Determining the Motives for a Positive Optimal Tax on Capital

William B. Peterman*
Board of Governors of the Federal Reserve System

May 1, 2012

Abstract

Previous literature demonstrates that in a computational life cycle model the optimal tax on capital is positive and large. Given the computational complexities of these overlapping generations models it is helpful to determine the relative importance of the economic factors driving this result. I highlight the effect of changing two common assumptions in a benchmark model that generates a large optimal tax on capital similar to the model in Conesa et al. (2009). First, the utility function is altered such that it implies an agent's Frisch labor supply elasticity is constant, as opposed to increasing, over his lifetime. Second, the government is allowed to tax accidental bequests at a separate rate from ordinary capital income. The main finding of this paper is that these two changes cause the optimal tax on capital to drop by almost half. Furthermore, I find that the welfare costs of adopting the high optimal tax on capital from the benchmark model in the model with the altered assumptions, which calls for a lower tax on capital, are equivalent to 0.35 percent of total lifetime consumption. Quantifying the effect of these assumptions in the benchmark model is important because the first has limited empirical evidence and the second, although included for tractability, confounds a motive for taxing capital with a motive for taxing accidental bequests.

JEL: E62, H21.

Key Words: Optimal Taxation; Capital Taxation.

1 Introduction

Receipts from taxes on individuals' capital income (capital gains and dividends) in 2005 were approximately $140 billion, or 15 percent of total personal income tax receipts.[1] Based on the sizable tax receipts from capital income in the U.S. economy and savings disincentives created by a capital tax, considerable research has been devoted to determining whether a non-zero tax on capital income is optimal.[2] In the seminal works on this topic, Chamley (1986) and Judd (1985) conclude that it is not optimal to tax capital in a model with infinitely lived agents who face no idiosyncratic risk. Atkeson et al. (1999) show that the optimal tax on capital is still zero in a two-period overlapping generations model when the government is allowed to condition the labor income tax on age. Other works, such as Aiyagari (1995), Hubbard and Judd (1986), İmrohoroğlu (1998), Erosa and Gervais (2002), Conesa et al. (2009), Garriga (2001), Jones et al. (1997) and Correia (1996), identify theoretical conditions under which it is optimal to tax capital.

When determining the optimal tax on capital, the policymaker must weigh the relevant benefits versus the distortions imposed by the tax. Since a tax on capital discourages saving, it is important to analyze the tax in an overlapping generations (OLG) model that includes the life cycle factors that motivate saving. One such study, Conesa et al. (2009) uses a calibrated life cycle model and finds that the optimal tax policy consists of flat tax rates on capital and labor income of 34 percent and 14 percent, respectively.[3] Additional studies such as Gervais (2010), Garriga (2001), Peterman (2012), Smyth (2006), and İmrohoroğlu (1998) find a non-zero tax on capital is optimal in an OLG model. Given the computational complexities of these OLG models, it is helpful to determine the economic factors driving these results. Studies that quantify the optimal tax on capital weigh the trade-off between realism and computational intensity when choosing simplifying assumptions. This paper quantifies the relative importance of two of the key modelling assumptions that motivate a positive tax on capital in a canonical OLG model. Understanding the effect of these assumptions is relevant in order to more accurately determine the optimal tax on capital.

I start by solving for optimal tax policy in a benchmark model similar to the model in Conesa et al. (2009). Next, in order to measure the assumptions effects on optimal tax policy, I solve for the optimal tax policy in an altered model in which I eliminate two commonly adopted assumptions that generate a non-zero tax on capital. First I eliminate the assumption that the Frisch elasticity varies over the life cycle.[4] Second I relax several restrictions in the benchmark model regarding how the government is allowed to tax accidental bequests. I test the effect of a varying Frisch elasticity since there is limited empirical evidence on whether the Frisch

---

[1] See www.treas.gov/offices/tax-policy/library/capgain3-2008.pdf and www.irs.gov/taxstats/indtaxstats/article/0,,id=129270,00.html.

[2] Following Conesa et al. (2009), I define an optimal tax policy as one that maximizes the expected lifetime utility of a newborn in a stationary equilibrium, holding tax revenue constant.

[3] This is model M4 in Conesa et al. (2009), which excludes idiosyncratic risk.

[4] The Frisch labor supply elasticity is the labor supply elasticity holding the marginal utility of wealth constant.

labor supply elasticity varies over the lifetime.[5] Therefore, it is important to understand the impact of this assumption on optimal tax policy. The restrictions regarding taxing accidental bequests are used to make the model more tractable, but these restrictions confound a motive for taxing ordinary capital with a motive to tax accidental bequests and are not consistent with actual policy.

The main finding of this paper is that these two assumptions are responsible for almost half of the positive optimal tax on capital in the benchmark OLG model. When these two assumptions are removed from the model the optimal tax on capital is reduced from approximately 30 percent to 16 percent. Additionally, I find that welfare losses are equivalent to 0.35 percent of total consumption if I implement the optimal tax policy from the benchmark model in the altered model. Altering just one of either two assumptions causes the optimal tax on capital to drop by approximately one-third. Therefore, the simplifying restrictions on the tax function regarding accidental bequests should not be included when determining optimal tax policy. These results also indicate that to more precisely determine the optimal tax on capital, one needs to empirically determine if the Frisch elasticity varies over the life cycle.

A varying Frisch elasticity over the life cycle motivates a positive optimal tax on capital because it causes the government to want to condition labor income taxes on age. If the government is disallowed from using age-dependent taxes, then a non-zero tax on capital can be used to mimic age-dependent taxes since a capital tax implicitly taxes younger labor income at a relatively higher rate. In a related work, Gervais (2010) demonstrates that a progressive labor income tax can also be used in tandem with a tax on capital to mimic an age-dependent tax policy.[6] The benchmark utility function in Conesa et al. (2009) is non-homothetic in labor, which implies that the Frisch elasticity varies over the life cycle with hours worked.[7] Therefore, in order to test the impact of this assumption, I determine the effect on optimal tax policy of changing the utility function such that it is homothetic in labor, which implies that the Frisch elasticity does not vary.[8]

Restricting how the government can tax accidental bequests confounds a motive for a non-zero tax on ordinary capital income. In the benchmark model, it is assumed that the government cannot distinguish accidental bequests from ordinary capital, which implies that the government has to tax the returns from both sources at the same rate. Additionally, the government is restricted to taxing only the return on the accidental bequests and not the principal. Since accidental bequests are inelastic income, the government would like to fully tax them. If they cannot distinguish between the two incomes, the optimal tax on capital is a mix of the

---

[5]Two exceptions are French (2005) and Clark and Summers (1981).

[6]In a similar model to mine that excludes any within cohort heterogeneity, Conesa et al. (2009) find that it is optimal to only use a tax on capital to mimic an age-dependent tax system. I find similar results in my model. Gervais (2010) finds a different result. He finds that the optimal age-independent tax system includes a both a large tax on capital and a progressive tax on labor income.

[7]Formally Erosa and Gervais (2002) demonstrate that in order to eliminate the motive to condition labor income taxes on age, the utility function must be both separable and homothetic in both consumption and labor.

[8]In order to fully eliminate this motive, the utility function must be separable and homothetic in both labor and consumption. Since the benchmark utility function is already separable, I focus on the homotheticity assumption. An alternative way to assess the impact of this assumption would be to allow the government to condition labor income taxes on age.

optimal tax on ordinary capital income and the optimal tax on accidental bequests. I test the effect of relaxing these tax restrictions by allowing the government to separately tax accidental bequests and ordinary capital income.

Given that these two assumption motivate approximately half of the optimal tax on capital in the benchmark model, it becomes relevant to quantify the individual effect of all the modelling features that could motivate a non-zero tax on capital within a common framework. Four common features in an OLG model that motivate a non-zero optimal tax on capital are: (i) a varying lifetime Frisch labor supply elasticity, (ii) restrictions on how the government can tax accidental bequests, (iii) the inability of individuals to borrow, and (iv) the inability of the government to facilitate a social security program.[9] I solve for the optimal tax policy in four other models with one of the four features that motivate a non-zero optimal tax on capital changed in order to determine the effect of each feature. Additionally, I solve for the optimal tax policy with an exogenously determined level of government debt or savings in order to ascertain its effect on optimal tax policy.[10]

In addition to the the non-constant Frisch elasticity and the government not being able to separately tax accidental bequests, I find that individual liquidity constraints also motivates a positive tax on capital but to a lesser extent. I find that assuming the government holds savings or debts has a dramatic effect on the optimal tax on capital. When I assume that the government holds savings (debt), the optimal tax on capital decreases (increases) a significant amount. There are only small changes to the optimal tax policy when I exclude the reduced-form social security program from the benchmark model; however, the life cycle profiles look less realistic. I find that the welfare cost of implementing the benchmark model's optimal tax policy in each of these alternative models ranges from .08 percent to 2.53 percent of total lifetime consumption. These results demonstrate that for some models there are large welfare consequences from adopting the optimal tax policy derived from a different model. Finally, this paper analyzes how the effect of the features change when the model is calibrated to match different targets for the Frisch elasticity since there is a large variance in the empirical estimates of this value. Generally, I find that these five features have a larger effect on the optimal tax on capital when the model is calibrated to match a medium or low Frisch elasticity as opposed to a high value.

This exercise is related to Conesa et al. (2009), however it includes three important differences. First, I exclude inter-cohort heterogeneity as a possible motive for a positive tax on capital. I abstract from this type of heterogeneity because Conesa et al. (2009) demonstrate that it does not affect the level of the optimal tax on capital.[11] Second, I examine how relaxing the restrictions on taxing accidental bequest affects optimal tax

---

[9]See Conesa et al. (2009), Smyth (2006), Conesa and Krueger (2006), Guvenen et al. (2009), Fuster et al. (2008), Garriga (2001), Erosa and Gervais (2002), and Nakajima (2010) for examples of papers that include similar assumptions when analyzing tax policy in an OLG framework.

[10]I test both savings and debt hence, I explore two additional models for a total of six models in addition to the benchmark.

[11]The authors find that including idiosyncratic uninsurable income shocks and productivity differences affect the progressivity of the optimal labor tax policy but not the optimal level of the tax on labor or capital. Gervais (2010) confirms that without

policy. The effects of these restrictions are not studied in Conesa et al. (2009). Third, I take an alternative approach to discern the effect of a varying Frisch elasticity. Conesa et al. (2009), similar to my benchmark model, use a utility specification in which the agent's Frisch labor supply elasticity is negatively related to hours worked. To determine the effect of a varying Frisch elasticity on optimal tax policy, the authors eliminate the variance by holding the labor supply exogenously constant. Using this approach eliminates any general equilibrium effects of endogenously determined labor supply on optimal tax policy. Instead, in this paper, I eliminate the variation in the Frisch elasticity by using a utility specification that implies that the Frisch labor supply elasticity is constant.[12] The advantage of this approach is that it isolates the effect of a constant Frisch while including general equilibrium effects from endogenously determined labor supply. In another related exercise, Garriga (2001) examines the effect on optimal tax policy of altering the social discount factor in a model where the government endogenously determines the level of savings or debt it holds. He finds that as the government decreases their discount rate, the optimal tax on capital increases. In contrast, this paper follows Conesa et al. (2009) and focuses on the relative effect of the five other common features on optimal tax policy when the government maximizes ex-ante welfare of a newborn in the steady state in a model where the level of government debt is exogenously determined.

This paper is organized as follows: Section 2 examines a simplified version of the model in order to provide analytical insights into the effect of the two channels eliminated in the alternative specification. I introduce the computational model and present the competitive equilibrium in section 3. Section 4 describes the functional forms and calibration parameters. Section 5 sets up the computational experiment, and section 6 reports the results of the computational experiment. Section 7 examines the sensitivity of the results with respect to the target that the Frisch elasticity is calibrated to match. Finally, section 8 summarizes the paper's findings.

2 Analytical Model

In this section, I demonstrate the intuition in an analytically tractable model for why the two main assumptions examined in this paper motivate a non-zero optimal tax on capital. The two assumptions I examine are a non-constant Frisch elasticity when the government cannot condition labor income taxes on age and the government not being able to distinguish between accidental bequests and ordinary capital.[13] In the analytically tractable model, I abstract from retirement, population growth, and progressive tax policies. Additionally, I assume that

---

idiosyncratic risk the optimal tax on capital can be large. Therefore, this paper abstracts from these sources of heterogeneity and focuses on models where agents are homogenous within the cohort.

[12]Formally, the condition to eliminate the desire to mimic age-dependent taxes is that the utility function is both separable and homothetic in consumption and labor. In order to focus on the effect of homotheticity, I focus on the separable utility function that Conesa et al. (2009) use in their sensitivity analysis.

[13]Conesa et al. (2009), Garriga (2001), Erosa and Gervais (2002) and Atkeson et al. (1999) demonstrate similar analytic results with regards to the first assumption. More specifically, this second assumption implies that the government is restricted to taxing accidental bequests in the same manner and at the same rate as ordinary capital.

the marginal products of capital and labor are constant. This assumption permits me to focus on the life cycle elements of the model, in that changes to the tax system do not affect the pre-tax wage or rate of return. Since the factor prices do not vary, I suppress their time subscripts in this section. All of these assumptions are relaxed in the computational model. Using the primal approach, I solve for the optimal tax policies in several versions of the simplified model in order to isolate the effect of the two different assumptions.

2.1 Agent's Problem

The analytical model is a simplified model where agents live for two periods, and their preferences over consumption and leisure are given by

$$U(c_{1,t}, 1 - h_{1,t}) + \Psi\beta U(c_{2,t+1}, 1 - h_{2,t+1}) \tag{1}$$

where β is the discount rate, Ψ is the probability of survival to the next period, $c_{j,t}$ is the consumption of an age j agent at time t, and $h_{j,t}$ is the percent of his time endowment he works.[14] Age-specific human capital is normalized to unity when the agent enters the model. At age two, age-specific human capital is ϵ_2. The agent maximizes equation 1 with respect to consumption and hours subject to the following constraints:

$$c_{1,t} + a_{1,t} = (1 - \tau_{h,1})h_{1,t}w_t + Tr_t(1 + r(1 - \tau_k)) \tag{2}$$

and

$$c_{2,t+1} = (1 + r_t(1 - \tau_k))(a_{1,t} + Tr_{t+1}) + (1 - \tau_{h,2})\epsilon_2 h_{2,t+1}w_{t+1}, \tag{3}$$

where $a_{1,t}$ is the amount saved at age 1, Tr_t is an accidental bequests that all living agents receive, $\tau_{h,j}$ is the tax rate on labor income for an agent of age j, τ_k is the tax rate on capital income, w_t is the efficiency wage for labor services, and r_t is the rental rate on capital. Accidental bequests are only non-zero when the model includes lifetime length uncertainty ($\Psi < 1$). When lifetime length uncertainty is introduced, some agents will die while holding savings.

A common treatment of these accidental bequests in OLG models is that when individuals die their assets are redistributed to all living agents (see Conesa et al. (2009) for an example). The living agents receive both the original assets (principal) and the capital returns on these assets. One tractable assumption that has been employed is to restrict the tax policy such that the government cannot distinguish between accidental bequests and ordinary capital nor can they tax the principal of the accidental bequests. Additionally, the government is forced to tax the return on both ordinary capital and accidental bequests at the same rate. In section 2.4, I examine the effect of these restrictions. I assume that the tax rate on labor income can be conditioned on age

---

[14]Time working is measured as a percentage of endowment and not in hours. However, for convenience, I refer to $h_{j,t}$ as hours.

in some of the models; however, the tax rate on capital income cannot.[15] I combine equations 2 and 3 to form a joint intertemporal budget constraint, which is necessary for the primal approach,

$$c_{1,t} + \frac{c_{2,t+1}}{1 + r_t(1 - \tau_k)} = w_t(1 - \tau_{h,1})h_{1,t} + \frac{w_{t+1}(1 - \tau_{h,2})\epsilon_2 h_{2,t+1}}{1 + r_t(1 - \tau_k)} + Tr_t(1 + r(1 - \tau_k)) + Tr_{t+1}. \quad (4)$$

The agent's problem is to maximize equation 1 subject to 4. The agent's first-order conditions are

$$\frac{U_{h1}(t)}{U_{c1}(t)} = -w_t(1 - \tau_{h,1}), \quad (5)$$

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -w_{t+1}\epsilon_2(1 - \tau_{h,2}), \quad (6)$$

and

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1 + r_t(1 - \tau_k)), \quad (7)$$

where $U_{c1}(t) \equiv \frac{\partial U(c_{1,t}, 1 - h_{1,t})}{\partial c_{1,t}}$. Given prices and taxes, these first-order conditions together with the intertemporal budget constraint determine the optimal allocation of $(a_{1,t}, c_{1,t}, h_{1,t}, c_{2,t+1}, h_{2,t+1})$.

2.2 Primal Approach

To determine the optimal tax policy, I use the primal approach. In the primal approach the benevolent government maximizes directly over allocations, discounting future generations with social discount factor θ, subject to the implementability constraint. The optimal tax policy is reverse-engineered from the optimal allocations. The implementability constraint is the agent's intertemporal budget constraint with the prices and taxes replaced by his first-order conditions (equations 5, 6, and 7). Including this constraint ensures that any allocation the government chooses can be supported by a competitive equilibrium.[16]

Formally, the government maximizes the objective function,

$$[U(c_{2,0}, 1 - h_{2,0})/\theta] + \sum_{t=0}^{\infty} \theta^t [U(c_{1,t}, 1 - h_{1,t}) + \beta U(c_{2,t+1}, 1 - h_{2,t+1})], \quad (8)$$

with respect to the implementability constraint,

$$c_{1,t}U_{c1}(t) + \beta c_{2,t+1}U_{c2}(t+1) + h_{1,t}U_{h1}(t) + \beta h_{2,t+1}U_{h2}(t+1) - Tr_t\frac{U_{c1}^2}{\beta \Psi U_{c2}} - U_{c1}Tr_{t+1} = 0 \quad (9)$$

---

[15]In a two-generation model, only young agents save, so this restriction would not bind.

[16]See Lucas and Stokey (1983) or Erosa and Gervais (2002) for a more-in depth discussion of the primal approach.

and the resource constraint,

$$c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t = rK_t + w(h_{1,t} + h_{2,t}\epsilon_2). \tag{10}$$

In section 2.3.2, I restrict the government from conditioning labor income taxes on age. In such a model the following constraint must be included:

$$\epsilon_2 \frac{U_{h1}(t)}{U_{c1}(t)} = \frac{U_{h2}(t+1)}{U_{c2}(t+1)}. \tag{11}$$

I use ρ as the Lagrange multiplier on the resource constraint, λ as the Lagrange multiplier on the implementability constraint, and η as the Lagrange multiplier on the additional constraint when age-dependent taxes are dissallowed. I assume the production technology is such that the marginal products of capital and labor are constant.[17] This assumption allows me to focus on the life cycle elements of the model because changes to the tax system do not affect the pre-tax wage or rate of return. Since the factor prices do not vary I suppress the time subscripts on these prices.

2.3 Effect of Non-Homothetic Utility Function

Garriga (2001) and Erosa and Gervais (2002) demonstrate that if the government cannot condition taxes on age, then in order for the optimal tax on capital to be zero, the utility function must be separable and homothetic in both consumption and labor. An implication of a non-homothetic utility function is the Frisch elasticity is not constant over the lifetime. Although the violation of the primal assumptions that motivates a non-zero tax on capital is a non-homothetic utility function, I refer to the modelling assumption as a non-constant Frisch elasticity since the intuition for the result comes from this variation. In this section, I demonstrate the intuition for why a varying Frish elasticity coupled with the government's inability to condition labor income taxes on age motivates a positive optimal tax on capital.

To demonstrate the intuition, I consider the optimal tax policy in a world where agents live with certainty for two periods ($\Psi = 1$, and $Tr = 0$) under two different utility functions:

$$U_{\text{constant Frisch}} = \frac{c^{1-\varsigma_1}}{1-\varsigma_1} - \chi \frac{(h)^{1+\frac{1}{\varsigma_2}}}{1+\frac{1}{\varsigma_2}}$$

$$U_{\text{non-constant Frisch}} = \frac{c^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h)^{1-\sigma_2}}{1-\sigma_2}.$$

The first utility function is homothetic in both consumption and hours worked, and the Frisch labor supply elasticity, ς_2, is not a function of time worked but instead is constant throughout the agent's life. The second

---

[17]As previously noted, in the computational model I relax this assumption.

utility function, which is non-homoethetic, has a varying Frisch labor supply elasticity, $\frac{1-h}{\sigma_2 h}$, as long as hours are not constant.

2.3.1 Non-Constant Frisch Utility

I start by solving for the model with the non-constant Frisch utility function where the government can condition labor income taxes on age. The formulation of the government's problem, resulting first order conditions, and derivations of the optimal tax policy can be found in appendix B.1.

Conesa et al. (2009), Atkeson et al. (1999), Erosa and Gervais (2002), and Garriga (2001) demonstrate with this type of utility function and age-dependent taxes that the optimal tax policy includes no tax on capital and different tax rates on different-aged labor income. Combining the agent's and government's first order conditions simplifies to the following expression for optimal labor income taxes:

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1+\frac{\sigma_2 h_{1,t}}{1-h_{1,t}})}{1+\lambda_t(1+\frac{\sigma_2 h_{2,t+1}}{1-h_{2,t+1}})}. \tag{12}$$

Equation 12 confirms that in a model with the non-constant Frisch utility function, the optimal tax on labor income varies by age as long as $h_{1,t} \neq h_{2,t+1}$. In the steady state, if $h_1 > h_2$, then the optimal tax on labor income is such that $\tau_{h,1} > \tau_{h,2}$. Recall that Frisch labor supply elasticity for the non-constant Frisch utility function is $\frac{1-h}{\sigma_2 h}$. Therefore, one can interpret this result as the optimal tax policy is such that $\tau_{h,1} > \tau_{h,2}$ when the Frisch labor supply elasticity rises over the agent's lifetime. The government prefers a higher tax on the labor that is supplied less elastically, as it limits the distortions imposed by the tax policy.

Furthermore, Conesa et al. (2009), Atkeson et al. (1999), Erosa and Gervais (2002), and Garriga (2001) demonstrate that in a similar set up, if the government cannot use age-dependent taxes, then the optimal tax on capital is no longer zero. The Lagrangian and first order conditions for this model where the government cannot use age-dependent taxes are in appendix B.2. Combining the governments first order conditions with respect to capital and consumption leads to the following expression:

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r)\left(\frac{1+\lambda_t(1-\sigma_1)-\frac{\eta_t\epsilon_2\sigma_1(1-h_{1,t})^{-\sigma_2}}{c_{2,t+1}}}{1+\lambda_t(1-\sigma_1)+\frac{\eta_t\sigma_1(1-h_{2,t+1})^{-\sigma_2}}{c_{1,t}}}\right). \tag{13}$$

In contrast, applying the non-constant Frisch utility function to equation 7 provides the following relationship:

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r(1-\tau_k)). \tag{14}$$

Equations 14 and 13 demonstrate that generally for the government to ensure the agent chooses the optimal

allocation, the tax on capital is not zero. In the absence of the ability to condition labor income taxes on age, the government chooses to tax capital in order to mimic an age-dependent tax on labor income.

Examining the Euler equation provides the intuition for why a tax on capital mimics an age-dependent tax on labor:

$$\epsilon_2 \frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1 + r(1 - \tau_k))\frac{1 - \tau_{h,1}}{1 - \tau_{h,2}}. \tag{15}$$

Equation 15 demonstrates that a tax on capital imperfectly mimics an age-dependent tax on labor income by creating a similar wedge on the marginal rate of substitution.[18] Specifically, a positive tax on capital mimics a relatively higher tax rate on young labor income since it creates a similar outcome on the left-hand side of equation 15.

2.3.2 Constant Frisch Utility and Age-Dependent Taxes

Next, I examine the optimal tax policy in a model with the constant Frisch utility function where the government is allowed to condition labor income taxes on age. Comparing the resulting optimal tax policy with the policy in section 2.3.1 isolates the effect of a varying Frisch elasticity on optimal tax policy. The formulation of the government's problem and their first order conditions for this model can be found in appendix B.3

Combining the agent's and government's first order equations generates the following expression for the optimal labor taxes:

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t(1 + \frac{1}{\varsigma_2})}{1 + \lambda_t(1 + \frac{1}{\varsigma_2})} = 1. \tag{16}$$

Equation 16 demonstrates that even if the government could condition taxes on age then they would tax labor income for different-aged individuals at the same rate. Under the constant Frisch utility specification, it is not optimal to vary the labor income tax rate based on age since the Frisch elasticity is constant. Therefore, using this utility function eliminates the motive of taxing capital in order to mimic an age-dependent tax when the age-de3pendent tax is unavailable.[19]

2.4 Effect of Incomplete Set of Tax Instruments on Accidental Bequests

In this section, I demonstrate the intuition why the assumption that the government cannot distinguish between accidental bequests and ordinary capital leads to a positive optimal tax on capital. In all of the models in this section (2.4) I use the constant Frisch utility function and allow for age-dependent taxes on labor income, which eliminates all of the motives for a non-zero tax on capital discussed in section 2.3. In this section, the models

---

[18] A non-zero tax on capital can only imperfectly mimic age-dependent taxes on labor income because the former provides one less degree of freedom so the government can no longer independently determine both the wedge and the overall revenue from the tax policy.

[19] See Erosa and Gervais (2002) and Garriga (2001) for a proof of a more general version of this result.

include lifetime length uncertainty so some agents die with positive levels of savings. These accidental bequests are transferred to living individuals. The assumption that the government cannot distinguish between ordinary capital and accidental bequests implies two restrictions. First, the government is restricted to only taxing the return on the accidental bequests and not the principal. Second, they are forced to tax the returns on accidental bequests at the same rate as the returns on ordinary capital. To demonstrate the intuition why these restrictions motivate a positive tax on capital, I start by solving for the optimal tax policy in an unrestricted model. I then examine the change to the optimal tax policies when I incrementally add the two restrictions to the tax policy.

2.4.1 No Restrictions on Accidental Bequest Tax Policy

I start by solving for the optimal tax policy in a simple model without restrictions on the accidental bequest tax policy. Specifically, the government taxes ordinary capital income, ra, at a rate of τ_k, and accidental bequests, $(1+r)Tr$, at a rate of τ_t. The intertemporal budget constraint, equation 4, becomes

$$c_{1,t} + \frac{c_{2,t+1}}{1+r(1-\tau_k)} = w(1-\tau_{h,1})h_{1,t} + \frac{w(1-\tau_{h,2})\epsilon_2 h_{2,t+1}}{1+r(1-\tau_k)} + Tr_t(1+r)(1-\tau_t) + Tr_{t+1}\frac{(1+r)(1-\tau_t)}{1+r(1-\tau_k)}. \quad (17)$$

The formulation of the government's problem and first order conditions can be found in appendix B.4. Combining the government's first order conditions for consumption and saving yields the following expression:

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1} = \Psi\beta(1+r)\left[\frac{1+\lambda_t(1-\varsigma_1) - c_{2,t+1}^{-1}\lambda_t\varsigma_1 Tr_{t+1}(1+r)(1-\tau_t)}{1+\lambda_t(1-\varsigma_1) - c_{1,t}^{-1}\lambda_t\varsigma_1 Tr_t(1+r)(1-\tau_t)}\right]. . \quad (18)$$

Combining the government's first order conditions for labor and savings yields the following expression for the optimal taxes on labor:

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1-\varsigma_1) - c_{2,t+1}^{-1}\lambda_t\varsigma_1 Tr_{t+1}(1+r)(1-\tau_t)}{1+\lambda_t(1-\varsigma_1) - c_{1,t}^{-1}\lambda_t\varsigma_1 Tr_t(1+r)(1-\tau_t)} \quad (19)$$

Examining equation 18, if the government fully taxes the principal on accidental bequests ($\tau_t = 1$), the equation simplifies to the same as equation 14 and the optimal tax on capital is zero. Additionally, if $\tau_t = 1$, then the right-hand side of equation 19 simplifies to one, and it is not optimal for the government to condition labor income taxes on age. Since accidental bequests are inelastic income, it is optimal for the government to fully tax these before using other forms of taxation that are distortionary. Therefore, if the government can tax the principal and return of the accidental bequests at a separate rate from ordinary capital income, then the optimal tax policy is similar to the policy in section 2.3.2; they do not include a tax on ordinary capital income nor age-dependent taxes on labor income.

2.4.2 Restricted Tax Policy: Cannot Tax Principal of Accidental Bequests

In this section, I solve for the optimal tax policy when the government can no longer tax the principal of accidental bequests. However, I still allow the government to set separate tax rates for the returns on accidental bequests (τ_t) and ordinary capital (τ_k). The intertemporal budget constraint in this model is

$$c_{1,t} + \frac{c_{2,t+1}}{1 + r(1 - \tau_k)} = w(1 - \tau_{h,1})h_{1,t} + \frac{w(1 - \tau_{h,2})\epsilon_2 h_{2,t+1}}{1 + r(1 - \tau_k)} + Tr_t(1 + r(1 - \tau_k)) + Tr_{t+1}\frac{(1 + r(1 - \tau_t))}{1 + r(1 - \tau_k)}. \quad (20)$$

The formulation of the government's problem and first order conditions can be found in appendix B.5.

Combining the government's first order conditions for consumption and savings yields the following expression:

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1} = \Psi\beta(1 + r)\left[\frac{1 + \lambda_t(1 - \varsigma_1) - c_{2,t+1}^{-1}\lambda_t\varsigma_1 Tr_{t+1}(1 + r(1 - \tau_t))}{1 + \lambda_t(1 - \varsigma_1) - c_{1,t}^{-1}\lambda_t\varsigma_1 Tr_t(1 + r(1 - \tau_t))}\right]. \quad (21)$$

Examining equation 21 and equation 14 demonstrates that generally the optimal tax on capital is no longer zero. Even if $\tau_t = 1$, individuals still receive some income from accidental bequests in the form of the principal. If they were not restricted from doing so, the government would like to fully tax the principal of the accidental bequests because it is perfectly inelastically supplied. Since young agents save these bequests in the form of ordinary capital, the government can indirectly tax these accidental bequests by taxing ordinary capital income. Hence, a non-zero tax on ordinary capital income is optimal.

Equation 22 is an expression for the optimal tax on labor income,

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t(1 - \varsigma_1) - c_{2,t+1}^{-1}\lambda_t\varsigma_1 Tr_{t+1}(1 + r(1 - \tau_t))}{1 + \lambda_t(1 - \varsigma_1) - c_{1,t}^{-1}\lambda_t\varsigma_1 Tr_t(1 + r(1 - \tau_t))}. \quad (22)$$

When the government is restricted from taxing the principal of accidental bequests, it now wants to condition labor income taxes on age. When the government taxes the return on ordinary capital it creates a wedge on the marginal rate of substitution. The government would like to condition labor income taxes on age in order to unwind that wedge. Combining equations 21, 43, and 22 yields the following ratio,

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + r(1 - \tau_k)}{1 + r}. \quad (23)$$

This expression demonstrates that the ratio of the wedges created by the labor taxes and the tax on capital are equal. As the tax on capital income increases, it is optimal to also decrease the relative tax on young labor income to unwind this wedge because the tax on capital is like a relatively larger tax on young labor.

2.4.3 Restricted Tax Policy: Cannot Observe Accidental Bequests

I finish by exploring the optimal tax policy in a model where the government cannot distinguish between income from ordinary capital and accidental bequests. This assumption implies that the government only observes the overall returns, $r(a + Tr)$, and it is restricted to tax all the returns at the same rate, τ_k. This treatment of accidental bequests is similar to Conesa et al. (2009). The agent's intertemporal budget constraint is the same as equation 4. The formulation of the government's problem and first order conditions can be found in appendix B.6

Utilizing the first order conditions from the Lagrangian with respect to capital and consumption leads to the following equation:

$$
\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1} = \Psi\beta(1+r)\left[\frac{1 + \lambda_t(1-\varsigma_1) - \frac{Tr_{t+1}\lambda_t\varsigma_1}{\beta\Psi}^{-2}c_{2,t+1}^{-1}c_{1,t}^{-2\varsigma_1}}{1 + \lambda_t(1-\varsigma_1) + \lambda_t\varsigma_1(Tr_tc_{1,t}^{-1} + \frac{Tr_{t+1}2c_{2,t+1}^{\varsigma_1}c_{1,t}^{-1-\varsigma_1}}{\beta\Psi})}\right]. \tag{24}
$$

Comparing equations 24 and 14 demonstrates that generally the optimal tax on capital will not be zero. Additionally, assuming λ_t is positive, then as the size of Tr increases, the optimal tax on capital also increases. Disallowing the government from taxing the returns on accidental bequests and capital at different rates creates an additional motive for a non-zero tax on capital.[20] Once again, the government would like to fully tax the whole accidental bequest (principal and return) since it is inelastic income. In this model, since the tax on capital income is a hybrid tax on returns from ordinary capital income and accidental bequests, the optimal tax rate is a weighted average of the optimal rates on each income. The modelling assumption that the government cannot tax capital and accidental bequests at separate rates creates an additional motivation for a non-zero optimal tax on capital.

2.5 Summary of Analytic Results

In this section, I demonstrate that with a non-constant Frisch elasticity profile, it is optimal to condition labor income taxes on age. Furthermore, if the government cannot use age-dependent taxes, then it is optimal to tax capital to mimic an age-dependent labor income tax. This motive for a non-zero tax on capital is eliminated in a model that uses the constant Frisch utility function. I also demonstrate that restricting the government from setting separate tax rates on ordinary capital and accidental bequests or disallowing them from taxing the principal of the accidental bequests leads to a non-zero optimal tax on capital.

---

[20]Analytically, the additional motives can be seen by comparing equations 24 and 21. Equation 24 has additional terms causing the ratio inside the brackets to not be equal to one.

3 Computational Model

To examine the relevant magnitude of the motives for a non-zero tax on capital, I examine a less parsimonious calibrated overlapping generations model that must be solved computationally. In this section, I describe the computational model (focusing on the benchmark model) and the definition of a stationary competitive equilibrium.

3.1 Demographics

In the computational 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. Agents are forced to retire at an exogenously set age 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

$$\mu_j = \frac{\Psi_{j-1}}{1+n}\mu_{j-1}, \text{for } i = 2, ..., J, \tag{25}$$

where μ_1 is normalized such that

$$\sum_{j=1}^{J}\mu_j = 1. \tag{26}$$

If agents die before J, these assets are treated as accidental bequests. In the benchmark model, the government is restricted to taxing just the returns on accidental bequests and at the same rate that it taxes returns on ordinary capital income. In order to test the strength of the motive for a non-zero tax on capital described in section 2.4, I also use a second treatment. In the second treatment, I allow the government to consume these assets removing this motive for a positive tax on capital.

3.2 Individual

An individual is endowed with one unit of productive time per period, which he splits between providing labor services and leisure in order to maximize his lifetime utility

$$\{\sum_{j=1}^{J}\beta^j \prod_{q=1}^{j-1}\Psi_{q-1}u(c_j, h_j)\}, \tag{27}$$

where c_j is the consumption of an agent at age j and h_j is the hours spent providing labor services. Agents discount the next period's utility by the product of Ψ_j and β. The discount factor conditional on surviving is β, and the unconditional discount rate is $\beta\Psi_j$.

14

An agent's age-specific human capital is ϵ_j so he receives labor income of $h_j\epsilon_j w_t$. Agents split their labor income between consumption and saving. An agent can save by purchasing a risk-free asset. An agent's level of assets are denoted by a_j, and he receives a pre-tax net return of r_t on the assets per period. Agents with binding liquidity constraints early in their life poses another potential motive for a positive tax on capital. In some of the iterations, I test this motive's strength by allowing agents to borrow. In these iterations of the model, agents pay the actuarially fair interest rate of $r_{b,j,t} = \frac{r_t}{\Psi_j}$ to borrow.

3.3 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

$$C_t + K_{t+1} - (1 - \delta)K_t + G_t \leq K_t^\alpha N_t^{1-\alpha}, \tag{28}$$

where K_t, C_t, and N_t represent the aggregate capital stock, aggregate consumption, and aggregate labor (measured in efficiency units), respectively. Additionally, α is the capital share and δ is the depreciation rate for physical capital.

3.4 Government Policy

The government consumes resources in an unproductive sector, G_t.[21] The government has two fiscal instruments to finance their consumption in the benchmark model ($G_t = T^K[y_k]+T^l[y_l]$). First, the government taxes capital income, y_k, according to a capital income tax schedule $T^K[y_k]$. In the benchmark, $y_k \equiv r_t(a + Tr_t)$. In other versions when the government can distinguish accidental bequests, $y_k \equiv r_t(a)$. Second, the government taxes each individual's taxable labor income. Part 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 is $y_l \equiv w_t s_j h_j(1 - .5\tau_{ss})$, which is taxed according to a labor income tax schedule $T^l[y_l]$. I impose two 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.

In some iterations of the model I force the government to borrow or save in order to quantify its effect on optimal tax policy. I solve for the optimal tax policies under a steady state equilibrium so the government's level of savings or debt cannot change over time. Therefore, in these iterations the government holds a fixed level of savings or debt but is still not allowed to run a deficit or surplus. When the government holds savings,

---

[21]A formulation that induces the same optimal tax policy is if the G_t enters the agents utility function in an additively separable manner.

the return on its capital is used to offset government consumption, and when the government is in debt it pays the interest on its debt by reducing government consumption. This assumption implies that the government is collecting the same amount of tax revenue regardless of the level of government savings or debt.

In addition to taxing income to finance G_t, the government runs a pay-as-you-go social security system in the benchmark model. The government pays SS_t to all individuals that are retired. Social security benefits are such that retired agents receive an exogenously determined fraction, b_t, of the average income of all working individuals. An agent's social security benefits are independent of his personal earnings history. Social security 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 other iterations of the model, I eliminate the social security program to determine its effect on the optimal tax policy.

3.5 Definition of Stationary Competitive Equilibrium

In this section, I define the competitive equilibria for the benchmark model. I do not present the definition of a competitive equilibrium for the other iterations of the model since they are similar to that of the benchmark model.

Given a social security replacement rate b, government expenditures G, and a sequence of population shares $\{\mu_j\}_{j=1}^{J}$, a stationary competitive equilibrium is a sequence of agent allocations, $\{c_j, a_{j+1}, h_j\}$, a production plan for the firm (N, K), a government labor tax function $T^l : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a social security tax rate τ_{ss}, a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, social security benefits SS, prices (w, r), and transfers Tr such that:

1. Given prices, policies, transfers, and benefits the agent maximizes the following:

$$\sum_{j=1}^{J} \text{Max}_{c_j, h_j, a_{j+1}} \beta^{j-1} [\prod_{q=0}^{j-1} \Psi_q] u(c_j, h_j)$$

subject to

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

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

$$c \geq 0, 0 \leq h \leq 1, a_j \geq 0, \text{ and } a_1 = 0.$$

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

Model	Non-Const. Frisch	Gov't Distributes Tr	Ind. Liq. Constraints	No Gov't Savings	No Gov't Borrowing	SS Program
A0 (Benchmark):	Yes	Yes	Yes	Yes	Yes	Yes
B0 (Alt. Specification):	No	No	Yes	Yes	Yes	Yes
A1 (Constant Frisch):	No	Yes	Yes	Yes	Yes	Yes
A2 (Gov't Consumes Tr):	Yes	No	Yes	Yes	Yes	Yes
A3 (Ind. Borrowing):	Yes	Yes	No	Yes	Yes	Yes
A4 (Gov't Saves):	Yes	Yes	Yes	No	Yes	Yes
A5 (Gov't Borrows):	Yes	Yes	Yes	Yes	No	Yes
A6 (No SS):	Yes	Yes	Yes	Yes	Yes	No

3. The social security policies satisfy:

$$SS = b\frac{wN}{\sum_{j=1}^{j_r-1} \mu_j} \text{ and } \tau_{ss} = \frac{ss \sum_{j=j_r}^{J} \mu_j}{w \sum_{j=1}^{j_r-1} \epsilon_j \mu_j}.$$

4. Transfers are given by:

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

5. Government budget balance:

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

6. Market clearing:

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

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

4   Calibration

In this section, I describe the the functional forms and calibration. Calibration involves two steps. The first step is choosing parameter values for which there are direct estimates in the data. Second, in order to calibrate the remaining parameters, I choose values such that under the baseline-fitted U.S. tax policy certain target values are the same in the models and the U.S. economy. I calibrate these parameters separately in the different iterations of the model.[22] Table 1 lists the different models and their features. Table 2 lists all the parameter values that are the same in the models. Table 3 lists the values for the parameters that are calibrated separately in each model.

---

[22]Choosing this approach is representative of the economists comparing models with different assumptions. In contrast, holding the calibration parameters constant between the models would be representative of the government making policy changes in an existing economy.

Table 2: **Calibration Parameters**

Parameter	Value	Target
Demographics		
Retire Age: j_r	65	By Assumption
Max Age: J	100	By Assumption
Surv. Prob: Ψ_j	Bell and Miller (2002)	Data
Pop. Growth: n	1.1%	Data
Preferences		
Risk aversion: σ_1, ς_1	2	Conesa et al. (2009)
Technology		
Capital Share: α	.36	Data
Depreciation: δ	8.33%	$\frac{I}{Y} = 25.5\%$
Productivity: A	1	Normalization
Government		
Tax Function: Υ_0	.258	Gouveia and Strauss (1994)
Tax Function: Υ_1	.768	Gouveia and Strauss (1994)

Table 3: **Calibration Parameters**

Model	Conditional Discount β	Frisch Elasticity $\sigma_2\ \varsigma_2$	Disutility to Labor χ	Govt Spending G
Target	$\frac{K}{Y} = 2.7$	Frisch$=\frac{2}{3}$	Avg. H $= \frac{1}{3}$	17% of Y
A0 (Benchmark):	0.993	3	1.9	0.137
B0 (Alt. Specification):	0.994	$\frac{2}{3}$	39.8	0.162
A1 (Constant Frisch):	0.993	$\frac{2}{3}$	35	0.136
A2 (Gov't Consumes Tr):	0.994	3	2.2	0.164
A3 (Ind. Borrowing):	0.993	3	1.9	0.137
A4 (Gov't Saves):	0.986	3	2.07	0.137
A5 (Gov't Borrows):	0.998	3	1.81	0.136
A6 (No SS):	0.969	3	1.86	0.139

4.1 Demographics

In the model, agents are born at a real-world age of 20 that corresponds to a model age of one. Agents are exogenously forced to retire at a real-world age of 65. If an individual survives until 100 (model age 80), then he dies the next period. I use Bell and Miller (2002) to determine the conditional survival probabilities. I assume a population growth rate of 1.1 percent.

4.2 Individual

As a benchmark specification I use the non-constant Frisch utility function. To determine the effect on optimal tax policy of the desire to mimic an age-dependent tax that arises because of a varying Frisch elasticity, I also find the optimal tax policy using the constant Frisch utility function. I determine β such that the capital-to-output ratio matches U.S. data of 2.7 in the benchmark model.[23] I determine χ such that under the baseline-fitted U.S. tax policy, agents work on average one-third of their time endowment in the benchmark model. Following Conesa et al. (2009), I set $\varsigma_1 = \sigma_1 = 2$, which controls the relative risk aversion. I set $\sigma_2 = 3$ for the non-constant Frisch utility function, which implies a Frisch labor supply elasticity of two-thirds when agents are

---

[23]This is the ratio of fixed assets and consumer durable goods less government fixed assets to GDP.

working one-third of their time endowment. Under the constant Frisch utility function, I set $\varsigma = \frac{2}{3}$, which also implies a Frisch elasticity of two-thirds. Past micro-econometric studies estimate the Frisch elasticity between 0 and 0.5. For examples, see Altonji (1986), MaCurdy (1981) and Domeij and Flodén (2006). However, more recent research has suggested that these estimates may be biased downward. Some of the reasons for the bias are: utilizing weak instruments, not accounting for borrowing constraints, disregarding the life cycle impact of endogenous-age specific human capital, and omitting correlated variables such as wage uncertainty. Some of these studies include Imai and Keane (2004), Wallenius (2011), Domeij and Flodén (2006), Pistaferri (2003), and Contreras and Sinclair (2008). Rogerson and Wallenius (2009) show that because individuals choose their labor supply on both the intensive and extensive margin "micro and macro elasticities need not be the same, and that macro elasticities can be significantly larger." Furthermore, Chetty (2009) shows that small frictions in the labor market can lead the observed Frisch elasticity to be much smaller. Since there is some uncertainty about this value, I test the sensitivity of the results with regards to this parameter in section 7. I calibrate $\{\epsilon_j\}_{j=0}^{j_r-1}$ such that the sequence matches a smoothed version of the relative hourly earnings estimated by age in Hansen (1993).

I focus on the effect of a varying Frisch elasticity since limited empirical evidence is available on whether the Frisch labor supply elasticity varies over the lifetime. Two exceptions are French (2005) and Clark and Summers (1981). French (2005) estimate that the labor supply elasticity is more than three times larger for 60-year-old individuals than 40-year-old individuals. However, the author notes that social security and pension incentives are responsible for this change. Therefore, the change in elasticity results from changes on the extensive margin and not the intensive margin. The focus of this study is changes in labor supply on the intensive margin since retirement is considered exogenous in the model. Clark and Summers (1981) suggest that teenagers may be more elastic than prime-aged workers. However, teenage workers are outside the scope of this study since agents enter the model once they are in their 20s.

4.3 Firm

I assume the capital share parameter, α, is 0.36. The depreciation rate is set to target the observed investment-output ratio of 25.5 percent.

4.4 Government Policy

To calibrate the parameters, I need a benchmark tax function to use when matching the targets in the models to the values in the data. I calibrate the model under a baseline tax function that mimics the U.S. tax code. I refer to this tax function as the baseline-fitted U.S. tax policy. I use the estimates from Gouveia and Strauss (1994) to determine the baseline-fitted U.S. tax policy. The authors match the U.S. tax code to the data using

a three parameter functional form,

$$T(y; \Upsilon_0, \Upsilon_1, \Upsilon_2) = \Upsilon_0(y - (y^{-\Upsilon_1} + \Upsilon_2)^{-\frac{1}{\Upsilon_1}}), \tag{29}$$

where y represents the sum of labor or capital income. The average tax rate is principally controlled by Υ_0, and Υ_1 governs the progressivity of the tax policy. Υ_2 is left free to ensure that the tax policy satisfies the budget constraint.

Gouveia and Strauss (1994) estimate values of $\Upsilon_0 = .258$ and $\Upsilon_1 = .768$ from the U.S. data. The authors do not fit separate tax functions for labor and capital income. Therefore, I use the same values on both sources of income for the baseline-fitted U.S. tax policy. I calibrate government consumption, G, such that it equals a percentage of output under the baseline-fitted U.S. tax policy, as observed in the U.S. data.[24] Hence, I set Υ_2 (for both sources of income) at the value that clears the government's budget constraint.

I calibrate the benchmark model such that G is 17 percent of output. In the model without the restrictions on taxing accidental bequests, the government raises more money from taxing accidental bequests because it fully confiscates them instead of only taxing the return on the bequests. To make these models comparable with the others, I adjust the government budget constraint such that the distortionary taxes on ordinary capital and labor income are the same percent of output in both models. This budget constraint implies that the government will raise more revenue in the models where it consumes accidental bequests; however, the part from distortionary taxes will be the same in all models. Additionally, in the models where the government holds debt (savings), I assume that the interest payments (income) offset government consumption. Therefore, traditional government consumption not including debt services is smaller (larger) in the model with government debt (savings) compared with the benchmark.

When determining the optimal tax policy, I restrict my attention to revenue neutral changes to the tax policy where the optimal tax policy is a separate flat tax rate on capital income and on labor income (τ_k and τ_h). When searching for the optimal tax policy, I limit my attention to flat taxes instead of searching over progressive tax policies. Conesa et al. (2009) and Peterman (2012) solve for the optimal tax policies in a model similar to the benchmark model. They both find that the optimal tax policies are flat taxes in models that do not include within-cohort heterogeneity. Therefore, I restrict my attention to flat taxes because all the agents within a cohort are homogenous. This experiment implies that, within a model, the government consumption is equal under the baseline-fitted U.S. tax policy and the optimal tax policies.

In the benchmark model, the social security system is chosen so that the replacement rate, b, is 50 percent.[25] The payroll tax, τ_{ss}, is determined such that the social security system contains a balanced budget each period.

---

[24]To determine this target, I use government expenditures less defense consumption.
[25]The replacement rate matches the rate in Conesa et al. (2009) and Conesa and Krueger (2006).

5 Computational Experiment

The computational experiment begins by solving for the optimal tax policy in my benchmark model. Next, I solve for the optimal tax policy in a model that has a utility function that implies a constant Frisch elasticity and allows the government to consume accidental bequests. I choose to examine these two features because the first confounds a motive for a positive tax on capital with the desire of the government to consume accidental bequests and the second has only limited empirical motivation.

Finally, I examine the strength of each of the other motives by eliminating each of them from the benchmark model. The aspects of the benchmark model that I change are: a varying Frisch labor supply elasticity profile; no separate tax rates on accidental bequests and ordinary capital income; individual borrowing constraints; excluding exogenously determined government savings or debt; and including a reduced form social security program. I solve for the optimal tax policy in a total of eight different iterations of the model. Table 1 lists the features in each iteration of the model.

To quantify the optimal tax policy, I need a social welfare function. Following Conesa et al. (2009) I choose a social welfare function that corresponds to a Rawlsian veil of ignorance (Rawls (1971)). Because in a stationary equilibrium, living agents face no earnings uncertainty, the social welfare is equal to the expected lifetime utility of a newborn,

$$SWF(\tau_h, \tau_k) = \sum_{j=1}^{J} \beta^{j-1} \left[\prod_{q=0}^{j-1} \Psi_q \right] u(c_j, h_j), \tag{30}$$

where τ_h is the flat tax rate on labor income and τ_k is the flat tax rate on capital income. When I determine the optimal tax policy, I search over τ_h and leave τ_k free to satisfy the government's budget constraint.[26] I require that any change in the tax policy is revenue neutral compared with the baseline-fitted U.S. tax policy.

6 Results

In this section, I start by solving for the optimal tax policies in the benchmark model (A0) and the alternative specification (B0) to test the effect of the varying Frisch elasticity and accidental bequests assumptions on the optimal tax policy. Next, I change one of the features of the benchmark model and solve for the optimal tax policy (models A1, A2, A3, A4, A5, and A6) to determine each feature's individual effect on optimal tax policy.

---

[26]Even when I exogenously impose a level of government savings or debt, I am able to solve for a unique τ_k because I solve for a tax policy with a specific level of government savings or debt.

6.1 Effect of Frisch Elasticity and Accidental Bequests

6.1.1 Optimal Policies

The first two columns of table 4 list the the optimal tax policies from the benchmark model, and the alternative specification model and the third column lists the ratio between the optimal tax on capital and labor income. I find that when I change the utility specification and allow the government to consume transfers, the optimal tax on capital drops by almost one-half. Although the tax on capital remains positive, it is no longer large. Therefore, two assumptions – the first which has limited empirical motivation, and the second that confounds a motive for a positive tax on capital with the governments desire to tax accidental bequests – are jointly responsible for approximately one-half of the large optimal tax on capital.

Table 4: **Optimal Tax Policies**

Model	τ_k	τ_h	$\frac{\tau_k}{\tau_h}$	CEV (baseline)	CEV (A0)
A0 (Benchmark):	29.3%	21.3%	1.4	0.73%	n/a
B0 (Alternative Specification):	16.6%	23.7%	0.7	0.95%	0.35%

The fourth column of table 4 lists the consumption equivalent variations based on the fitted U.S. tax policy (CEV (baseline)). The CEV (baseline) is the uniform increase in an agent's lifetime consumption that is necessary to make him indifferent between being born under the baseline-fitted U.S. tax policy and the optimal tax policy. I find that the CEV (baseline) is 0.73 percent and 0.95 percent in the benchmark and alternative specification models, respectively. Total consumption for all individuals in the United States was approximately $10,245.5 billion in 2010, so the CEV (baseline) represents approximately $74 billion and $98 billion dollars in the benchmark and alternative specification models, respectively.[27] Next, I assess how much welfare would be lost by using the optimal tax policy from the wrong model. Specifically, I impose the optimal tax policy from A0 in model B0 and determine the increase in consumption required to return an individual to the same level of utility as under the optimal tax policy in B0.[28] I refer to this welfare effect as CEV (A0) since it measures the welfare impact of the optimal tax policy from A0 in other models. I find that the CEV (A0) for model B0 is 0.35 percent or approximately one-third of the size of CEV (baseline).

6.1.2 Transition

When examining the welfare effects of adopting the optimal tax policy, it is useful to analyze the transition from the steady state under the current tax policy to the steady state under the optimal tax policy. To assess

---

[27]See personal consumption expenditures from Bureau of Economic Analysis.

[28]Implementing the optimal tax policy from model A0 in model B0 will not raise the right amount of revenue to satisfy the government's budget constraint. Therefore, when implementing the optimal tax policy from A0 in model B0, I allow the government to adjust the levels of the tax on capital and labor but require the ratio of the tax rates to remain consistent with the optimal tax policy in A0.

the transition path, it is necessary to take a stand on how the tax policy adjusts as the economy moves toward the new steady state. I assume that the economy starts at the steady state under the baseline-fitted U.S. tax policy. In the first period of the analysis, the government imposes the optimal labor income tax rate. In each period the government adjusts the tax on capital in order to fulfill their budget constraint. Over time the economy converges to its new steady state, which is the one computed under the optimal tax policy. I analyze the transitions in both the benchmark and alternative specification models.

Starting with the benchmark model (A0), figure 1 plots the capital and labor tax rates during the transition. The upper-left and right panels of figure 1 plot the tax rate for capital and labor income, respectively. The circles represent the average marginal tax rates under the baseline-fitted U.S. tax policy, and the lines are the marginal tax rate throughout the transition.[29] The bottom panel of figure 1 excludes the baseline-fitted U.S. tax policy rate so that one can focus on the movement of the tax on capital throughout the transition. Under the baseline-fitted U.S. tax policy, the average marginal tax rate on labor and capital income are approximately 25.5 percent and 19.3 percent, respectively. In the first period of the transition, the tax rate on labor income drops to 21.3 percent and the tax on capital income increases to approximately 29.5 percent. After the first period, the tax on labor is constant, while over time, the tax rate on capital falls a small amount before it stabilizes at its new steady state value of 29.3 percent.

Figure 2 plots the evolution of the aggregate economic variables in the benchmark model during the transition. After the initial large increase in the tax on capital, agents save less and capital falls. Because of the change in the tax on labor, agents gradually increase the amount they work throughout the transition. Additionally, total consumption increases while aggregate savings decreases. Over time, as capital levels out, so does consumption to its new level, which is approximately 1 percent higher than under the baseline-fitted U.S. tax policy. The decrease in capital and increase in labor over time causes wages to decline and rental rates to increase.

Welfare in the transition needs to be examined for two separate groups. The first group is those that are alive at the time of the tax policy change and the second group is those that are born during the transition. The left panel of figure 3 plots the CEV (baseline) for all living individuals at the time the tax change is enacted (the first group). The x-axis for this plot is the age of the cohort when the tax change is enacted. Within the living individuals, the channels by which the tax change effects welfare differ between retired and working individuals. During the transition, there are two counteracting effects on the previously retired generation's welfare. First, the tax on capital increases, which reduces the after-tax return to their savings. However, since the new steady state level of capital is lower, agents increase consume in order to deplete their savings throughout the transition. The increase in consumption increases their utility.[30] For the younger retired generations who live for more

---

[29]Since the baseline-fitted U.S. tax policy is progressive, these values are the average marginal rates paid by all individuals.

[30]The lower level of savings implies that individuals are dying with less assets. The reduction in lifetime savings lowers the

Figure 1: **Tax Rates in Transition in Model A0**
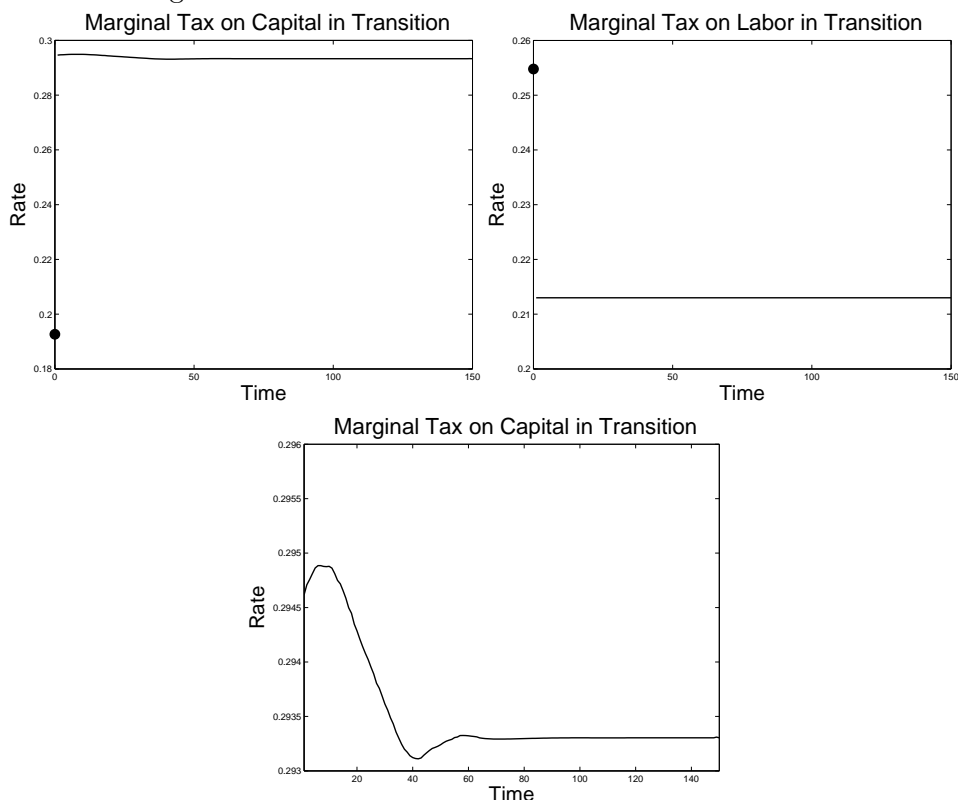
Note: The circles are the tax rates in the steady state under the baseline-fitted U.S. tax policy.

periods, the decrease in the value of savings dominates, and their welfare drops. Conversely, for the older retired generations, the increase in consumption dominates, and their welfare increases.

The working generations who are alive at the time of the tax change also experience two counteracting forces on their utility. First, the tax on labor decreases, which increases the value of the individual's time endowment. However, the increase in the tax on capital decreases the value of their savings. For younger working individuals who have not saved much, the increase in the value of their time endowment dominates, and they experience a welfare increase. The older working individuals do not have as many periods to work, so the increase in the value of their time endowment is less important. Therefore, the decrease in the value of the older working individuals' savings dominates, and they experience a decline in their welfare during the transition.

Overall, I find that approximately 36 percent of the living population experiences an increase in welfare during the transition compared with the baseline steady state. The population weighted average CEV (baseline) for the living individuals at the time the new tax policy is enacted is −0.36 percent. While the new steady state welfare will be higher in model A0, these results demonstrate that during the transition, individuals generally

distortion imposed by the individuals inability to insure against lifetime length uncertainty. Therefore, it is intuitive that the older individuals, with a higher probability of dying, experience the largest increase in welfare from this effect.

Figure 2: **Aggregates in Transition in Model A0**
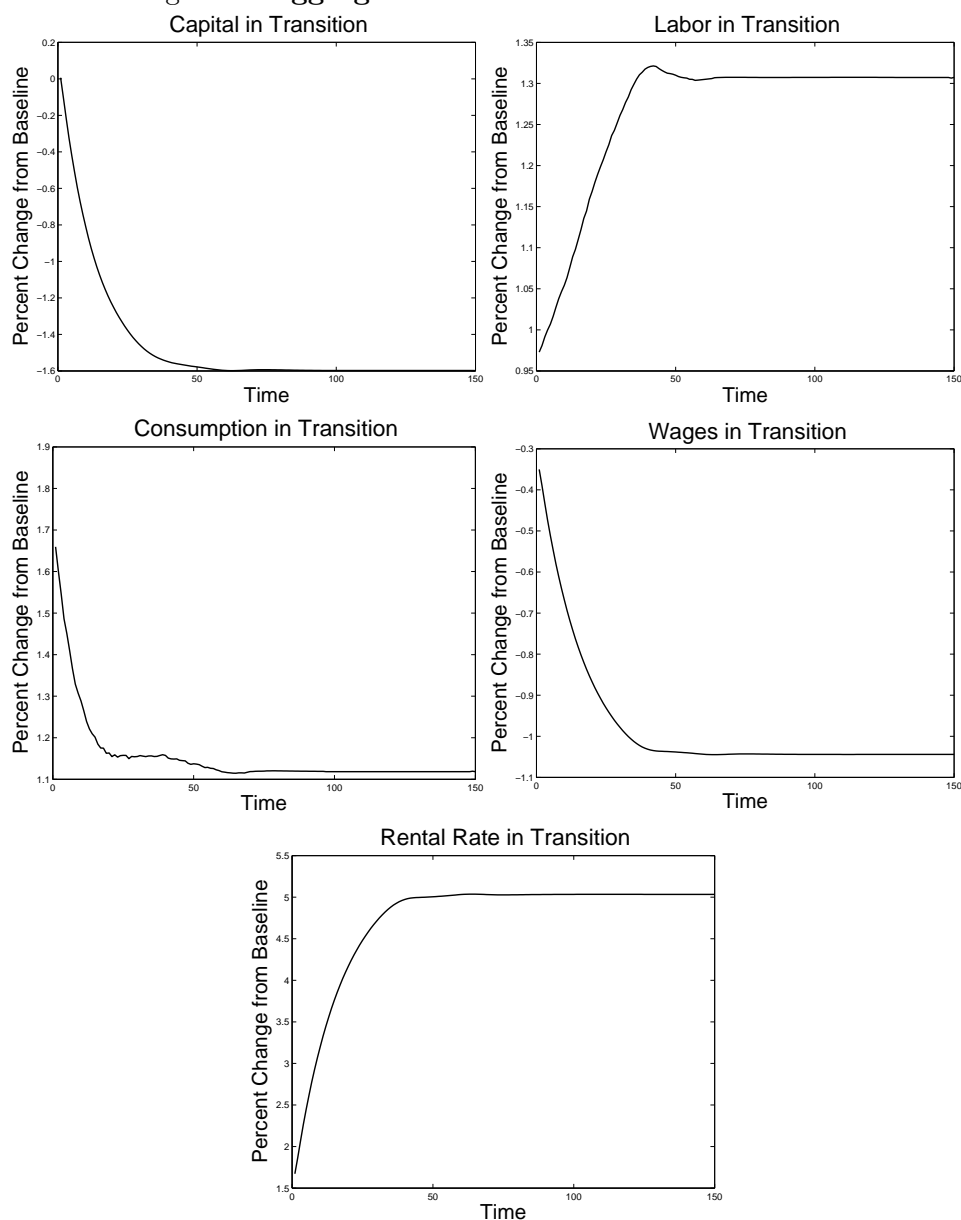
Note: The figures are the percent change from the steady state under the baseline-fitted U.S. tax policy.

will suffer welfare losses. The right panel of figure 3 plots the welfare effects on individuals who are born during the transition. The x-axis of this plot is the number of years after the tax policy was enacted when the cohort is born. Specifically, the figure plots the CEV (baseline) for each cohort born after the new tax policy is enacted. In contrast to the individuals born prior to the policy change, I find that all of these cohorts born later benefit from the tax policy change.

Next, turning to the alternative model (B0), figure 4 plots the capital and labor tax rates during the transition. The upper-left and right panels plot the tax rate for capital and labor income, respectively. The

Figure 3: **Transition CEV (baseline) in Model A0**
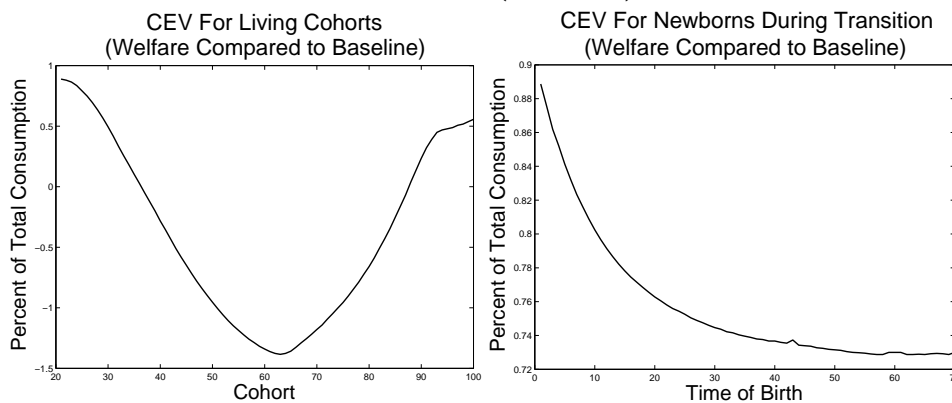
Note: The left panel is the CEV for all living individuals at the time of the tax change. The right panel is the CEV for all newborns after the tax change is enacted. The x-axis for the left panel is the age of the cohort when the tax policy change is enacted. The x-axis for the right panel is the number of years after the tax policy was enacted that the cohort is born.

circles represent the average marginal tax rates under the baseline-fitted U.S. tax policy and the lines are the marginal tax rates after the tax policy is adjusted. The lower panel of figure 4 excludes the average marginal tax rates under the baseline so that one can focus on the changes in the marginal tax on capital over the transition. The average marginal tax rates on both labor and capital under the baseline tax policy are higher than under the optimal tax policy. It is possible to have a revenue neutral change in the tax policy that consists of a decrease in the average marginal rates on both sources of income for two reasons. First, the baseline-fitted U.S. tax policy is progressive while the optimal tax policy consists of flat taxes. Second, the economy is larger under the optimal tax policy so total income is also larger. After the marginal tax on capital falls in the first period, it continues to fall over the transition until it levels out at the new steady state value of just below 17 percent.

Figure 5 plots the evolution of the aggregate economic variables in the alternative model during the transition. During the transition, I find that capital rises at a decreasing rate in response to the decreasing tax rate on capital. Additionally, labor rises over 1.5 percent in the first period after the marginal tax rate on labor decreases. However, over the rest of the transition, total labor supply reverts toward its new steady state level, which is approximately 0.8 percent higher than under the baseline-fitted U.S. tax policy. Overall, since both labor and capital increase, the size of the economy increases during the transition. These changes cause the rental (wage) rate to jump up (down) immediately after the change in tax policy but then to fall (rise) to its new level.

Unlike in model A0, I find that welfare improves for all living individuals when the tax policy is enacted (see left panel of figure 6). Overall, the average CEV (baseline) for living individuals is 0.71, percent indicating that the welfare increase during the transition is similar to the welfare increase in the new steady state after the tax policy change (0.71 percent is similar to the transitionless steady state welfare change of 0.95 percent).

26

Figure 4: **Tax Rates in Transition in Model B0**
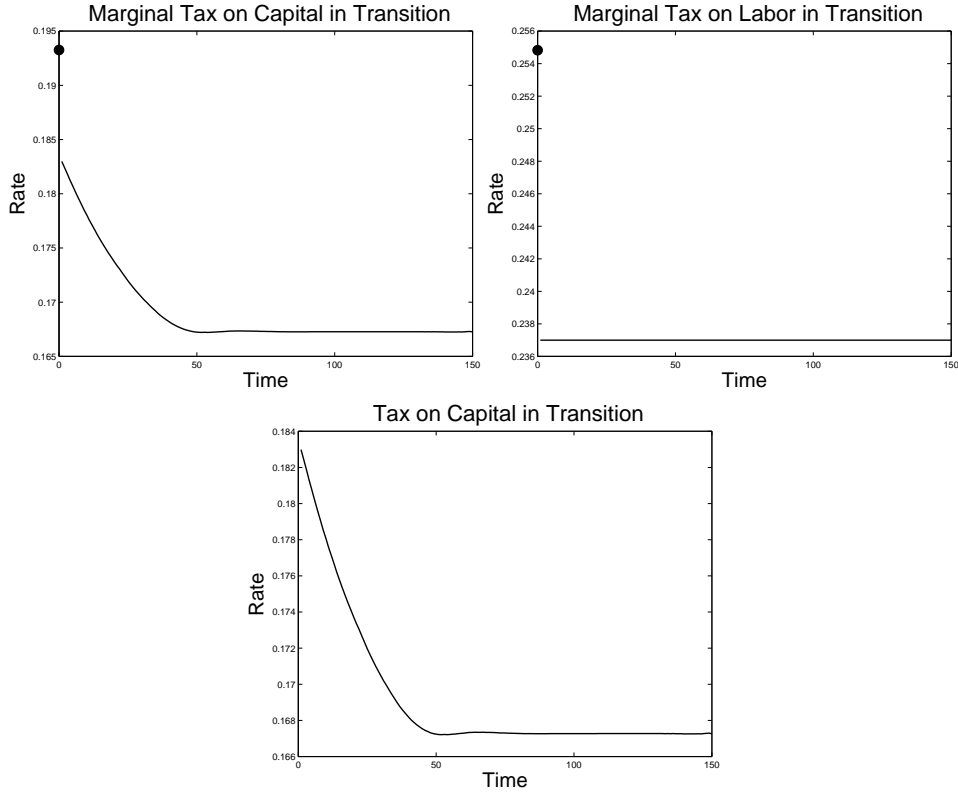
Note: The circles are the tax rates in the steady state under the baseline-fitted U.S. tax policy.

Additionally, I find that as the economy transitions toward the new steady state, the welfare of new cohorts born increases (see right panel of figure 6) compared with those born at the time the new policy is enacted. These two panels indicate that old individuals at the time of the policy change and individuals born after the policy change are the groups that experience the greatest gain in welfare. The older individuals who are already retired at the time that the tax policy changes benefit from the initial large increase in the after-tax rental rate on capital. The individuals born later in the transition benefit from higher wages resulting from the larger capital stock. These results indicate that the welfare improvements from the optimal tax policy in the alternative model are similar in the new steady state and over the transition. However, the welfare effects of the new tax policy during the transition are different in model A0 and B0. The next section documents the individual effect of all the model features on optimal tax policy and the economy.

6.2 Determining the Individual Effect of Each Assumption

Table 5 describes the optimal tax policies and the aggregate economic variables in the seven iterations of the model that test each feature individually. Columns three and four focus on the welfare consequences of adopting the optimal tax policies. Column three, CEV (baseline), is the uniform increase in consumption necessary to

27

Figure 5: **Aggregates in Transition in Model B0**
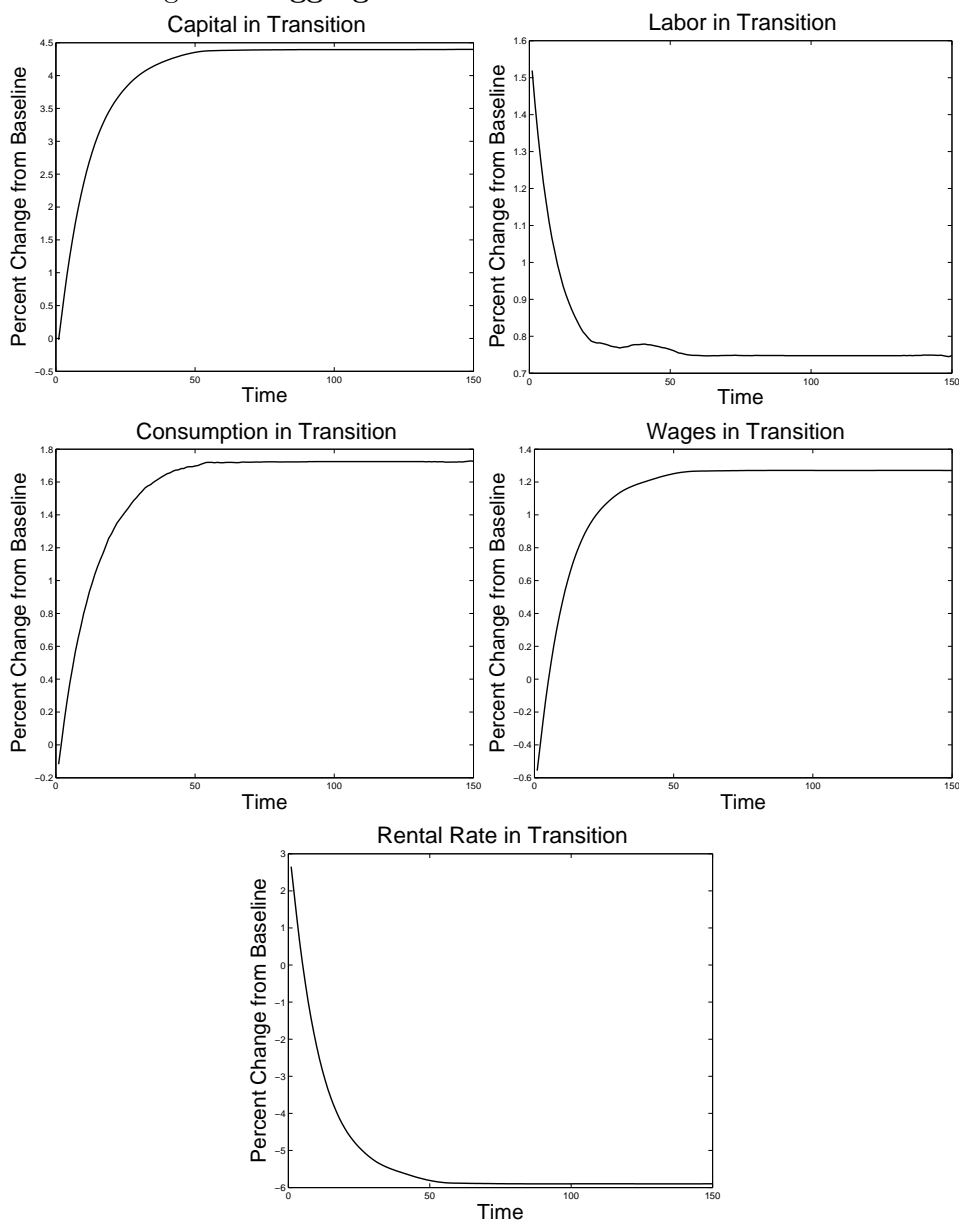
Note: The figures are the percent change from the steady state under the baseline-fitted U.S. tax policy.

make an agent indifferent between the baseline-fitted U.S. tax policy and the optimal tax policy. Within the specific model, column four, CEV (A0), is the percent increase in consumption needed to make an individual indifferent between the inferior optimal tax policy for A0 and the optimal tax policy for that model. Table 6 reports, for each model, what the percent differences in the aggregate economic variables compared with model A0 under the optimal tax policies. In this section, I examine the effect of each feature on the optimal tax policy, the aggregate economic variables, and life cycle profiles.

Figure 6: **Transition CEV (baseline) in Model B0**
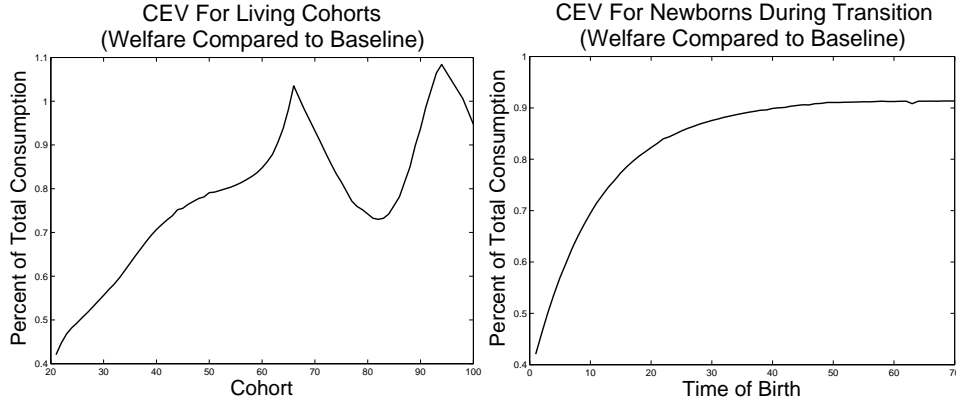
Note: The left panel is the CEV for all living individuals at the time of the tax change. The right panel is the CEV for all newborns after the tax change is enacted.

Table 5: **Aggregate Economic Variables (Under Optimal Tax Policy)**

Model	τ_k	τ_h	CEV (Baseline)	CEV (A0)	Y	K	N	w	r	tr	τ_{ss}
A0 (Benchmark):	29.3%	21.3%	0.73%	n/a	0.81	2.15	0.47	1.11	0.052	0.024	11.7%
A1 (Constant Frisch):	19.9%	23.3%	0.95%	0.6%	0.82	2.23	0.46	1.13	0.048	0.026	11.5%
A2 (Gov't Consumes Tr):	19.4%	23.5%	0.77%	0.29%	0.82	2.25	0.46	1.13	0.047	0.026	11.5%
A3 (Ind. Borrowing):	21.7%	23%	0.71%	0.11%	0.82	2.22	0.47	1.12	0.049	0.026	11.6%
A4 (Gov't Saves):	-17.1%	29.3%	1.77%	2.53%	0.84	2.51	0.46	1.18	0.038	0.019	11.7%
A5 (Gov't Borrows):	53.8%	9%	3.01%	1.57%	0.77	1.85	0.47	1.05	0.067	0.03	11.7%
A6 (No SS):	31.3%	19.6%	0.77%	0.08%	0.81	2.14	0.47	1.1	0.054	0.04	0%

Table 6: **Percent Changes Compared to A0**

Model	Y	K	N	w	r	tr
A1 (Constant Frisch):	0.7%	3.8%	-1%	1.7%	-7.7%	5.7%
A2 (Gov't Consumes Tr):	1%	4.8%	-1.1%	2.2%	-9.7%	6.1%
A3 (Ind. Borrowing):	1.1%	3.3%	-0.2%	1.3%	-6%	4.6%
A4 (Gov't Saves):	3.9%	16.7%	-2.7%	6.7%	-28.4%	-20.6%
A5 (Gov't Borrows):	-4.8%	-13.8%	0.6%	-5.4%	27%	24%
A6 (No SS):	0.3%	-0.5%	0.8%	-0.5%	2.2%	64.6%

Notes: Each row is the percent change from the benchmark model (A0). For example, A1 is the percent change between A0 and A1.

6.2.1 Desire to Mimic Age-Dependent Tax

The first assumption I alter is changing the utility function such that the Frisch elasticity is constant. I demonstrated in section 2.3.2 that utilizing the constant Frisch utility function instead of the non-constant Frisch utility function eliminates the government's desire to condition taxes on age. Models A0 and A1 are identical except A0 uses the non-constant Frisch utility function, and A1 uses the constant Frisch utility function. I find that eliminating this channel reduces the optimal tax on capital by almost 10 percentage points (see table 5).

The different optimal tax policies lead to a lower level of aggregate labor and higher capital stock in model A1 than A0. The different levels of capital and labor translate into a higher wage rate and a lower pre-tax return to capital in A1. The welfare gain from adopting the optimal tax policy is larger in A1 compared with A0. In model A1, the CEV (baseline) is 0.95 percent and CEV (A0) is 0.6 percent.

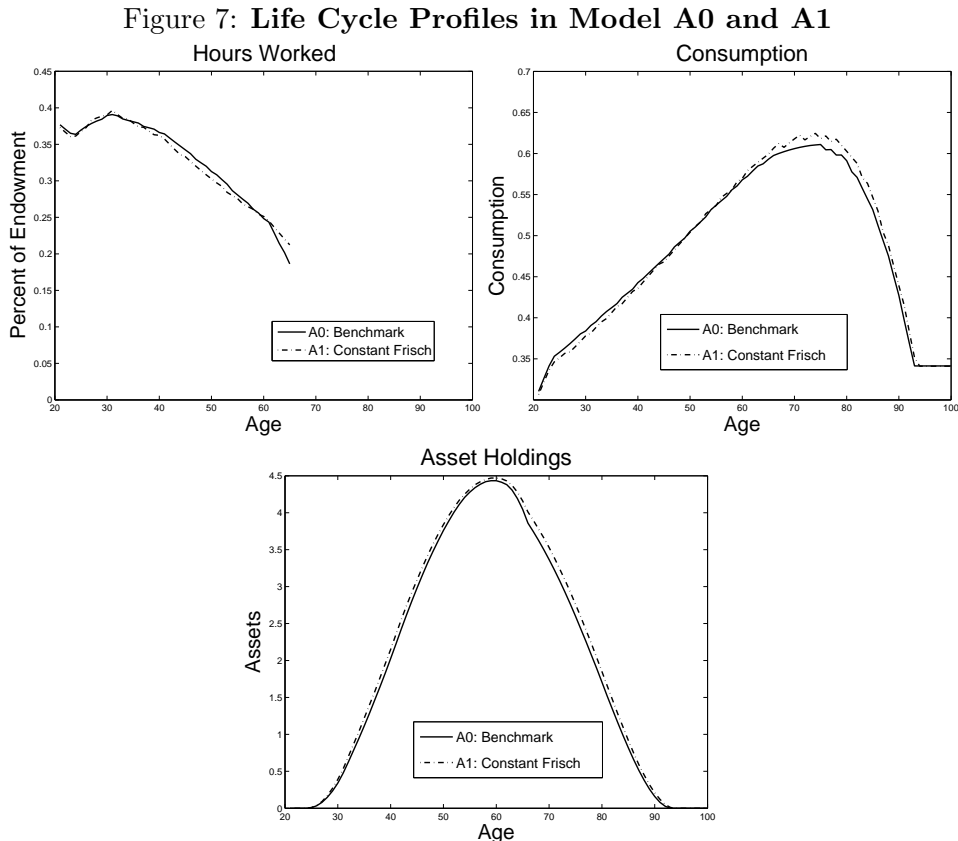Figure 7: **Life Cycle Profiles in Model A0 and A1**

Figure 7 plots the life cycle profiles for labor supply, consumption, and savings in the benchmark model (A0) and the model that eliminates the desire to condition taxes on age (A1). Generally, the life cycle profiles in the two models look similar. In the benchmark model, an agent's Frisch labor supply elasticity is negatively related to the hours they work. Therefore, in the benchmark model, agents become more elastic toward the end of their life when their hours decrease. In model A1, an agent's Frisch labor supply elasticity is constant. Therefore, agents tend to be relatively more elastic in their middle years and less elastic in their later years in model A1 compared with model A0. Since an agent's wage drops with his human capital late in his working life, agents work fewer hours in their middle years and more hours in their later working years in model A1 (see upper-left panel of figure 7).

The change in the marginal after-tax return in A1 affects the shape of the lifetime consumption profile. The intertemporal Euler equation controls the slope of consumption profile over an agent's lifetime. The relationship

30

is

$$\left(\frac{c_{j+1}}{c_j}\right)^{\varsigma_1} = \Psi_j \beta \widetilde{r}_t, \tag{31}$$

where $\widetilde{r}_t$ is the marginal after-tax return on capital. Since the marginal after-tax return on capital is larger in A1 than in A0, the consumption profile, in the upper-right panel of figure 7, is steeper. The larger after-tax return on capital in A1 also causes agents to have more savings (see bottom panel of figure 7).

6.2.2 Government Consumption of Accidental Bequests

Next, I examine the effect of relaxing the assumption that the government taxes accidental bequests at the same rate as ordinary capital income and that the government is not allowed to tax the principal of the bequests. I demonstrated in section 2.4 why these restrictions cause the optimal tax on capital to be non-zero. In model A2, I allow the government to tax these incomes at different rates. In model A2, the government fully consumes these accidental bequests as opposed to redistributing them to living agents. In this model the government raises more revenue than in A0 because it consumes accidental bequests instead of only taxing the return on the bequests. To make the models comparable, I adjust the government budget constraint in A2 such that the distortionary taxes on ordinary capital and labor income are the same percent of output as in A0. This budget constraint implies that the government will raise more revenue in model A;, however, the part from distortionary taxes will be the same.

Comparing line one and three in table 5 shows that eliminating this motive for a positive tax on capital causes the optimal tax on capital income to drop by approximately 10 percentage points. Similar to model A1, the tax on labor income increases. Even in this model, where I include a varying Frisch elasticity, the optimal tax on capital drops by over one-third. Therefore, a large optimal tax on capital is not robust to allowing the government a richer policy set.[31]

Examining tables 5 and 6, the smaller optimal tax on capital and larger optimal tax on labor in model A2 cause agents to save more and work slightly less, so aggregate output is slightly larger than in model A0. The larger capital stock and smaller amount of labor causes the pre-tax return on capital to be smaller and the wage rate to be larger. Since agents have higher levels of savings, their accidental bequests are also larger. The CEV (baseline) from adopting the optimal tax policies are similar in A0 and A2. CEV (A0) in A2 is approximately one-third as large as CEV (baseline).

Figure 8 plots the life cycle profiles in model A0 and A2. The lower optimal tax on capital in A2 implies that the tax on young labor income is relatively lower. Therefore, the agents work more hours earlier in their life. A2 has a larger overall tax burden since the government confiscates accidental bequests. The increase in

---

[31]If model A2 is calibrated with a budget constraint equal to A0, then using accidental bequests to finance government consumption as opposed to other distortionary taxes is welfare improving.

Figure 8: **Life Cycle Profiles in Model A0 and A2**
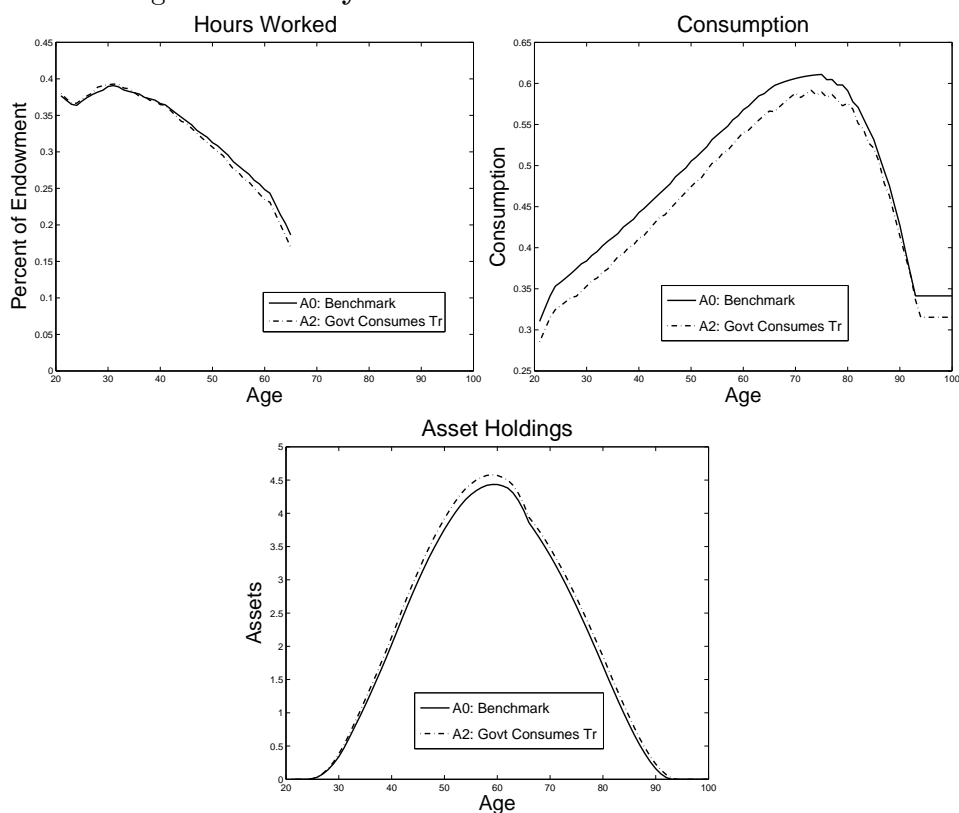
government consumption causes the consumption profile to be lower in A2. Agents compensate for the lack of income from transfers by accumulating more assets in A2.

6.2.3   Individual Liquidity Constraints

Two forces affect the optimal tax on capital in opposite directions when I change the model to allow individual's to borrow. First, agents prefer to smooth their consumption. Therefore, when an agent faces a hump-shaped lifetime earnings profile, he would prefer to smooth his consumption by borrowing against earnings from later years to facilitate consumption in earlier years. Borrowing constraints hinder an agent's ability to shift consumption, creating a role for tax policy to help facilitate this shift. Since an individual typically accumulates more assets later in their life, increasing the tax on capital income and decreasing the tax on labor income will allocate more of the lifetime tax burden to an individual's later years, which facilitates consumption smoothing. Therefore, restricting agents from borrowing can motivate a positive tax on capital. However, second, when agents are allowed to borrow, I find that agents decrease their labor supply early in their life because they are able to utilize borrowing (see the upper-left panel of figure 9). This shift in hours affects the labor supply elasticity in model A3, causing young agents to supply labor more elastically than in model A0. This change in relative elasticity leads to a decrease in the desire to mimic an age-dependent tax on labor income and, in turn

a decrease in the optimal tax on capital. To determine the overall affect, I compare the optimal tax policies in a model where agents are not able to borrow (A0) and one where agents can borrow at the actuarially fair rate (A3). I find that when I eliminate individual liquidity constraints, the optimal tax on capital falls nearly 7 percentage points.

The competing effects on the optimal tax policy also mean that the aggregate economic variables are similar in A3 to those in A0. Overall, the optimal tax policy in A3 looks similar to A1, so the aggregate economic variables are also similar. CEV (baseline) is similar in A0 and A3. CEV (A0) is small in model A3 since the optimal tax policies in A0 and A3 are similar.

Figure 9: **Life Cycle Profiles in Model A0 and A3**
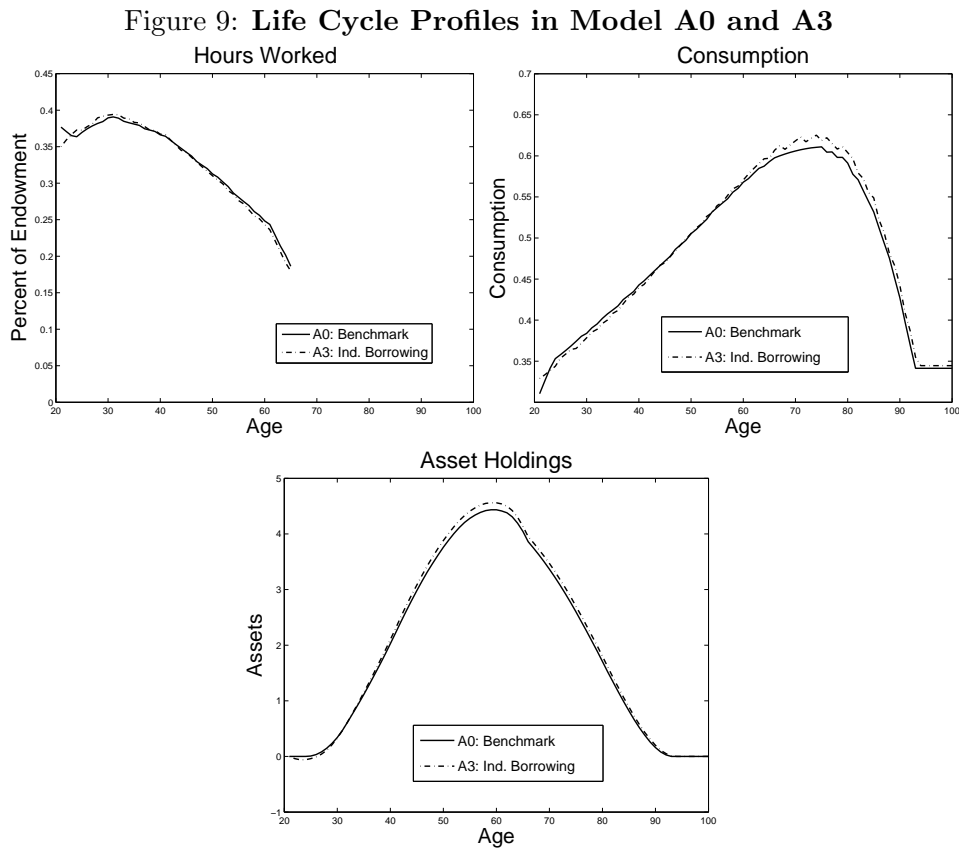
The lower panel of figure 9 demonstrates that an agent's borrowing constraint is only binding in the first few years of their life. Therefore, eliminating borrowing constraints alters an agent's hours and consumption decisions in the first few years of their life (see the upper-left panel of 9). After the first five years, the life cycle profiles in model A0 and A3 look similar.

6.2.4 Government Savings and Debt

Assuming that the government has an exogenously set level of savings or debt also alters the optimal tax on capital (see Conesa et al. (2009) for an analytical derivation). To quantify the strength of this motive, I

examine the optimal tax policy when the government has savings (A4) or debt (A5). I examine the model when government savings or debt equals 550 percent of their annual consumption. I use this number because the relative government debt-to-government expenditures (less defense consumption) was approximately 550 percent in 2008.[32] I assume that the government borrows or saves in the form of productive capital.[33] Additionally, I assume that the amount the government raises from distortionary taxes is unaffected by the debt (savings) and only government consumption is affected by the interest payments (rebates).

Savings Comparing models A0 and A4, it is clear that this level of government savings has large effects on the optimal tax policy. Including government savings causes over a 45 percentage point drop in the optimal tax on capital and an 8 percentage point increase in the optimal tax on labor income (see table 5). The aggregate capital stock is over 17 percent larger in A4 because the government now holds capital in addition to private agents (see table 6). The higher tax rate on labor income reduces aggregate labor supply by almost 3 percent. Due to the larger capital stock, output is almost 4 percent larger in A4 compared with A0. Since the capital-labor ratio in A4 is higher, wages are approximately 6.5 percent larger, and the pre-tax return to capital is almost 30 percent smaller. The smaller tax on capital and lower rental rate on capital have opposing affects on the after-tax return. However, overall the after-tax return to capital is larger in A4. Although total output is larger in A4, the amount devoted to private saving is smaller by approximately 19 percent. The lower level of private savings causes a lower level of bequests. Compared with the baseline-fitted U.S. tax policy, the optimal tax in A4 is much different, hence the CEV (baseline) is almost two and a half times as large as in A0. Additionally, CEV (A0) is larger than CEV (baseline).

Examining the upper-right and bottom panels of figure 10, it is clear that the lower private saving in A4 causes the consumption and asset accumulation profiles to be lower. Overall, the after-tax return to labor is lower in A4 which translates into the labor supply profile being lower.

Debt Including government debt in the model causes an opposite reaction in the aggregate economic variables compared with the model in which the government holds savings. The tax on capital increases over 24 percentage points, and the tax on labor decreases by approximately 12 percentage points (see table 5) in model A5. Comparing the fourth and fifth lines of table 6, it is clear that adding government debt causes the aggregate economic variables to have an opposite reaction (generally with a similar magnitude) compared with the model in which the government holds savings. Once again, the optimal tax policy in model A5 is much different than the baseline-fitted U.S. tax policy compared with the optimal tax policy in A0, so the CEV (baseline) is almost

---

[32]Previous studies such as Garriga (2001) and Conesa et al. (2009) use the debt-to-GDP ratio to calibrate debt as opposed to using the debt-to-government consumption ratio. A ratio of 550 percent of debt-to-government consumption is equivalent to a debt-to-GDP ratio of 94 percent.

[33]This assumption implicitly rules out the government issuing bonds.

Figure 10: **Life Cycle Profiles in Model A0 and A4**
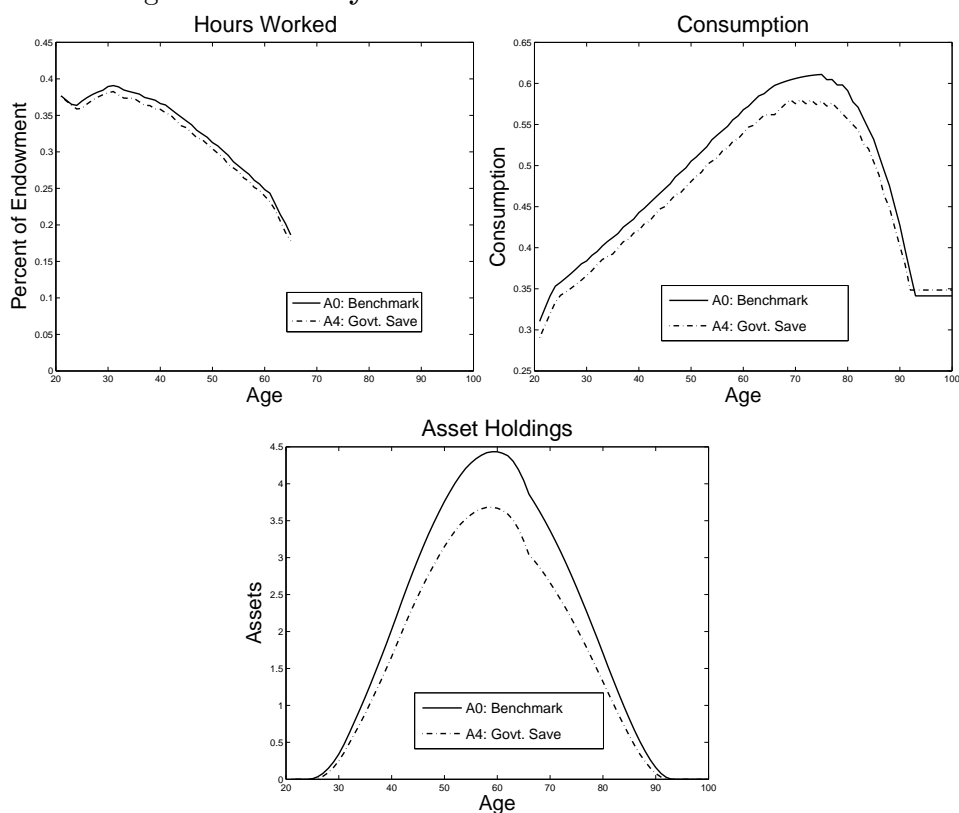
four times as large in A4. CEV (A0) is also larger in this model.

The effect of the government holding debt on the life cycle profiles under the optimal tax policy is also opposite to the effect of the government holding savings (compare figures 10 and 11). The government holding debt causes an increase in the amount of private saving, hence both the consumption and asset profiles are higher in A5 compared with A0. The hours profiles look almost identical in A0 and A5.

6.3 Social Security Program

Model A6 examines the effect of the social security program on the optimal tax policy by eliminating it from model A0. Eliminating the social security program has a small effect on the optimal tax policy. When the program is excluded, the optimal tax on capital increases approximately 1.5 percent. With the exception of transfers, the aggregate economic variables look similar in A0 and A5. Since agents no longer receive social security benefits, they need to increase their level of saving in order to finance consumption once they retire. Therefore, agents hold substantially more savings after they retire to finance their consumption. Additionally, the larger savings causes an increase in accidental bequests. CEV (baseline) is similar in A6 as to that in A0. CEV (A0) is small (0.08 percent) since the optimal tax policies in A0 and A6 are so similar.

Without a social security program, the government would like to decrease the tax on capital (or provide

Figure 11: **Life Cycle Profiles in Model A0 and A5**
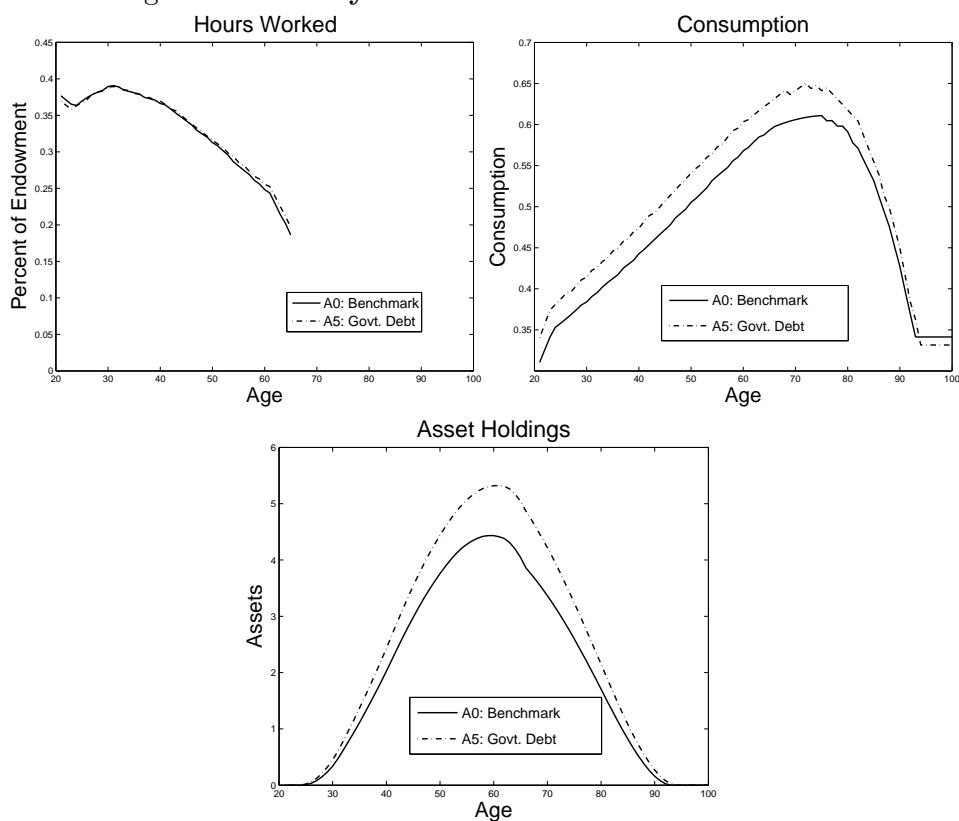
a rebate on capital income) in order to mimic the welfare improving social security program. However, the motives for a positive tax on capital demonstrated in A2 and A3 are enhanced in this model and overall cause the optimal tax on capital to increase. Accidental bequests are higher in this model, so the motive for a positive tax on capital demonstrated in A2 is enhanced. Also, agents must finance their own retirement with personal savings, so the agents have more incentive to save. However, I recalibrate the model such that capital to output is consistent across all the models under the baseline-fitted U.S. tax policy. To induce a similar level of capital, the discount rate falls in model A6 (see table 3). The change in the discount rate coupled with the increased need for savings late in an agent's life causes the life cycle savings profile to shift to the right (see the lower box in figure 12). This shift means that agents face binding liquidity constraints for more years and the motive for a positive tax on capital demonstrated in model A3 is also enhanced. Overall, these three motives generally cancel each other out.

Excluding the social security program causes the life cycle consumption and savings profiles to have less realistic shapes. The upper-left and right panels in figure 12 demonstrate that without a social security program, the labor profile and consumption profile are flatter when agents are working. The profiles are flatter because of the decrease in β. Since agents face lifetime uncertainty and finance their own retirement consumption in model A6, their consumption falls much more dramatically toward the end of their life. Additionally, agents

Figure 12: **Life Cycle Profiles in Model A0 and A6**
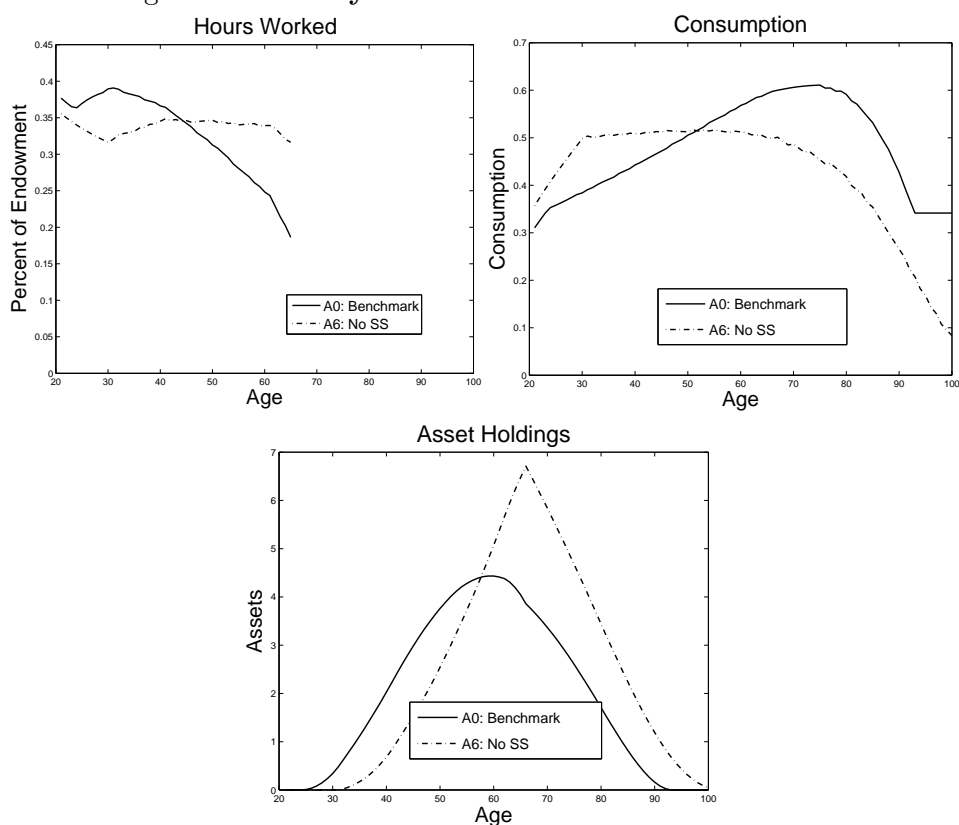
accumulate more assets to finance retirement, so their lifetime savings profile shifts to the right in model A6.

6.4 Summary of Results

Overall, I find that assuming the Frisch elasticity is non-constant and that the government cannot distinguish accidental bequests are significant motives for a positive tax on capital. Including individual budget constraints also motivates a positive tax on capital but to a lesser extent. Including exogenously determined levels of government savings (borrowing) causes the optimal tax on capital to decrease (increase). Additionally, including a reduced form social security program is important because it causes the life cycle profiles to be more realistic.

7 Sensitivity Analysis

Next, I check the sensitivity of the results with respect to the parameter value that governs the Frisch labor supply elasticity. I choose to examine the sensitivity of the results with respect to this parameter because there is some uncertainty about the actual value of the Frisch elasticity. In this section, I test how using different Frisch elasticity parameters affects the optimal tax policy in the benchmark model and the effect of each of the model features on optimal tax policy. I solve for the optimal tax policy in model A0 with three different

values of σ_2 of 6, 3, and 2. These values imply a Frisch elasticity of $\frac{1}{3}$, $\frac{2}{3}$, and 1, respectively, if an agent works one-third of their time endowment. I also solve for models A1-A6 with the different values.[34] Prior to solving the models, I calibrate the models with the three different targets for the Frisch elasticity. Tables 7 and 8 list these parameters. Generally, I find that a lower elasticity target implies a higher discount rate parameter and lower disutility-to-labor parameter.

Table 7: **Calibration Parameters (Low Elasticity)**

Model	Conditional Discount β	Frisch Elasticity $\sigma_2\ \varsigma_2$	Disutility to Labor χ
Target	$\frac{K}{Y} = 2.7$	Frisch=$\frac{1}{3}$	Avg. H = $\frac{1}{3}$
A0 (Benchmark):	0.996	6	0.59
A1 (Constant Frisch):	0.996	$\frac{1}{3}$	182.4
A2 (Gov't Consumes Tr):	0.997	6	0.67
A3 (Ind. Borrowing):	0.996	6	0.59
A4 (Gov't Saves):	0.989	6	0.62
A5 (Gov't Borrows):	1.002	6	0.56
A6 (No SS):	0.969	6	0.55

Table 8: **Calibration Parameters (High Elasticity)**

Model	Conditional Discount β	Frisch Elasticity $\sigma_2\ \varsigma_2$	Disutility to Labor χ
Target	$\frac{K}{Y} = 2.7$	Frisch=1	Avg. H = $\frac{1}{3}$
A0 (Benchmark):	0.99	2	2.8
A1 (Constant Frisch):	0.991	1	19.9
A2 (Gov't Consumes Tr):	0.991	2	3.2
A3 (Ind. Borrowing):	0.99	2	2.8
A4 (Gov't Saves):	0.984	2	3.1
A5 (Gov't Borrows):	0.995	2	2.6
A6 (No SS):	0.969	2	2.8

7.1 Effect on Optimal Tax Policies in Benchmark Models

Table 9 presents the optimal tax policies in the benchmark model calibrated to target the three different Frisch elasticities. I find that the optimal tax on capital is larger when the model is calibrated to a higher Frisch elasticity. The optimal tax on capital increases for two reasons. First, when the government is deciding between taxing capital and labor income, it is weighing the relative distortions that each tax induces on the economy. An agent will be more sensitive to a tax on labor income when the Frisch elasticity is higher. Therefore, the government prefers to reduce the tax on labor income and increase the tax on capital under the two calibrations that target a higher Frisch elasticity.

Second, the higher Frisch elasticities enhances the motive for an age-dependent tax on labor income. Figure 13 plots the life cycle profiles for the three different calibrations. The upper-left panel of the figure demonstrates

---

[34]In model A1 I use the values of $\frac{1}{3}$, $\frac{2}{3}$, and 1 for ς.

Table 9: **Optimal Tax Policies in Benchmark Models with Different Frisch Elasticities**

Tax Rates	Frisch= $\frac{1}{3}$	Frisch= $\frac{2}{3}$	Frisch= 1
τ_k	22.5%	29.3%	30.5%
τ_h	22.8%	21.3%	20.9%

Figure 13: **Life Cycle Profiles in Benchmark Models with Different Elasticities**
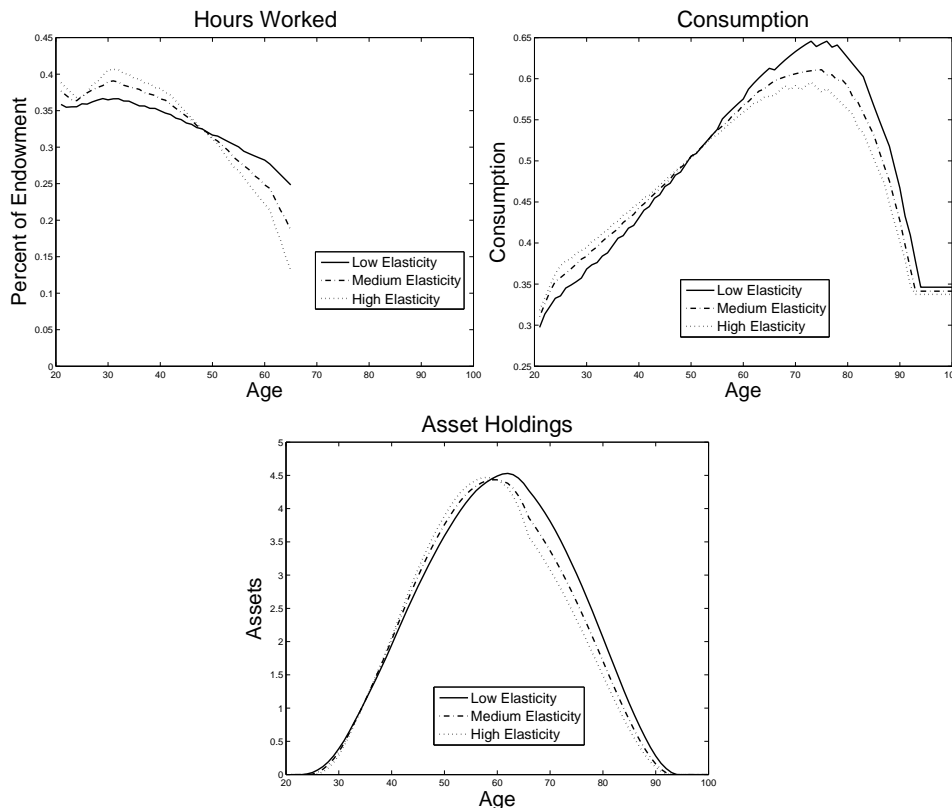
that as the model is calibrated to match a higher average Frisch elasticity, the relative change between the hours he works when he is young and old increases. A larger drop in hours enhances the motive for an age-dependent tax on labor income. Since the government cannot condition labor income taxes on age, they increase the tax on capital.

7.2 Frisch Elasticity Effect on Channels' Impact

To determine how changing the Frisch elasticity alters the effect on optimal tax policy of each of the channels, I solve for models A0, A1, A2, A3, A4, A5, and A6 under the three different Frisch elasticity targets. Table 10 describes the optimal tax policies in the six models under the three different calibrations. Table 11 presents the percentage changes in the optimal tax policies between the benchmark model (A0) and the various models (A1-A6) under all three calibrations. Generally, as the model is calibrated to match a lower Frisch elasticity,

each of the channels have a larger effect on the optimal tax policy.

Table 10: **Optimal Tax Policy in Sensitivity Analysis Under Different Calibrations**

Model	τ_k			τ_h		
	Low	**Medium**	**High**	**Low**	**Medium**	**High**
A0 (Benchmark):	22.5%	29.3%	30.5%	22.8%	21.3%	20.9%
A1 (Constant Frisch):	15.5%	19.9%	24.6%	24.3%	23.3%	22.3%
A2 (Gov't Consumes Tr):	9.8%	19.4%	24.1%	24.6%	23.5%	22.4%
A3 (Ind. Borrowing):	19.7%	21.7%	25.6%	23.4%	23%	22.1%
A4 (Gov't Saves):	-24.6%	-17.1%	-14.2%	30%	29.3%	29%
A5 (Gov't Borrows):	57.3%	53.8%	57%	7.8%	9%	7.1%
A6 (No SS):	32.8%	31.3%	29%	19.3%	19.6%	19.8%

Notes: Each row is the optimal tax policy for a model similar to the benchmark model with one channel removed. Low is calibrated with a target Frisch elasticity of $\frac{1}{3}$, medium is calibrated with a target Frisch elasticity of $\frac{2}{3}$, and high is calibrated with a target Frisch elasticity of 1.

Table 11: **Percentage Changes In Optimal Taxes Induced by Change in Model**

Model	τ_k			τ_h		
	Low	**Medium**	**High**	**Low**	**Medium**	**High**
A1 (Constant Frisch):	-31.1%	-32.1%	-19.3%	6.6%	9.4%	6.7%
A2 (Gov't Consumes Tr):	-56.4%	-33.8%	-21%	7.9%	10.3%	7.2%
A3 (Ind. Borrowing):	-12.4%	-25.9%	-16.1%	2.6%	8%	5.7%
A4 (Gov't Saves):	-209.3%	-158.4%	-146.6%	31.6%	37.6%	38.8%
A5 (Gov't Borrows):	154.7%	83.6%	86.9%	-65.8%	-57.7%	-66%
A6 (No SS):	45.8%	6.8%	-4.9%	-15.4%	-8%	-5.3%

Notes: Each row is the percentage change from the benchmark model. For example, A1 is the percentage change in the optimal tax policy between A0 and A1.

The optimal tax on capital decreases in models A1, A2, A3, and A4 for all of the calibrations. The optimal tax on capital increases in model A5 for all of the calibrations. The optimal tax on capital increases in A6 for the low and medium calibrations and decreases for the high calibration. When the social security program is removed, the motives for a positive tax on capital from the non-constant Frisch elasticity profile and restrictions on taxing capital are enhanced (see section 6.3). However, the government also wants to mimic a welfare improving social security program by reducing the tax on capital. These three motives have competing effects on the optimal tax on capital. In the models calibrated to match the high target, the desire to mimic the social security program dominates and the optimal tax on capital drops. When the models are calibrated to match the low and medium targets, the effect of the non-constant Frisch elasticity profile and the restrictions on how the government can tax accidental bequests dominate, so the optimal tax on capital increases.

8   Conclusion

Through an analysis of the optimal tax on capital in a standard life cycle model, this paper concludes that if one alters the utility function such that the Frisch elasticity profile is constant and allows the government to

tax accidental bequests at a separate rate from ordinary capital income, then the optimal tax on capital falls from 29.3 percent to 16.4 percent. It is important to quantify the effect of these two model features because there is not a consensus on whether the labor supply elasticity profile is upward sloping and prohibiting the government from taxing accidental bequests at a different rate from ordinary capital income confounds the government's desire to confiscate the bequests with a positive optimal tax on capital. Although the optimal tax on capital is not zero in the model without these features, it is no longer large. Comparing steady states under the baseline-fitted U.S. tax policy and the optimal tax policy, the CEV is 0.73 percent and 0.95 percent for the benchmark and alternative models, respectively. In the case of the benchmark model, over the transition, I find that adopting the optimal tax policy causes welfare to decrease for agents who are already living at the time of the tax policy change. In contrast, I find that in the alternative specification adopting the new tax policy increases the welfare of living individuals.

I also find that if the government holds savings (debt), then the optimal tax on capital decreases (increases). Removing individual liquidity constraints cause the optimal tax on capital to fall. I show that it is important to include at least a reduced-form social security program in a life cycle analysis of optimal tax policy, otherwise the life cycle profiles will be unrealistic. Overall, I find that in the various models, the welfare loss from adopting the optimal tax policy determined in the benchmark model as opposed to the actual optimal tax policy for that specific model range from 0.08 percent to 2.53 percent of total lifetime consumption. Finally, I find that generally as the models are calibrated to match a lower Frisch elasticity, the effect of changing the various features is larger.

When modelling certain aspects of the economy, economists try to balance realism and tractability. I demonstrate that some of these simplifying assumptions have a sizable effect on optimal tax policy. For example, assuming that the government cannot distinguish between ordinary capital and accidental bequests has large implications for optimal tax policy. Therefore, further research should focus on modelling this feature more realistically. Additionally, the shape of the Frisch labor supply elasticity profile has a large effect on the optimal tax policy. Since little empirical evidence addresses whether the Frisch elasticity varies, it is an important question for economists to examine.

A   Computational Algorithm

To determine the competitive equilibrium for each set of tax parameters, I use a modified algorithm based on Heer and Maussner's algorithm to compute a stationary equilibrium for the overlapping generation model.[35] The algorithm consists of the following steps:

1. Make initial guesses of the steady state values of the aggregate capital stock, labor, transfers, and social security benefits.

2. Make initial guess for tax rate.

3. Solve for the prices and social security tax rate.

4. Compute the optimal path for consumption, savings, and employment for the new-born generation by backward induction given the initial capital stock is zero.[36]

5. Compute the tax rate that clears the markets and compare with initial guess. If change of tax rate is not within the tolerance, then return to step 4.[37]

6. Update the aggregate capital stock and labor and return to step 3 until convergence.

   To determine the optimal tax policy, I solve each model's steady state in Matlab. I use a grid search method to determine which tax policy is optimal.

B   Analytical Derivations

B.1   Benchmark Simple Model

The Lagrangian for the benchmark simple model is

$$
\begin{aligned}
\mathscr{L} = & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h_{1,t})^{1-\sigma_2}}{1-\sigma_2} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h_{2,t+1})^{1-\sigma_2}}{1-\sigma_2} \\
& - \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\
& - \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\
& + \lambda_t(c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi(1-h_{1,t})^{-\sigma_2}h_{1,t} - \beta\chi(1-h_{2,t+1})^{-\sigma_2}h_{2,t+1}).
\end{aligned}
\tag{32}
$$

---

[35]See Heer and Maussner (2005).

[36]Even when agents are allowed to borrow it is assumed that they cannot borrow until the second period of their life.

[37]This extra loop to get convergence of the tax parameters prior to updating the aggregates is not included in Heer and Maussner's algorithm. I found that convergence was more stable when this step was included in some of the models.

The first order conditions with respect to $h_{1,t}$, $h_{2,t+1}$, K_{t+1}, $c_{1,t}$, and $c_{2,t+1}$ are

$$\rho_t = \chi(1 - h_{1,t})^{-\sigma_2}\left[1 + \lambda_t\left(1 + \frac{\sigma_2 h_{1,t}}{(1 - h_{1,t})}\right)\right],\tag{33}$$

$$\rho_{t+1}\theta = \chi(1 - h_{2,t+1})^{-\sigma_2}\frac{\beta}{\epsilon_2}\left[1 + \lambda_t\left(1 + \frac{\sigma_2 h_{2,t+1}}{(1 - h_{2,t+1})}\right)\right],\tag{34}$$

$$\rho_t = \theta(1 + r)\rho_{t+1},\tag{35}$$

$$\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t(1 - \sigma_1)c_{1,t}^{-\sigma_1},\tag{36}$$

and

$$\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta\lambda_t(1 - \sigma_1)c_{2,t+1}^{-\sigma_1}.\tag{37}$$

Combining the first order equations for the government's problem with respect to consumption (equations 36 and 37) yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta}.\tag{38}$$

Further, combining the agent's first order conditions, equations 5, and 6, under the non-constant Frisch utility specification yields

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1}{\epsilon_2}\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1}\left(\frac{1 - h_{2,t+1}}{1 - h_{1,t}}\right)^{-\sigma_2}.\tag{39}$$

Combining equation 39 and 38 gives

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \left(\frac{\beta\rho_t}{\epsilon_2\rho_{t+1}\theta}\right)\left(\frac{1 - h_{2,t+1}}{1 - h_{1,t}}\right)^{-\sigma_2}.\tag{40}$$

Next, I combine the first order conditions for the government with respect to young and old hours,

$$\frac{1 + \lambda_t(1 + \frac{\sigma_2 h_{1,t}}{1 - h_{1,t}})}{1 + \lambda_{t+1}(1 + \frac{\sigma_2 h_{2,t+1}}{1 - h_{2,t+1}})} = \frac{\beta\rho_t}{\epsilon_2\rho_{t+1}\theta}\left(\frac{1 - h_{2,t+1}}{1 - h_{1,t}}\right)^{-\sigma_2}.\tag{41}$$

Therefore, equation 40 and equation 41 simplify to

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t(1 + \frac{\sigma_2 h_{1,t}}{1 - h_{1,t}})}{1 + \lambda_t(1 + \frac{\sigma_2 h_{2,t+1}}{1 - h_{2,t+1}})}.\tag{42}$$

Utilizing the first order conditions from the Lagrangian with respect to capital and consumption leads to the following equation:

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1 + r).\tag{43}$$

Applying the non-constant Frisch utility function to equation 7 provides the following relationship:

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1 + r(1 - \tau_k)). \tag{44}$$

Equations 43 and 44 demonstrate that for the agent to choose the optimal allocation determined from the primal approach, the tax on capital must equal zero. Therefore, if the government can condition labor income taxes on age, then the optimal tax on capital is zero for the non-constant Frisch utility specification.

B.2   No Age Conditional Labor Income Taxes

When the government cannot condition labor income taxes on age, then equation 11 must be included as a constraint in the Lagrangian. The Lagrangian for this model is

$$\mathcal{L} = \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} + \chi\frac{(1-h_{1,t})^{1-\sigma_2}}{1-\sigma_2} + \beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} + \chi\frac{(1-h_{2,t+1})^{1-\sigma_2}}{1-\sigma_2} \tag{45}$$

$$- \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2))$$

$$- \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2))$$

$$+ \lambda_t(c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi(1-h_{1,t})^{-\sigma_2}h_{1,t} - \beta\chi(1-h_{2,t+1})^{-\sigma_2}h_{2,t+1})$$

$$\eta_t(\epsilon_2 c_{2,t+1}^{-\sigma_1}(1-h_{1,t})^{-\sigma_2} - c_{1,t}^{-\sigma_1}(1-h_{2,t+1})^{-\sigma_2}). \tag{46}$$

The first order conditions with respect to $h_{1,t}$, $h_{2,t+1}$, K_{t+1}, $c_{1,t}$, and $c_{2,t+1}$ are

$$\rho_t = \chi(1-h_{1,t})^{-\sigma_2}\left[1 - \eta_t\epsilon_2\sigma_2\frac{c_{2,t+1}^{-\sigma_1}}{(1-h_{1,t})} + \lambda_t(1 + \frac{\sigma_2 h_{1,t}}{(1-h_{1,t})})\right], \tag{47}$$

$$\rho_{t+1}\theta\epsilon_2 = \chi(1-h_{2,t+1})^{-\sigma_2}\beta\left[1 + \eta_t\sigma_2\frac{c_{1,t}^{-\sigma_1}}{(1-h_{2,t+1})} + \lambda_t(1 + \frac{\sigma_2 h_{2,t+1}}{(1-h_{2,t+1})})\right], \tag{48}$$

$$\rho_t = \theta(1+r)\rho_{t+1} \tag{49}$$

$$\rho_t = c_{1,t}^{-\sigma_1}\left[1 + \lambda_t(1-\sigma_1) + \frac{\eta_t\sigma_1(1-h_{2,t+1})^{-\sigma_2}}{c_{1,t}}\right], \tag{50}$$

and

$$\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1}\left[1 + \lambda_t(1-\sigma_1) - \frac{\eta_t\epsilon_2\sigma_1(1-h_{1,t})^{-\sigma_2}}{c_{2,t+1}}\right]. \tag{51}$$

Combining equations 49, 50, and 51 yields

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r)\left(\frac{1 + \lambda_t(1-\sigma_1) - \frac{\eta_t \epsilon_2 \sigma_1 (1-h_{1,t})^{-\sigma_2}}{c_{2,t+1}}}{1 + \lambda_t(1-\sigma_1) + \frac{\eta_t \sigma_1 (1-h_{2,t+1})^{-\sigma_2}}{c_{1,t}}}\right). \tag{52}$$

B.3 Constant Frisch Utility Function

The Lagrangian for this specification is

$$\begin{aligned}
\mathcal{L} = {} & \frac{c_{1,t}^{1-\varsigma_1}}{1-\varsigma_1} - \chi \frac{h_{1,t}^{1+\frac{1}{\varsigma_2}}}{1+\frac{1}{\varsigma_2}} + \beta \frac{c_{2,t+1}^{1-\varsigma_1}}{1-\varsigma_1} - \chi \frac{h_{2,t+1}^{1+\frac{1}{\varsigma_2}}}{1+\frac{1}{\varsigma_2}} \\
& - \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\
& - \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\
& + \lambda_t(c_{1,t}^{1-\varsigma_1} + \beta c_{2,t+1}^{1-\varsigma_1} + \chi h_{1,t}^{1+\frac{1}{\varsigma_2}} + \beta \chi h_{2,t+1}^{1+\frac{1}{\varsigma_2}}).
\end{aligned} \tag{53}$$

The first order conditions with respect to labor, capital, and consumption are

$$\rho_t = \chi h_{1,t}^{\frac{1}{\varsigma_2}}\left[1 + \lambda_t(1 + \frac{1}{\varsigma_2})\right], \tag{54}$$

$$\rho_{t+1}\theta\epsilon_2 = \beta \chi h_{2,t+1}^{\frac{1}{\varsigma_2}}\left[1 + \lambda_t(1 + \frac{1}{\varsigma_2})\right], \tag{55}$$

$$\rho_t = \theta(1+r)\rho_{t+1}, \tag{56}$$

$$\rho_t = c_{1,t}^{-\varsigma_1} + \lambda_t(1-\varsigma_1)c_{1,t}^{-\varsigma_1}, \tag{57}$$

and

$$\theta \rho_{t+1} = \beta c_{2,t+1}^{-\varsigma_1} + \beta \lambda_t(1-\varsigma_1)c_{2,t+1}^{-\varsigma_1}. \tag{58}$$

Combining the first order equations for the government's problem with consumption (equations 57 and 58) yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\varsigma_1} = \frac{\beta \rho_t}{\rho_{t+1}\theta}. \tag{59}$$

Taking the ratio of the agent's first order conditions, equations 5 and 6, under the constant Frisch utility specification gives

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1}{\epsilon_2}\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\varsigma_2}}. \tag{60}$$

Combining equation 59 and 60 yields

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{\beta \rho_t}{\epsilon_2 \rho_{t+1} \theta} \left(\frac{h_{2,t+1}}{h_{1,t}} \right)^{\frac{1}{\varsigma_2}}. \tag{61}$$

The ratio of first order equations for the government with respect to young and old hours is

$$\frac{\rho_t \beta}{\epsilon_2 \rho_{t+1} \theta} \left(\frac{h_{2,t+1}}{h_{1,t}} \right)^{\frac{1}{\varsigma_2}} = \frac{1 + \lambda_t (1 + \frac{1}{\varsigma_2})}{1 + \lambda_t (1 + \frac{1}{\varsigma_2})}. \tag{62}$$

Combining equation 62 and 61 generates the following expression for the optimal labor taxes:

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t (1 + \frac{1}{\varsigma_2})}{1 + \lambda_t (1 + \frac{1}{\varsigma_2})} = 1. \tag{63}$$

B.4 Unrestricted Tax Policy with Accidental Bequests

The Lagrangian for this specification is

$$\begin{aligned}
\mathcal{L} = & \frac{c_{1,t}^{1-\varsigma_1}}{1 - \varsigma_1} - \chi \frac{h_{1,t}^{1+\frac{1}{\varsigma_2}}}{1 + \frac{1}{\varsigma_2}} + \beta \Psi \frac{c_{2,t+1}^{1-\varsigma_1}}{1 - \varsigma_1} - \chi \frac{h_{2,t+1}^{1+\frac{1}{\varsigma_2}}}{1 + \frac{1}{\varsigma_2}} \\
& - \rho_t (c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - r K_t - w(h_{1,t} + h_{2,t} \epsilon_2)) \\
& - \rho_{t+1} \theta (c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - r K_{t+1} - w(h_{1,t+1} + h_{2,t+1} \epsilon_2)) \\
& + \lambda_t (-T r_t c_{1,t}^{-\varsigma_1} (1 + r)(1 - \tau_t) - \beta \Psi T r_{t+1} (1 + r)(1 - \tau_t) c_{2,t+1}^{-\varsigma_1} + c_{1,t}^{1-\varsigma_1} + \psi \beta c_{2,t+1}^{1-\varsigma_1} - \chi h_{1,t}^{1+\frac{1}{\varsigma_2}} - \psi \beta \chi h_{2,t+1}^{1+\frac{1}{\varsigma_2}}).
\end{aligned} \tag{64}$$

The first order conditions with respect to labor, capital, and consumption are

$$\rho_t = \chi h_{1,t}^{\frac{1}{\varsigma_2}} \left[1 + \lambda_t (1 + \frac{1}{\varsigma_2}) \right], \tag{65}$$

$$\rho_{t+1} \theta \epsilon_2 = \beta \Psi \chi h_{2,t+1}^{\frac{1}{\varsigma_2}} \left[1 + \lambda_t (1 + \frac{1}{\varsigma_2}) \right], \tag{66}$$

$$\rho_t = \theta (1 + r) \rho_{t+1}, \tag{67}$$

$$\rho_t = c_{1,t}^{-\varsigma_1} (1 + \lambda_t (1 - \varsigma_1) - c_{1,t}^{-1} \lambda_t \varsigma_1 (1 + r)(1 - \tau_t)), \tag{68}$$

and

$$\theta \rho_{t+1} = \psi \beta c_{2,t+1}^{-\varsigma_1} (1 + \lambda_t (1 - \varsigma_1) - c_{1,t}^{-1} \lambda_t \varsigma_1 (1 + r)(1 - \tau_t)). \tag{69}$$

Combining the first order equations for the government's problem with consumption (equations 68 and 69)

yields

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1} = \Psi\beta(1+r)\left[\frac{1+\lambda_t(1-\varsigma_1)-c_{2,t+1}^{-1}\lambda_t\varsigma_1 Tr_{t+1}(1+r)(1-\tau_t)}{1+\lambda_t(1-\varsigma_1)-c_{1,t}^{-1}\lambda_t\varsigma_1 Tr_t(1+r)(1-\tau_t)}\right]. \tag{70}$$

Combining the government's first order conditions with respect to labor yields,

$$\frac{\rho_t\beta\psi}{\theta\epsilon_2\rho_{t+1|}}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\varsigma_2}} = 1. \tag{71}$$

Taking the ratio of the agent's first order conditions, equations 5 and 6, under the constant Frisch utility specification gives

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1}{\epsilon_2}\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\varsigma_2}}. \tag{72}$$

Combining equations 70, 71, and 72 yields the following expression for the optimal labor income taxes:

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1-\varsigma_1)-c_{2,t+1}^{-1}\lambda_t\varsigma_1 Tr_{t+1}(1+r)(1-\tau_t)}{1+\lambda_t(1-\varsigma_1)-c_{1,t}^{-1}\lambda_t\varsigma_1 Tr_t(1+r)(1-\tau_t)} \tag{73}$$

B.5 Restricted to Taxing Only Return on Accidental Bequests

The Lagrangian for this specification is

$$\begin{aligned}
\mathscr{L} =& \frac{c_{1,t}^{1-\varsigma_1}}{1-\varsigma_1} - \chi\frac{h_{1,t}^{1+\frac{1}{\varsigma_2}}}{1+\frac{1}{\varsigma_2}} + \beta\Psi\frac{c_{2,t+1}^{1-\varsigma_1}}{1-\varsigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\varsigma_2}}}{1+\frac{1}{\varsigma_2}} \\
&- \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\
&- \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\
&+ \lambda_t(-Tr_t c_{1,t}^{-\varsigma_1}(1+r(1-\tau_t)) - \beta\Psi Tr_{t+1}(1+r(1-\tau_t))c_{2,t+1}^{-\varsigma_1} + c_{1,t}^{1-\varsigma_1} + \psi\beta c_{2,t+1}^{1-\varsigma_1} - \chi h_{1,t}^{1+\frac{1}{\varsigma_2}} - \psi\beta\chi h_{2,t+1}^{1+\frac{1}{\varsigma_2}}).
\end{aligned} \tag{74}$$

The first order conditions with respect to labor, capital, and consumption are

$$\rho_t = \chi h_{1,t}^{\frac{1}{\varsigma_2}}\left[1 + \lambda_t(1+\frac{1}{\varsigma_2})\right], \tag{75}$$

$$\rho_{t+1}\theta\epsilon_2 = \beta\Psi\chi h_{2,t+1}^{\frac{1}{\varsigma_2}}\left[1 + \lambda_t(1+\frac{1}{\varsigma_2})\right], \tag{76}$$

$$\rho_t = \theta(1+r)\rho_{t+1}, \tag{77}$$

$$\rho_t = c_{1,t}^{-\varsigma_1}(1 + \lambda_t(1-\varsigma_1) - c_{1,t}^{-1}\lambda_t\varsigma_1(1+r(1-\tau_t))), \tag{78}$$

and

$$\theta\rho_{t+1} = \psi\beta c_{2,t+1}^{-\varsigma_1}(1 + \lambda_t(1 - \varsigma_1) - c_{1,t}^{-1}\lambda_t\varsigma_1(1 + r(1 - \tau_t))).\tag{79}$$

Combining the first order equations for the government's problem with consumption (equations 78 and 79) yields

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1} = \Psi\beta(1+r)\left[\frac{1 + \lambda_t(1 - \varsigma_1) - c_{2,t+1}^{-1}\lambda_t\varsigma_1 Tr_{t+1}(1 + r(1 - \tau_t))}{1 + \lambda_t(1 - \varsigma_1) - c_{1,t}^{-1}\lambda_t\varsigma_1 Tr_t(1 + r(1 - \tau_t))}\right].\tag{80}$$

Combining the government's first order conditions with respect to labor yields,

$$\frac{\rho_t\beta\psi}{\theta\epsilon_2\rho_{t+1|}}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\varsigma_2}} = 1.\tag{81}$$

Taking the ratio of the agent's first order conditions, equations 5 and 6, under the constant Frisch utility specification gives

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1}{\epsilon_2}\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\varsigma_2}}.\tag{82}$$

Combining equations 80, 81, and 82 yields the following expression for the optimal labor income taxes:

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t(1 - \varsigma_1) - c_{2,t+1}^{-1}\lambda_t\varsigma_1 Tr_{t+1}(1 + r(1 - \tau_t))}{1 + \lambda_t(1 - \varsigma_1) - c_{1,t}^{-1}\lambda_t\varsigma_1 Tr_t(1 + r(1 - \tau_t))}\tag{83}$$

B.6 One Rate on All Returns

The Lagrangian for this specification is

$$\begin{aligned}
\mathscr{L} =& \frac{c_{1,t}^{1-\varsigma_1}}{1 - \varsigma_1} - \chi\frac{h_{1,t}^{1+\frac{1}{\varsigma_2}}}{1 + \frac{1}{\varsigma_2}} + \beta\Psi\frac{c_{2,t+1}^{1-\varsigma_1}}{1 - \varsigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\varsigma_2}}}{1 + \frac{1}{\varsigma_2}}\\
&- \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2))\\
&- \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2))\\
&+ \lambda_t(-Tr_t c_{1,t}^{-\varsigma_1} - \frac{Tr_{t+1}c_{1,t}^{-2\varsigma_1}c_{2,t+1}^{\varsigma_1}}{\beta\Psi} + c_{1,t}^{1-\varsigma_1} + \psi\beta c_{2,t+1}^{1-\varsigma_1} - \chi h_{1,t}^{1+\frac{1}{\varsigma_2}} - \psi\beta\chi h_{2,t+1}^{1+\frac{1}{\varsigma_2}}).
\end{aligned}\tag{84}$$

The first order conditions with respect to labor, capital, and consumption are

$$\rho_t = \chi h_{1,t}^{\frac{1}{\varsigma_2}}\left[1 + \lambda_t(1 + \frac{1}{\varsigma_2})\right],\tag{85}$$

$$\rho_{t+1}\theta\epsilon_2 = \beta\Psi\chi h_{2,t+1}^{\frac{1}{\varsigma_2}}\left[1 + \lambda_t(1 + \frac{1}{\varsigma_2})\right],\tag{86}$$

$$\rho_t = \theta(1+r)\rho_{t+1},\tag{87}$$

48

$$\rho_t = c_{1,t}^{-\varsigma_1}(1 + \lambda_t(1 - \varsigma_1) + Tr_t\lambda_t\varsigma_1 c_{1,t}^{-1} + \frac{2Tr_{t+1}\lambda_t\varsigma_1}{\beta\Psi}c_{2,t+1}^{\varsigma_1}c_{1,t}^{-1-\varsigma_1}), \tag{88}$$

and

$$\theta\rho_{t+1} = c_{2,t+1}^{-\varsigma_1}(\Psi\beta(1 + \lambda_t(1 - \varsigma_1)) - \frac{Tr_{t+1}\lambda_t\varsigma_1}{\beta\Psi}c_{2,t+1}^{-1}c_{1,t}^{-2\varsigma_1}). \tag{89}$$

Combining the first order equations for the government's problem with consumption (equations 88 and 89) yields

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\varsigma_1} = \Psi\beta(1 + r)\left[\frac{1 + \lambda_t(1 - \varsigma_1) - \frac{Tr_{t+1}\lambda_t\varsigma_1}{\beta\Psi}^{-2}c_{2,t+1}^{-1}c_{1,t}^{-2\varsigma_1}}{1 + \lambda_t(1 - \varsigma_1) + \lambda_t\varsigma_1(Tr_t c_{1,t}^{-1} + \frac{Tr_{t+1}2c_{2,t+1}^{\varsigma_1}c_{1,t}^{-1-\varsigma_1}}{\beta\Psi}}\right] \tag{90}$$

References

Aiyagari, S. Rao, "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *The Journal of Political Economy*, December 1995, *103* (6), 1158–1175.

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.

Chamley, Christopher, "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, May 1986, *54* (3), 607–622.

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

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 and Dirk Krueger, "On the Optimal Progressivity of the Income Tax Code," *The Journal of Monetary Economics*, October 2006, *53* (7), 1425–1450.

_ , **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.

Correia, Isabel H., "Should Capital Income Be Taxed in the Steady State?," *Journal of Public Economics*, 1996, *60* (1), 147 – 151.

Domeij, David and Martin Flodén, "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.

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.

Fuster, Luisa, Ayse Imrohoroglu, and Selahattin İmrohoroğlu, "Altruism, incomplete markets, and tax reform," *Journal of Monetary Economics*, 2008, *55* (1), 65 – 90. Carnegie-Rochester Conference Series on Public Policy: Advances in Dynamic Public Finance April 20-21, 2007.

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

Gervais, Martin, Working Paper, University of Southampton February 2010.

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

Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan, "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," Working Paper 438, Federal Reserve Bank of Minneapolis November 2009.

Hansen, G.D., "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked," *Journal of Applied Econometrics*, January - March 1993, *8* (1), 71–80.

Heer, Burkhard and Alfred Maussner, *Dynamic General Equilibrium Modeling*, Berlin: Springer, 2005.

Hubbard, Glenn and Kenneth Judd, "Liquidity Constraints, Fiscal Policy, and Consumption," *Brookings Papers on Economic Activity*, 1986, (1), 1–59.

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

İmrohoroğlu, Selahattin, "A Quantitative Analysis of Capital Income Taxation," *International Economic Review*, May 1998, *39* (2), 307–328.

Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi, "On the Optimal Taxation of Capital Income," *Jounal of Economic Theory*, 1997, *73*, 93 – 117.

Judd, Kenneth, "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 1985, *28*, 59–83.

Lucas, Robert E. and Nancy L. Stokey, "Optimal Fiscal and Monetary Policy in an Economy without Capital," *Journal of Monetary Economics*, 1983, *12* (1), 55 – 93.

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

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

Peterman, William, "The Effect of Endogenous Human Capital Accumulation on Optimal Taxation," Finance and Economics Discussion Series 2012-03, Federal Reserve Boark of Governors 2012.

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.

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.

Smyth, Seamus, "A Balancing Act: Optimal Nonlinear Taxation in Overlapping Generations Models," Mimeo, Harvard May 2006.

Wallenius, Johanna, "Human Capital Accumulation and the Intertemporal Elasticity of Substitution of Labor: How Large is the Bias?," *Review of Economic Dynamics*, 2011, *14* (4), 577 – 591.