
-	0.04	68.02	-1.11	3.59	1.50	-2.93
-	0.02	68.28	-1.26	4.03	1.71	-3.67
-	0.00	68.54	-1.43	4.47	1.92	-4.48

Note: The percentage change is in relation to the calibrated profile without any age-differentiation in the tax code.

That is, the decrease in pre-treatment hours is 27 % of the total increase in labor supply. The model in this paper predicts a 2403 hour increase in total labor supply, and a 709 hour decrease in pre-treatment labor supply for the same increase in net wage at age 60 (I set $R = 35$, and $\tau_2 = 0.04$). The model thus predicts that the decrease in pre-treatment labor

supply is 29 % of the increase in total labor supply. This suggests that the intertemporal substitution effects are of similar proportion. In addition to slight differences in behavioral parameterization,[20] the larger treatment effect found in this paper could possibly be explained by differences in preference specification and the absence of policy uncertainty. In the next section I therefore conduct numerical illustrations with the inclusion of anticipation effects by considering the risk that the policy will be abandoned. Yet, it is interesting to note that the tractable model in this paper delivers results that agrees with the predictions of a highly sophisticated empirical model framework to a large extent.

When considering treatment at age 65, tax revenues increase relative to the control scenario for favorable treatment up to a magnitude of 11 percentage points. For more generous treatments, tax revenues begin to decrease compared with the control scenario. A maximum increase in tax revenues is achieved for an 8 percentage point tax treatment. Overall, the simulations suggest that the favorable tax treatment is close to revenue neutral for treatment at age 65. A small to moderate treatment (2–12 percentage points) leads to both delayed retirement and increased tax revenue. While more generous tax credits will increase incentives to delay retirement even more, this will come at the expense of less tax revenue collected relative to the control scenario of no favorable tax treatment. The magnitude of the treatment effect in terms of labor supply response on both margins could however be exaggerated due to a lack of policy uncertainty in the modeling framework. If instead considering treatment at age 60, the tax treatment has a negative effect on revenues for any magnitude of tax treatment. If larger than 8 percentage points, the revenue loss is larger than 1%.

---

[20]I assume a relatively less risk averse agent relative French and Jones (their estimates vary between 2.2 and 5.1 for the inverse of intertemporal elasticity of consumption). The relative weight of consumption preferences is also estimated to be larger in his paper relative the parameterization in this paper (their estimates vary between 0.533 and 0.615).

3.2 Modeling experiments

In this section I perform a number of modeling experiments to complement the analysis of the calibrated profile. The first is a sensitivity analysis of changes in the behavioral parameterization. More specifically, I consider one more impatient and one more risk-averse agent relative to the calibrated profile. Next, I introduce policy uncertainty in two ways: first by introducing the probability that the tax treatment is abandoned, which the agents effectively weigh into their planned behavior, and second by considering an agent who rationalizes the future tax treatment later in life when the treatment is within a more foreseeable future. Finally, the model is augmented with a defined contribution social security scheme to relax the assumption of an entirely self-financed retirement and to provide a more realistic institutional context.

Two additional experiments have been relegated to the appendix. One refers to the assumption of exponential discounting, which has been challenged by findings in the behavioral economic literature.[21] I therefore test whether the intertemporal effects under the assumption of hyperbolic discounting are consistent with the main analysis. In the second, the assumption of zero inherited and bequeathed wealth is relaxed by introducing heterogeneity in initial endowments and bequest amounts.

3.2.1 Sensitivity analysis

I conduct a sensitivity analysis with regard to the behavioral parameterization by varying the degree of impatience and elasticity of intertemporal substitution. Figure (5) illustrates the effect of a more impatient agent by considering a higher intrinsic discount rate. Figure (6) illustrates an agent with a smaller elasticity of intertemporal substitution, i.e. a more risk-averse agent.

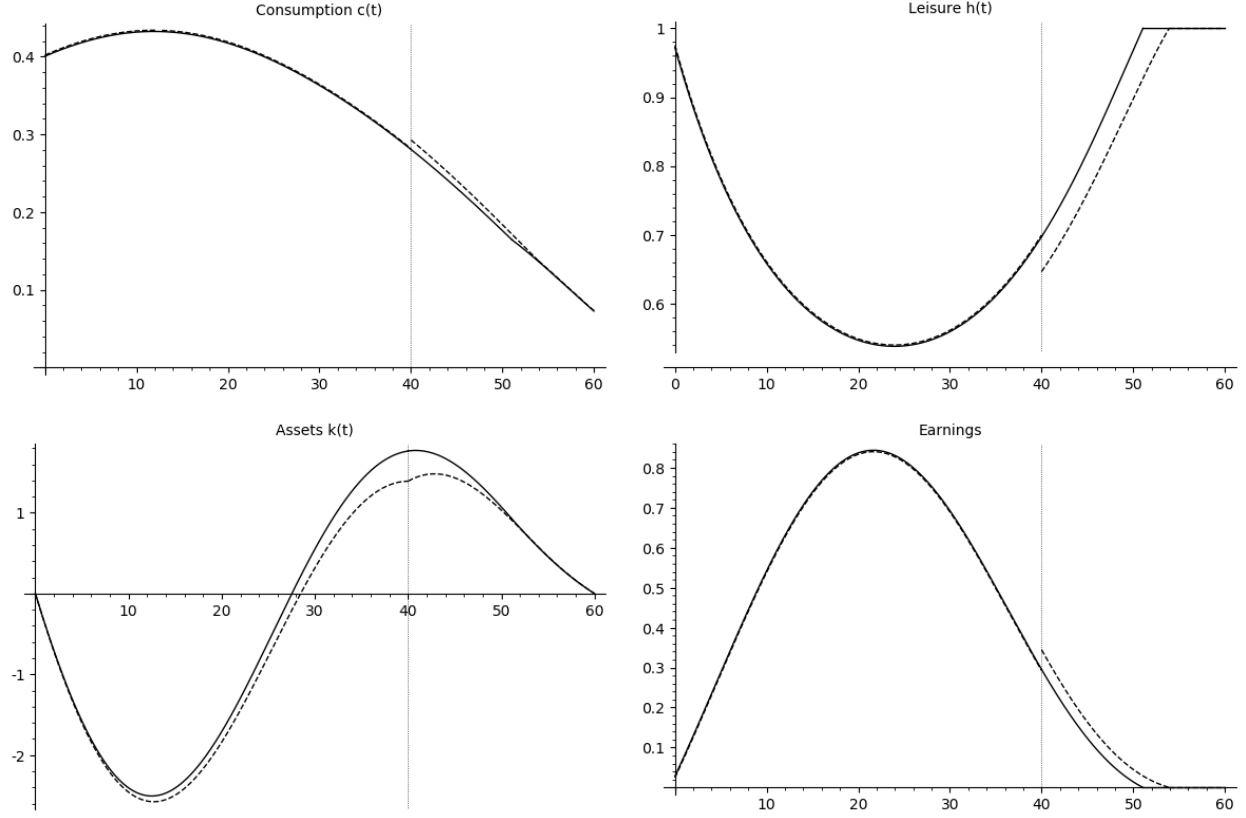[21]The interested reader is referred to DellaVigna (2009) for an extensive survey.

Figure 5: "Impatient" life-cycle profile. $r = 3.5\%, \theta = 5\%, \phi = 0.34, \sigma = 2$. The dashed line illustrates the treated profile ($\tau_1 = 0.2$, $\tau_2 = 0.1$).

The immediate take-away from the more impatient profile is the substantial increase in borrowing compared withthe baseline simulation in the previous subsection. It follows naturally that when the intrinsic discount rate is close, equal to, or larger than the interest rate, consumption peaks early in life when the labor payoff is low. Since the increase in the payoff following treatment coincides with a period that the agent now puts less utility weight on, the utility-maximizing choice is to delay more labor supply to old age.

Concerning the more risk-averse agent illustrated in Figure 6, the working life is prolonged and the labor supply varies less dramatically over the life-cycle relative to the baseline simulation. The savings behavior is similar as the agent does not accrue any major debt during their younger years. That is, relative to the impatient profile, the agent does not need to compensate for a low labor income stream by hefty borrowing, and thus swiftly becomes
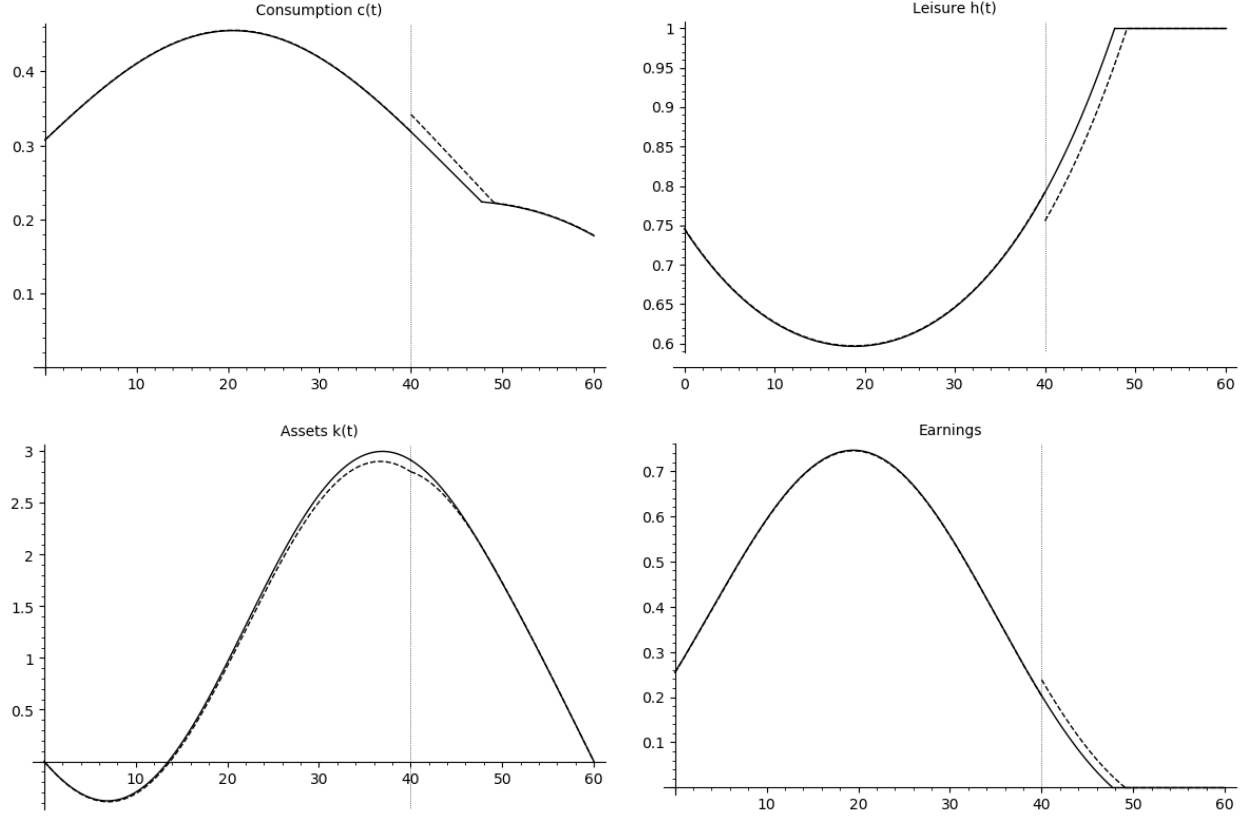
Figure 6: More risk-averse life-cycle profile. $r = 3.5\%, \theta = 2\%, \phi = 0.34, \sigma = 10$. The dashed line illustrates the treated profile ($\tau_1 = 0.2$, $\tau_2 = 0.1$).

solvent once capital returns start to accumulate.

3.2.2 Anticipation effects

Let $\mathbb{Q} \in [0, 1]$ be the probability that the policy of favorable tax treatment of older workers is abandoned before the agent is eligible for treatment, and $(1 - \mathbb{Q})$ the probability that the policy remains. The optimal control profile of the representative agent can then be computed as a weighted average of the profiles when the policy is abandoned and when it remains. For example, the consumption profile will be equal to $c_a(t) = \mathbb{Q}c(t)|_{\tau_1 = \tau_2} + (1 - \mathbb{Q})c(t)|_{\tau_1 > \tau_2}$. Table 2 reports the results for $\mathbb{Q} = 0.25, 0.5, 0.75$.

Given treatment at age 65, if the probability of abandonment is 75%, a 10 percentage point tax treatment leads to a delayed retirement age by approximately 0.42 years. This can

31

be compared with the guaranteed treatment scenario in which the agent delays retirement by approximately 1.68 years. The effect becomes even more modest when considering the change in total labor supply. While the scenario with guaranteed treatment suggests that total working hours are increased by approximately 398 hours, a 75% risk of the policy being abandoned largely offsets the increase and results in approximately 45 additional working hours.

Table 2: Comparison of control agent, and treatment effects for different probabilities that the policy will be abandoned. Tax treatment implies tax rates $\tau_1 = 0.2$, $\tau_2 = 0.1$.

Treatment at age 65			
	Total labor supply	Pre-treatment labor supply	Retirement age
Control	67022	66991	65.55
$\mathbb{Q} = 0\%$	67420	66913	67.23
$\mathbb{Q} = 25\%$	67265	66933	66.81
$\mathbb{Q} = 50\%$	67147	66952	66.39
$\mathbb{Q} = 75\%$	67067	66972	65.97
Treatment at age 60			
	Total labor supply	Pre-treatment labor supply	Retirement age
Control	67022	64152	65.55
$\mathbb{Q} = 0\%$	68520	63733	67.16
$\mathbb{Q} = 25\%$	68095	63838	66.75
$\mathbb{Q} = 50\%$	67703	63943	66.35
$\mathbb{Q} = 75\%$	67346	64047	65.95

An alternative approach to model anticipation effects is to consider an agent who waits until the treatment lies within a foreseeable time horizon before responding. For this, I consider three individuals who differ in their accumulated wealth at age 55: the representative agent, a relatively wealthier agent, and a relatively less wealthy agent. The simulated profiles follow the calibrated profile but instead start from age 55 and only differ in assets accumulated until that age, ceteris paribus. A wealthy enough agent will likely retire earlier than the representative agent, and the tax treatment may not be large enough to provide incentives for re-entry after reaching the eligibility age. Therefore, interpreting the effects on

the labor supply of the representative agent as an average effect might exaggerate the policy's effect on life-cycle behavior. In the example illustrated in Figure 7, the representative agent's response will be quantitatively larger than the average effect. This follows from the fact that the incentive to increase labor supply due to the tax treatment is not large enough for the wealthier agent to re-enter the labor market.
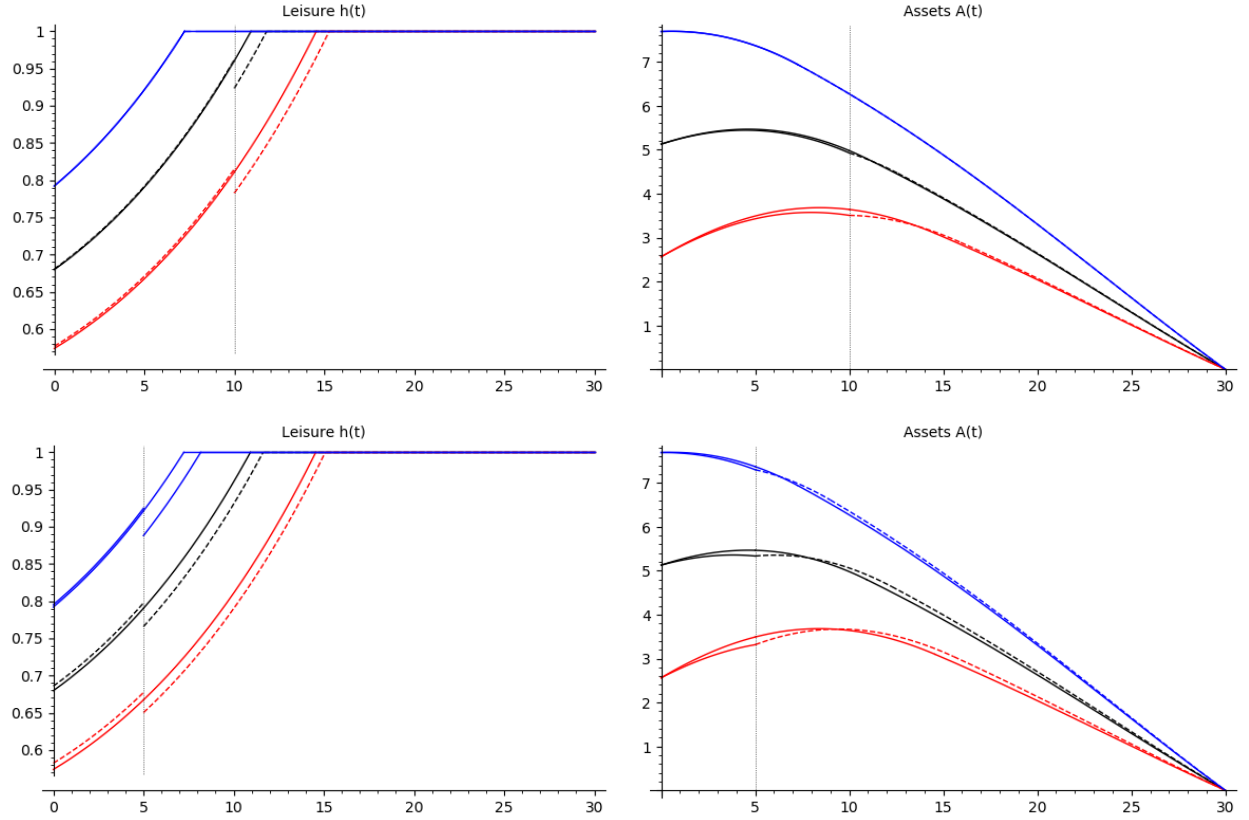

Figure 7: Heterogeneity in assets at age 55. $r = 3.5\%, \theta = 2\%, \phi = 0.34, \sigma = 2$. Black lines illustrate the representative agent, blue lines the relatively wealthier agent, and red the relatively less wealthy agent. Solid lines represent the control profiles and dashed lines illustrate treated profiles. Note: Observe that t $= 0$ corresponds to age 55. The vertical line indicates timing of treatment after age 55.

Quantifying the results illustrated in Figure 7, the average effect of the treatment on retirement timing at age 65 is a 1.01 year delay, while the representative agent delays retirement by 1.65 year, which suggests a large discrepancy between the average and the representative profiles. Besides illustrating the role of anticipation, this exercise also shows the potential

33

discrepancy between the average effect of the tax treatment on labor supply response and the response of the representative agent. More elaborate modeling of individual heterogeneity is therefore suggested for future analyses.

3.3 Social security

In this section I introduce a defined contribution social security scheme that redistributes a fraction of labor income tax contributions to annuities for withdrawal during retirement. With the inclusion of retirement benefits, I redefine the asset accumulation function governing the savings dynamics:

$$
\dot{k} = \begin{cases} [1 - h(t)](1 - \tau_1 - \tau_s)w(t) + rk(t) - c(t) & \text{for } t \in [0, R), \\ [1 - h(t)](1 - \tau_2 - \tau_s)w(t) + rk(t) - c(t) & \text{for } t \in [R, F), \\ \bar{b}(1 - \tau_b) + rk(t) - c(t) & \text{for } t \in [F, T]. \end{cases} \tag{39}
$$

where $\bar{b}$ is the annuity amount received during retirement, financed by a social security tax τ_s which is not age-differentiated. τ_b is a proportional tax on pension income.[22] The government ensures a solvent social security budget by determining the total benefit amount b such that the following condition is satisfied:

$$
b = \tau_s \int_0^F \mathbb{P}(t)(1 - h(t))w(t)dt, \tag{40}
$$

where the RHS of equation (40) is the total social security tax contributions. By dividing equation (40) by the survival probability, the account is effectively annuitized:

$$
\bar{b} = \frac{\tau_s[\int_0^R \mathbb{P}(t)(1 - h(t))w(t)dt + \int_R^F \mathbb{P}(t)(1 - h(t))w(t)dt]}{\int_F^T \mathbb{P}(t)dt}. \tag{41}
$$

---

[22]The reason for including a tax on pension income is to prevent an unrealistically early retirement given the behavioral parameterization.

The numerical procedure follows the principles of the numerical iteration algorithm of Auerbach et al. (1987):

1. *Guess the values of retirement age, F, and labor supply of subdomains 1 and 2, and store these in a vector v_G;*

2. *Compute retirement benefits $\bar{b}$ given arguments in vector v_G.*

3. *Given the annuity value, I re-optimize to solve for new values of F and $\mu(0)$;*

4. *From the new life-cycle profile, I compute the new values of labor supply and retirement age, and store these in a new vector, v_F;*

5. *The algorithm iterates on v_F. I consider the model to have converged when $(v_F - v_G)^2 < 0.01$.*

I model the effects of treatment at age 60, since the inclusion of social security will increase incentives for earlier retirement. I consider a slightly modified behavioral parameterization by ascribing the agent a lower elasticity of intertemporal substitution, $\sigma = 6$. The reason for this modification is that a representative agent enrolling in social security would retire at a counterfactually young age if adhering to the same preference parameters as the self-financing counterpart. I set $\tau_s = 0.1$, $\tau_b = 0.25$, $\tau_1 = 0.2$, where a treated profile will face $\tau_2 = 0.1$.

Figure (8) illustrates the effects of tax treatment. Optimal retirement occurs when the accumulated benefits reach a critical value, after which it is welfare maximizing to exit the labor market and consume out of savings and pension annuities. This results in a discrete increase in leisure at retirement contrary to the smooth reduction in working hours predicted by the model without social security.
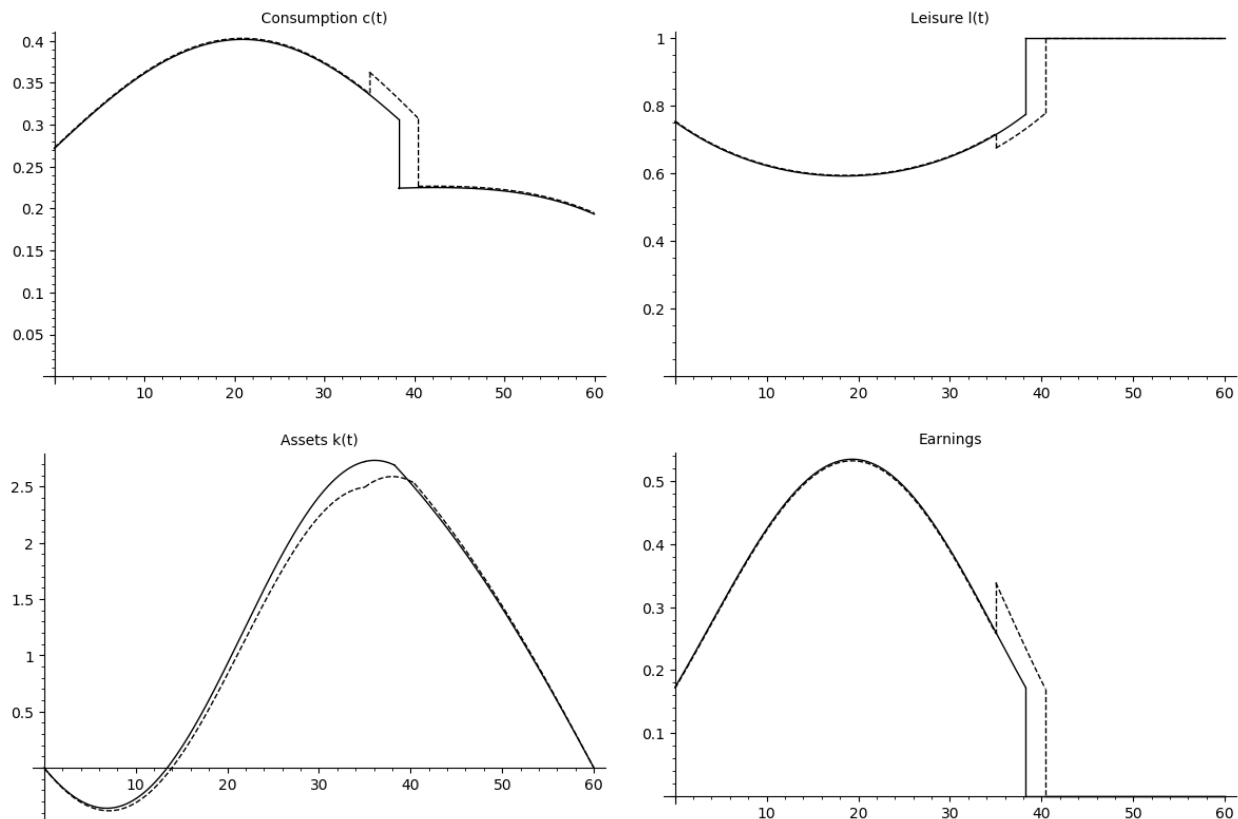
35

Figure 8: Effect of social security. Control (solid lines) and treated (dashed lines) profiles. $\tau_s = 0.1$, $\tau_b = 0.25$, $\tau_1 = 0.2$, $\sigma = 6$, $r = 3.5\%$, $\theta = 2\%$, $\phi = 0.34$

The output is consistent with the commonly observed consumption drop at retirement, a result of smoothing the consumption of the composite good as opposed to just smoothing consumption. The main intuition of the self-financed profile still holds: labor supply decreases before and increases following treatment. Consumption and total labor supply increase, and retirement is effectively delayed. A 10 percentage point treatment at age 60 leads to an increase in life-cycle labor supply by 2,465.03 hours from 64,540.19 to 67,005.22 hours, and delays retirement by 2.16 years from 63.24 to 65.4. More specifically, labor supply during the first subdomain decreases by 341.29 hours from 60,921.31 to 60,580.02 hours.

36

4 Concluding remarks

A favorable marginal tax treatment of older workers is a potential policy approach to promote a longer working life. This form of earned labor tax credits was implemented in Sweden in 2007 and aimed to engage the labor supply reserve of agents on the verge of retirement by increasing the net returns to labor later in life. A static model of labor supply would predict such policy to increase the extensive margin labor supply and lead to delayed retirement. However, in a dynamic model accounting for the change in opportunity cost of leisure following the tax treatment, the behavioral response is more ambiguous.

I model the effects of age-differentiated labor income taxes on life-cycle labor supply, consumption/savings behavior, and retirement timing. By augmenting the standard Heckman type model of life-cycle labor supply to a life-cycle model of three subdomains, I am able to provide fully analytical life-cycle profiles for consumption, savings, intensive margin labor supply, and retirement timing. In a life-cycle model, the labor-leisure tradeoff includes the intertemporal substitution effect arising because leisure is relatively cheaper before treatment.

The results of my numerical illustrations suggest that favorable marginal tax treatment of older workers promotes longer working lives. The agent realizes that the tax change implies that leisure becomes relatively more expensive post-treatment and will therefore allocate more time to leisure during the primary working life. The intertemporal substitution effect on savings behavior is more evident when considering a more impatient agent since it puts more utility weight on leisure activities and consumption at earlier stages of the economic life. Tax treatment simply coincides with a period far into the future that the agent assigns less weight to in terms of disutility from work compared with a more patient agent. Therefore, the agent accumulates a larger debt and compensates by working relatively more when old.

The model is able to simultaneously reproduce several salient life-cycle facts such as the timing of the consumption hump, the timing of retirement, and the average work hours over the lifecycle. However, the observed earnings profile is visibly flatter during the middle-aged portion of the life-cycle arc compared with calibrated earnings, the profile of which has a more distinct peak. This result follows from the fact that the model allows for a completely flexible labor supply. What follows is that optimal labor supply co-varies with a distinctly hump-shaped efficiency wage profile, which results in a large variation in hours worked over the life cycle. Furthermore, borrowing is almost completely absent when agents are moderately impatient. With regard to the discrepancy between observed and calibrated earnings, a better fit could potentially be achieved by the inclusion of fixed costs of labor supply or time-varying leisure preferences. This observation could potentially be explained by the absence of durable goods and any loans associated with housing purchases.

The model in this paper provides a simple, tractable framework to analyze age-targeted labor income taxation without compromising on the analytics on any of the labor supply margins. The analysis is not limited to the tradeoff faced by older workers as in a static framework. Furthermore, the results are robust to the inclusion of social security in the model as well as to various changes in behavioral assumptions and the effective discounting function.

An interesting avenue for future studies would be to extend the analysis into a general equilibrium framework. This would allow for studying the transitional dynamics of economic aggregates and fiscal solvency of social security when introducing the age-targeted income taxation policy. Regarding the microfoundations, it naturally follows to consider other dimensions of the labor-leisure tradeoff such as human capital and health capital accumulation. Since age-differentiated labor income taxation affects the retirement margin, it will also af-

fect incentives for human capital investments by changing the life-cycle financial returns to education. Endogenous health capital and longevity would likely provide a more realistic representation of the intertemporal tradeoffs determining life-cycle labor supply and savings. I leave these suggestions for future research to explore.

References

Alpert, A. and Powell, D. (2013). Estimating intensive and extensive tax responsiveness: Do older workers respond to income taxes?

Auerbach, A. J., Kotlikoff, L. J., et al. (1987). *Dynamic fiscal policy*. Cambridge University Press.

Biggs, A. (2012). A payroll tax cut could help social security. *Wall Street Journal*, 25.

Browning, M. and Crossley, T. F. (2001). The life-cycle model of consumption and saving. *Journal of Economic Perspectives*, 15(3):3–22.

Bütler, M. (2001). Neoclassical life-cycle consumption: A textbook example. *Economic Theory*, 17(1):209–221.

Caliendo, F. N. and Findley, T. S. (2020). Dynamic consistency and regret. *Journal of Economic Behavior & Organization*.

Carroll, C. D. and Summers, L. H. (1991). Consumption growth parallels income growth: some new evidence. In *National saving and economic performance*, pages 305–348. University of Chicago Press.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic literature*, 47(2):315–72.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, 72(2):395–427.

French, E. and Jones, J. (2012). Public pensions and labor supply over the life cycle. *International Tax and Public Finance*, 19(2):268–287.

Gahramanov, E. and Tang, X. (2016). Labor-leisure choice: Is everything as straightforward as one might have thought? *Theoretical Economics Letters*, 6(04):750.

Hall, R. E. and Jones, C. I. (2007). The value of life and the rise in health spending. *The Quarterly Journal of Economics*, 122(1):39–72.

Hartl, R. F. and Sethi, S. P. (1983). A note on the free terminal time transversality condition. *Zeitschrift für Operations-Research*, 27(1):203–208.

Heckman, J. (1974). Life cycle consumption and labor supply: An explanation of the relationship between income and consumption over the life cycle. *The American Economic Review*, 64(1):188–194.

Keane, M. P. and Wasi, N. (2016). Labour supply: the roles of human capital and the extensive margin. *The Economic Journal*, 126(592):578–617.

Kuhn, M., Wrzaczek, S., Prskawetz, A., and Feichtinger, G. (2015). Optimal choice of health and retirement in a life-cycle model. *Journal of Economic Theory*, 158:186–212.

Lacomba, J. A. and Lagos, F. (2006). Population aging and legal retirement age. *Journal of Population Economics*, 19(3):507–519.

Laitner, J. and Silverman, D. (2012). Consumption, retirement and social security: Evaluating the efficiency of reform that encourages longer careers. *Journal of Public Economics*, 96(7-8):615–634.

Laun, L. (2017). The effect of age-targeted tax credits on labor force participation of older workers. *Journal of Public Economics*, 152:102–118.

Mao, H., Ostaszewski, K. M., and Wang, Y. (2014). Optimal retirement age, leisure and consumption. *Economic Modelling*, 43:458–464.

Mao, H., Ostaszewski, K. M., and Wen, Z. (2019). The retirement decision with consideration of part-time work after retirement. *Journal of Insurance Issues*, 42(1):86–112.

Runge, C. (1901). Über empirische funktionen und die interpolation zwischen äquidistanten ordinaten. *Zeitschrift für Mathematik und Physik*, 46(224–243):20.

Seierstad, A. and Sydsaeter, K. (1987). *Optimal control theory with economic applications*. Elsevier North-Holland, Inc.

Strulik, H. and Trimborn, T. (2018). Hyperbolic discounting can be good for your health. *Journal of Economic Psychology*, 69:44–57.

Tomiyama, K. (1985). Two-stage optimal control problems and optimality conditions. *Journal of Economic Dynamics and Control*, 9(3):317–337.

Appendix

Hyperbolic discounting

To explore the effect of the discounting method, I compare the profile of the exponentially discounting individual with a hyperbolically discounting agent. Following Strulik and Trimborn (2018) the effective discount factor $\Theta(t) = \mathbb{P}(t)e^{-\theta t}$ is replaced with the following specification: $\frac{\mathbb{P}(t)}{1+\alpha t}$. To make a direct comparison of a hyperbolic agent with the benchmark, exponentially discounting representative agent, I apply the finite-time equivalent-present-value argument (e.g., Strulik and Trimborn, 2018). Consider a constant stream of one unit of an argument[23] $x \in \mathbb{R}^+$ in a performance index $\Omega(\Theta(t), x, t)$ of a finite lifetime problem with $t \in [0, T]$. The difference between the present value performance conditional on the two discounting methods is then given by:

$$\Delta\Omega(\Theta(t), x = 1, t) = \int_0^T \mathbb{P}(t)e^{-\theta t}dt - \int_0^T \frac{\mathbb{P}(t)}{1+\alpha t}dt. \tag{42}$$

Any difference in the performance index $\Delta\Omega(\Theta(t), x, t) \neq 0$ can be attributed to differences in the discounting method, $\Delta\Omega(t) \neq 0$. Thus, the condition $\Delta\Omega(t) = 0$ will ensure that the present value of the constant stream over the lifetime is the same under both exponential and hyperbolic discounting. For any chosen value of the exponential discount rate θ, one is able to identify the corresponding value of α. Ultimately the normalization method allows the comparison of numerical illustrations of an exponential agent with an "equivalent" hyperbolic agent.

I begin by comparing the control simulations of the benchmark agent and the hyperbolic agent. These profiles are illustrated in Figure (9). Since the hyperbolic agent adds more weight to late-life welfare, they retire earlier and enjoys more consumption when retired. This is financed by more labor supply during their middle-age phase when the labor payoff is the largest. Comparing life-cycle assets, the hyperbolic agent borrows slightly more when young, but accumulates a higher level of wealth following a period of more labor supply and less consumption. This savings buffer allows for notably higher expenditures when retired.

Given the convex decline in its degree of impatience, the hyperbolic agent puts more weight on late-life utility, compared with the exponential agent. Had there been a discrepancy between planned and actual behavior following a propensity for immediate gratification, the conventional solution associated with undersaving and lower old-age welfare would likely be the realized outcome.[24] The comparison of treated profiles is illustrated in Figure (9).

The perhaps most interesting finding is that a tax treatment of 10 percentage points

---

[23]For example 1 unit of income, or 1 unit of expenditure.

[24]I abstract from time-inconsistent solutions in this paper since it follows from continuous re-optimization of the performance index over the life-cycle domain which is a much more demanding computational exercise. While one can certainly elaborate further on the role of the discounting method, the key take-away from this exercise is that foregoing the exponential discount factor does not yield perverse results. That is, the results are not ad hoc to the choice of discounting method.
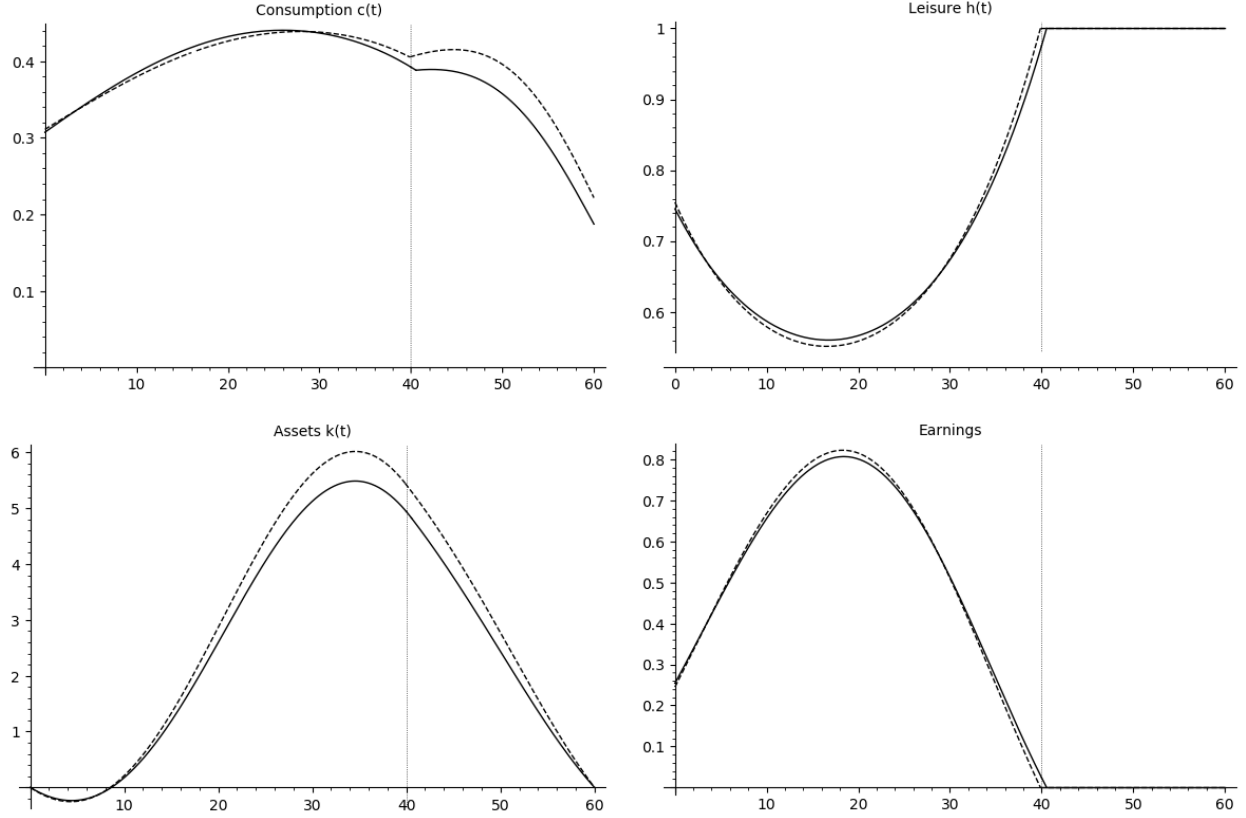
Figure 9: Exponential (solid lines) and hyperbolic (dashed lines) non-treated profiles. The value of α corresponding to θ=2% was computed to be 2.86%. $r = 3.5\%, \phi = 0.34, \sigma = 2$.

generates large enough monetary incentives for the hyperbolic agent to return to supplying labor during the treatment period. However, the return is brief and the hyperbolic agent still full-time retires earlier than the exponential one.

Inherited and bequeathed wealth

I now consider three agents who receive the tax treatment but differ in inherited wealth and bequeathed wealth: 1) one who inherits but does not bequeath any wealth ($k_0 = 10000$, $k_T = 0$), 2) pme who does not inherit but bequeaths wealth ($k_0 = 0, k_T = 10000$), and 3) one who both inherits and bequeaths wealth ($k_0 = k_T = 10000$). The life-cycle profiles are illustrated in Figure (10).

The results are intuitive. Agent 1 utilizes the inherited wealth as a buffer to increase life-cycle leisure without compromising on consumption expenditure relative to the calibrated profile. This results in retirement prior to the period when the the tax treatment takes place, together with a brief return to the labor market following treatment. Agent 2 also leaves the labor market prior to treatment and returns post-treatment, but works more in total and consumes less to ensure sufficient savings for the bequeathed amount. Agent 3 acts similar to the calibrated profile, but has to work longer and consumes less to reach the bequeathed
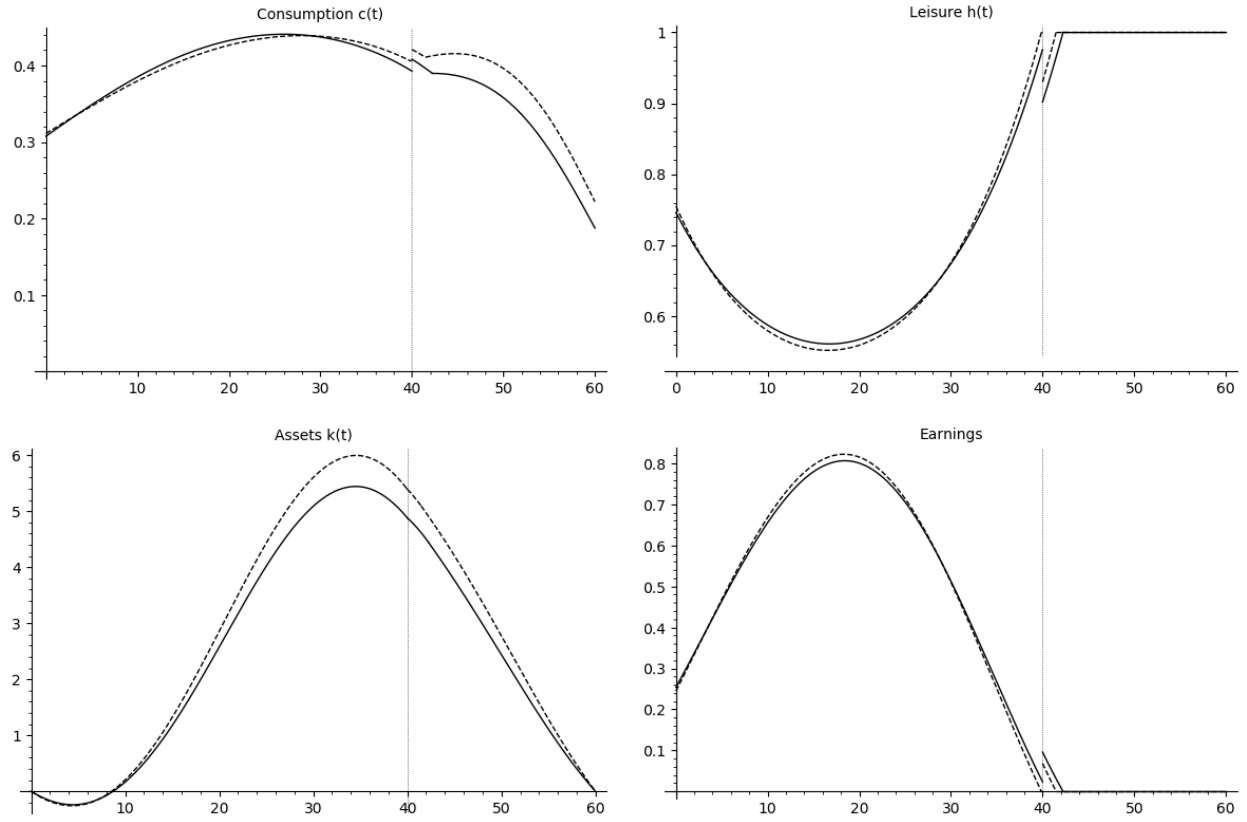
42

Figure 10: Exponential (solid lines) and hyperbolic (dashed lines) treated profiles ($\tau_1 = 0.2$, $\tau_2 = 0.1$). The value of α corresponding to θ=0.02 was computed to be 0.0286. $r = 3.5\%, \phi = 0.34, \sigma = 2$.
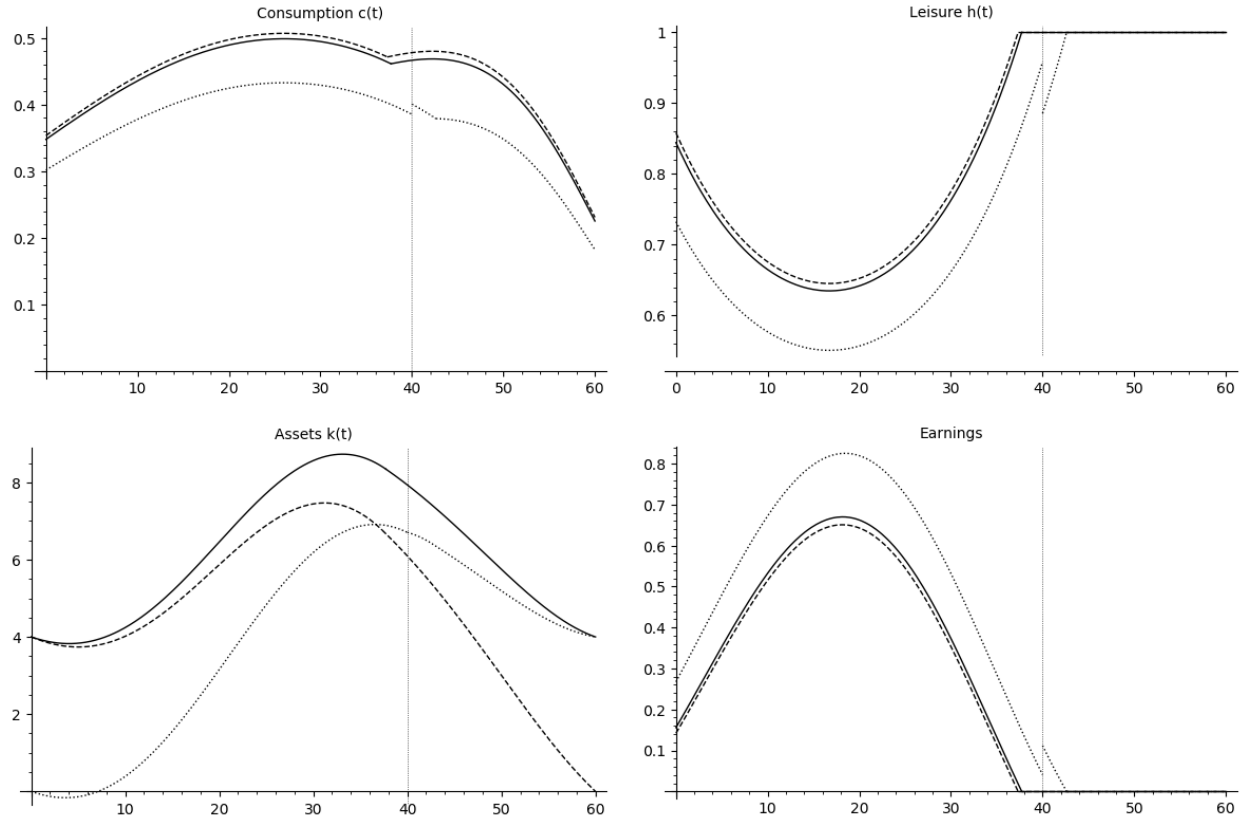
wealth.

Figure 11: Variations in inherited and bequeathed wealth among treated profiles ($\tau_1 = 0.2$, $\tau_2 = 0.1$). Agent 1 (dashed line), Agent 2 (dotted line), and Agent 3 (solid line). $r = 3.5\%, \theta = 2\%, \phi = 0.34, \sigma = 2$.