Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

# Taxing Capital? The Importance of How Human Capital is Accumulated

**William B. Peterman**

**2015-117**

Please cite this paper as:
Peterman, William B. (2015). "Taxing Capital? The Importance of How Human Capital is Accumulated," Finance and Economics Discussion Series 2015-117. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.117.

# Taxing Capital? The Importance of How Human Capital is Accumulated

William B Peterman*

December 17, 2015

**Abstract**

This paper considers the impact of how human capital is accumulated on optimal capital tax policy in a life cycle model. In particular, it compares the optimal capital tax when human capital is accumulated exogenously, endogenously through learning-*by*-doing, and endogenously through learning-*or*-doing. Previous work demonstrates that in a simple two generation life cycle model with exogenous human capital accumulation, if the utility function is separable and homothetic in each consumption and labor, then the government has no motive to condition taxes on age or tax capital. In contrast, this paper demonstrates analytically that adding either form of endogenous human capital accumulation creates a motive for the government to use age-dependent labor income taxes. Moreover, if the government cannot condition taxes on age, then a capital tax can be optimal in order to mimic such taxes. This paper quantitatively explores the strength of this channel and finds that, including human capital accumulation with learning-by-doing, as opposed to exogenously, causes the optimal capital tax to increase by between 7.3 and 14.5 percentage points. In contrast, introducing learning-or-doing causes a much smaller increase in the optimal capital tax of between 0.7 and 3.7 percentage points. Taken as a whole, this paper finds that the specific formulation by which human capital is accumulated can have notable implications on the optimal capital tax.

JEL: E24, E62, H21.

Key Words: Optimal Taxation, Capital Taxation, Human Capital.

*20th and C Street NW, Washington DC 20551. Tel: 202-452-3703. E-mail: william.b.peterman@frb.gov. Views expressed on this site are my own and do not reflect the view of the Federal Reserve System or its staff. For extensive discussions and helpful comments, I thank the anonymous referee, Vasia Panousi, Irina Telyukova, Valerie Ramey, Roger Gordon, and Scott Borger, as well as seminar participants at University of California at San Diego, Madrid Macroeconomic Workshop, the Federal Reserve Board of Governors, the Federal Reserve Bank of Philadelphia, the Eastern Economics Association Conference, the Missouri Economics Conference, the Midwestern Macroeconomics Conference, and the Conference in Computing in Economics and Finance.

# 1 Introduction

In their seminal works, Chamley (1986) and Judd (1985) determine that it is not optimal to tax capital in an infinitely-lived agent model. In contrast, Peterman (2013) and Conesa et al. (2009) demonstrate that in a life cycle model the optimal tax on capital is positive. The authors show that, in part, the non-zero optimal capital tax is driven by the government wanting to condition taxes on age due to variation in consumption and labor over the life cycle.[1] This variation in consumption and labor is partially due to fluctuations in an agent's productivity over his life cycle, or age-specific human capital. Despite the importance of age-specific human capital for the non-zero optimal capital tax result in life cycle models, previous research tends to assume that it is accumulated exogenously. One exception, Peterman (2015) demonstrates that incorporating a specific form of endogenous human capital accumulation can cause considerable effects on the optimal capital tax. However, less is known about how different forms of endogenous human capital accumulations affect the optimal capital tax. Thus, this paper revisits optimal capital taxation by, analytically and quantitatively, assessing the effect of various different human capital accumulation processes on the optimal capital tax. Overall, this paper finds that the way in which human capital is accumulated can have considerable effects on the optimal capital tax.

Specifically, this paper explores the change in the optimal capital tax when human capital is accumulated exogenously, endogenously with learning-*by*-doing (LBD), or endogenously with learning-*or*-doing (LOD). As opposed to being pre-determined with exogenous human capital accumulation, with LBD an agent acquires human capital by working. Alternatively, in LOD, which is also referred to as Ben Porath type skill accumulation or on-the-job training, an agent acquires human capital by spending time training in periods in which he is also working.[2] Thus, with LBD, an agent determines his level of age-specific human capital by choosing the hours he works, while with LOD, an agent determines his human capital by choosing the hours he trains. I analyze the effects of all three forms since each is commonly employed in quantitative life cycle models so understanding the effect on the optimal capital tax of the different human capital assumptions is important.[3]

First I analytically assess the implications of how human capital is accumulated in simple overlapping

[1] Atkeson et al. (1999), Erosa and Gervais (2002), and Garriga (2001) demonstrate this result analytically in a simple life cycle model.

[2] This paper does not evaluate the effect of formal education on optimal tax policy but instead focuses on human capital acquired after an individuals begins working. Although, the quantitative model is calibrate to exclude time spent in school, the mechanisms by which LOD changes the optimal tax policy would be similar for formal education. For a discussion of the effects of formal education on optimal taxation see Jacobs and Bovenberg (2009).

[3] Examples of life cycle studies that include these three forms of human capital accumulation are Conesa et al. (2009), Conesa and Krueger (2006), Huggett et al. (2007)Hansen and İmrohoroğlu (2009), Imai and Keane (2004), Chang et al. (2002), Jones et al. (1997), Jones and Manuelli (1999), Guvenen et al. (2009), Kuruscu (2006), Kapicka (2006), and Kapicka (2009).

generations model (OLG) model where the utility function is both separable and homothetic with respect to consumption and hours worked. Garriga (2001) finds that in this type of model with exogenous human capital accumulation the optimal tax policy does not include age-dependent taxes on labor income and the optimal capital tax is zero.[4] In contrast, I find adding LBD or LOD causes the optimal tax policy to include age-dependent taxes. Moreover, if age-dependent taxes are not available then a non-zero capital tax can be used to mimic the wedge created by conditioning labor income taxes on age. Specifically, a positive (negative) tax on capital can be used to impose the same wedge on the marginal rate of substitution as a relatively larger (smaller) tax on young labor income.

Adding LBD alters the optimal tax policy because it alters an agent's incentives to work over his life cycle. In a model with exogenous skill accumulation, an agent's only incentive to work is his wage. In a model with LBD, the benefits from working are current wages as well as an increase in future age-specific human capital. I refer to these benefits as the "wage benefit" and the "human capital benefit," respectively. The importance of the human capital benefit, which is unique to LBD, decreases as an agent approaches retirement. Thus, adding LBD causes the agent to supply labor relatively less elastically early in his life compared with later in his life. Relying more heavily on a capital tax reduces the distortions that this tax policy imposes on the economy, since it implicitly taxes this less elastically supplied labor income from agents when they are younger at a relatively higher rate than when they are older. I refer to this channel as the elasticity channel since it is transmitted through changes in the labor supply elasticity profile.

Adding LOD to the model also causes changes to the optimal capital tax. There are two channels through which LOD affects the optimal tax policy: the elasticity channel and the savings channel. First, adding LOD changes an agent's elasticity profile. Training is an imperfect substitute for labor as both involve forfeiting leisure in exchange for higher lifetime income. The substitutability of training decreases as an agent ages since he has less time to take advantage of the accumulated skills. Therefore, introducing LOD causes a young agent to supply labor relatively more elastically. With LOD, the elasticity channel causes the optimal capital tax to be lower in order to decreases the implicit taxes on labor income when agents are younger. The second channel, the savings channel, arises because training is an alternative method of saving, as opposed to accumulating physical capital. When the government taxes labor they implicitly decrease the desirability of saving via training as opposed to ordinary capital. In order to mitigate this distortion, the government increases the capital tax and reduces the labor tax. Since these two channels have counteracting effects, one cannot analytically determine the cumulative direction of their impact on the optimal tax policy.[5]

[4]A host of work demonstrates a similar set of results in a two generation model with a single cohort. Two examples of these works include Atkinson and Stiglitz (1976), and Deaton (1979).

[5]It is assumed that the government cannot directly tax human capital since it is unobservable.

Next, I quantitatively assess the effect of the form by which age-specific human capital is accumulated on the optimal capital tax in a calibrated life cycle model which includes exogenously determined retirement, a reduced form social security program, lifetime length uncertainty. I find that adding either form of endogenous human capital increases the optimal capital tax compared to the optimal capital tax with exogenous human capital accumulation.[6] When LBD is included, I find that the optimal tax on capital increases between 7.3 and 14.5 percentage points.[7] LBD has In contrast, when LOD is included, the optimal tax on capital increases only between 0.7 and 4.7 percentage points compared to the rates with exogenous human capital accumulation. Thus, the optimal tax on capital varies by up to 14.5 percentage points depending on how human capital accumulation is modeled. Therefore, this modeling choice is of first order importance when determining optimal capital tax policy.

This paper is generally related to a class of research which demonstrates that in a model where the government has an incomplete set of tax instruments a non-zero capital tax may be optimal in order to mimic the missing taxes (see Correia (1996), Armenter and Albanesi (2009), and Jones et al. (1997)). This paper combines two related strands of the literature within this class of research. The first strand examines the optimal capital tax in a calibrated life cycle model but does not assess the importance of how human capital is accumulated. Conesa et al. (2009), henceforth CKK, solve a calibrated life cycle model to determine the optimal capital tax in a model with exogenous human capital accumulation. They determine that the optimal tax policy is a flat 34 percent capital tax and a flat 14 percent labor income tax.[8] They state that a primary motive for imposing a high capital tax is to mimic a relatively larger labor income tax on younger agents when they supply labor relatively less elastically. An agent supplies labor more elastically as he ages because his labor supply is decreasing, and the authors use a utility specification in which the agent's Frisch labor supply elasticity is a negative function of hours worked. Peterman (2013) confirms that this is an economically significant motive for the positive capital tax in a model similar to CKK's model, but concludes that the restriction on the government from being able to tax accidental bequests at a different

[6] Unlike the analytically tractable model, I find that in the model with exogenous human capital accumulation the optimal tax on capital is between 18.2 and 31.8 percent depending on the form of the utility function. See Peterman (2013) for an in depth discussion of motives for a positive capital tax in calibrated OLG model with exogenous human capital accumulation.

[7] The range of the increases in the optimal capital tax is because I find that the size of the effect is different depending on the utility function.

[8] This is model M4 in Conesa et al. (2009). I refer to CKK's model that abstracts from idiosyncratic earnings risk and within-cohort heterogeneity because they find that these features do not affect the level of the optimal capital tax. Moreover, Peterman (2015) finds that the effect of LBD on the optimal capital tax with idiosyncratic earnings risk is similar to the effect in this paper. However, both studies find that the within cohort heterogeneity from idiosyncratic productivity shocks does affect the optimal progressivity of the labor tax. Therefore, I exclude these features in my benchmark analysis in order to focus on the effects of human capital accumulation on the optimal capital tax and abstract form the effect of endogenous human capital accumulation on the optimal progressivity.

rate from ordinary capital income is also a large contribution to the positive optimal capital tax.[9] Moreover, Cespedes and Kuklik (2012) find that when a non-linear mapping between hours and wages is incorporated into a model similar to CKK hours become more persistent and the optimal capital tax fall significantly. All of these studies assume human capital is accumulated exogenously. Thus, this paper extends these previous life cycle studies of optimal tax policy by determining how all three forms of human capital accumulation affect the optimal capital tax policy.

This paper is related to a second strand of the literature that analyzes the effect of how human capital is accumulated on the tradeoff between labor and capital taxes but not in a life cycle model.[10] For example, both Jones et al. (1997) and Judd (1999) examine optimal capital tax in an infinitely lived agent model in which agents are required to use market goods to acquire human capital similar to ordinary capital. They find that if the government can distinguish between pure consumption and human capital investment, then, similar to a model with exogenous human capital accumulation, it is not optimal to distort either human or physical capital accumulation in the long run. Moreover, Reis (2007) shows in a similar model that if the government cannot distinguish between consumption and human capital investment, then similar to a model with exogenous human capital accumulation, the optimal capital tax is still zero as long as the level of capital does not influence the relative productivity of human capital. Chen et al. (2010) find in an infinitely lived agent model with labor search, that including endogenous human capital accumulation causes the optimal capital tax to increase, relative to a model with exogenous human capital accumulation, because a higher capital tax unravels the labor market frictions.[11] This second strand of literature is unable to account for the effects of endogenous human capital accumulation through life cycle channels. Since CKK and Peterman (2013) demonstrate that these life cycle channels are quantitatively important for motivating a positive capital tax this paper includes them. Thus, this paper combines both strands of the literature and determines the effect on optimal capital tax policy of how human capital is accumulated in a life cycle model.[12]

[9]Further work, such as, Karabarbounis (2012) and Peterman (2012), demonstrate that incorporating endogenous fluctuations in labor supply on the extensive margin can enhance this motive for the government to use a capital tax to mimic age-dependent taxes on labor income.

[10]In a related paper, Best and Kleven (2012) examine how introducing LBD changes the optimal general income tax in a model without savings. Best and Kleven (2012) show that introducing LBD causes the government to change the progressivity of the tax rates such that the relative tax on young income increases. This result is similar to the result in this paper. However, in this paper when the government can use either a progressive tax on labor or a non-zero tax on capital to mimic age-dependent taxes they choose the tax on capital.

[11]The labor market frictions in Chen et al. (2010) cause a lower level of employment in their economy. A capital tax causes the wage discount to increase, thus causing firms to post more vacancies which in turn causes an increase in worker participation.

[12]Two exceptions that examine the effects of endogenous human capital accumulation in a life cycle model are Peterman (2015) and da Costa and Santos (2015) on the optimal tax policy. However, these other two studies include within cohort heterogeneity and focus on the optimal progressivity of the labor tax and optimal level of the capital tax that balances equity versus efficiency. In contrast, this study excludes the within cohort heterogeneity and focuses on how the optimal capital tax that maximizes efficiency

This paper is organized as follows: Section 2 examines an analytically tractable version of the model to demonstrate that including endogenous human capital accumulation creates a motive for the government to condition labor income taxes on age. Section 3 describes the full model and the competitive equilibrium used in the quantitative exercises. The calibration and functional forms are discussed in section 4. Section 5 describes the computational experiment, and section 6 presents the results. Section 7 tests the sensitivity of the results with respect to calibration parameters and utility specifications, while section 8 concludes.

# 2 Analytical Model

In this section, I demonstrate that adding endogenous human capital accumulation overturns the result from Garriga (2001). In particular including either form of endogenous human capital accumulation in a model with a utility function that is separable and homothetic in each consumption and labor the government creates an incentive to condition labor income taxes on age. In contrast, Garriga (2001) finds that with exogenous human capital accumulation the government does not want to condition labor income taxes on age.[13] It is useful to determine if the government wants to use age-dependent taxes because both Garriga (2001) and Erosa and Gervais (2002) show that if the government wants to condition taxes on age and cannot do so, then the optimal capital tax will generally be non-zero in order to mimic this age-dependent tax.

I derive these analytical results in a tractable two-period version of the computational model. For tractability purposes, the features I abstract from include: retirement, population growth, progressive tax policy, and conditional survivability. Additionally, in order to focus on the life cycle elements of the model I assume that the marginal products of capital and labor are constant and thus factor prices are constant. 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.

## 2.1 Exogenous Age-Specific Human Capital

### 2.1.1 General Set-up

In the analytically tractable model, agents live with certainty for two periods, and their preferences over consumption and labor are represented by

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

changes with different forms of human capital is accumulation.

[13] A similar set of results for the exogenous and LBD model are in Peterman (2015). I include them in this paper for completeness.

where $\beta$ is the discount rate, $c_{j,t}$ is the consumption of an age $j$ agent at time $t$, and $h_{j,t}$ is the percent of the time endowment the agent works.[14] Age-specific human capital is normalized to unity when the agent is young. At age two, age-specific human capital is $\varepsilon_2$. The agent maximizes equation 1 with respect to consumption and hours subject to the following constraints,

$$c_{1,t} + a_{1,t} = (1 - \tau_{h,1}) h_{1,t} w \tag{2}$$

and

$$c_{2,t+1} = (1 + r(1 - \tau_k)) a_{1,t} + (1 - \tau_{h,2}) \varepsilon_2 h_{2,t+1} w, \tag{3}$$

where $a_{1,t}$ is the amount young agents save, $\tau_{h,j}$ is the tax rate on labor income for an agent of age $j$, $\tau_k$ is the tax rate on capital income, $w$ is the efficiency wage for labor services, and $r$ is the rental rate on capital. I assume that the tax rate on labor income can be conditioned on age; however, the tax rate on capital income cannot.[15] I combine equations 2 and 3 to form a joint intertemporal budget constraint,

$$c_{1,t} + \frac{c_{2,t+1}}{1 + r(1 - \tau_k)} = w(1 - \tau_{h,1}) h_{1,t} + \frac{w(1 - \tau_{h,2}) \varepsilon_2 h_{2,t+1}}{1 + r(1 - \tau_k)}. \tag{4}$$

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

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

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -w\varepsilon_2(1 - \tau_{h,2}), \tag{6}$$

and

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

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

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

[15] Agents only live for two periods in the analytically tractable model so they choose not to save when they are old. Therefore, in this model the restriction on the capital tax policy is not binding.

### 2.1.2 Tax on Capital Mimics Age-Dependent Tax on Labor

When examining the optimal capital tax it useful to determine if it is optimal to condition labor income taxes on age. Examining the intertemporal Euler equation,

$$\varepsilon_2 \frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1+r(1-\tau_k))\frac{1-\tau_{h,1}}{1-\tau_{h,2}}. \tag{8}$$

it is clear that if the government wants to create a wedge on the marginal rate of substitution by varying the labor income tax rate by age, then $\tau_k$ is an alternative option. A positive (negative) capital tax induces a wedge on the marginal rate of substitution that is similar to a relatively larger tax on young (old) labor income. Thus, throughout the analytical analysis, it will of importance to determine whether age-dependent labor income taxes are optimal.

### 2.1.3 Optimal Tax Policy

Next, I solve for the optimal tax policy in the exogenous model, with a benchmark utility function that is homothetic with respect to consumption and hours worked, $U(c,h) = \frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$. I solve for the optimal tax policy using the primal approach which implies I solve for the optimal allocation (see Appendix A.1 for details of the approach). From this optimal allocation, one can determine the optimal tax policy. In particular, I find that the optimal allocation implies the following ratio for the optimal labor taxes (see Appendix A.2 for the formulation of the problem),

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})} = 1. \tag{9}$$

Equation 9 demonstrates that in this model the government has no incentive to condition labor income taxes on age when age-specific human capital is exogenous.[16]

Moreover, using the primal approach, the optimal allocation of consumption is represented by the following expression,

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r). \tag{10}$$

[16] $\lambda$ is the Lagrange multiplier on the implementability constraint. See Appendix A.1 for more details. This result is specific to this utility function. See Garriga (2001) for further details.

Assuming the benchmark utility function, the optimal allocation indicated by the primal approach is,

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r(1-\tau_k)). \tag{11}$$

Thus, the optimal capital tax is zero. As Garriga (2001) points out, since there is no desire to condition labor income taxes on age in this exogenous model, the optimal tax on capital is zero regardless of whether the government can condition labor income taxes on age.[17]

## 2.2 Learning-by-Doing

### 2.2.1 Including LBD Creates Motive for Age-Dependent Taxes on Labor Income

Next, I examine the LBD model. In the LBD model, age-specific human capital for a young agent is normalized to one. Age-specific human capital for an old agent is determined by the function $s_2(h_{1,t})$. The function $s_2(h_{1,t})$ is a positive and concave function of the hours worked when young. In this model agents maximize the same utility function subject to,

$$c_{1,t} + a_{1,t} = (1-\tau_{h,1})h_{1,t}w, \tag{13}$$

and

$$c_{2,t+1} = (1+r(1-\tau_k))a_{1,t} + (1-\tau_{h,2})s_2(h_{1,t})h_{2,t+1}w. \tag{14}$$

The agent's first order conditions are given by,

$$\frac{U_{h1}(t)}{U_{c1}(t)} = -[w(1-\tau_{h,1}) + \beta\frac{U_{c2}(t+1)}{U_{c1}(t)}w(1-\tau_{h,2})h_{2,t+1}s_{h1}(t+1))], \tag{15}$$

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -ws_2(h_{1,t})(1-\tau_{h,2}), \tag{16}$$

and

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1+r(1-\tau_k)). \tag{17}$$

[17] When the government cannot condition labor income taxes on age then the Lagrangian includes an additional constraint,

$$\varepsilon_2\frac{U_{h1}(t)}{U_{c1}(t)} = \frac{U_{h2}(t+1)}{U_{c2}(t+1)}. \tag{12}$$

However, in the analytically tractable model with exogenous human capital accumulation, this constraint is not binding and thus the Lagrange multiplier on this constraint would be equal to zero.

The first order conditions with respect to $h_2$ and $a_1$ are similar in the LBD (equations 16 and 17) and exogenous models (equations 6 and 7). However, the first order condition with respect to $h_1$ is different in the two models (equations 15 and 5) because working has the additional human capital benefit in the LBD model.

The formulation for the government's problem and the resulting first order conditions (utilizing the benchmark utility function) are in appendix A.3. The optimal allocation is represented by,

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}} = \frac{\left(1+\lambda_t(1+\frac{1}{\sigma_2})-\lambda_t(1+\frac{1}{\sigma_2})\frac{h_{1,t}s_{h1}(t+1)}{s_2}\right)\left(1+\frac{h_{2,t+1}s_2}{1+r(1-\tau_k)}\right)}{1+\lambda_t(1+\frac{1}{\sigma_2})+h_{2,t+1}^{1+\frac{1}{\sigma_2}}h_{1,t}^{1+\frac{-1}{\sigma_2}}\frac{\lambda_t}{s_2}\left(\frac{s_{h1}(t+1)}{s_2}-s_{h1,h1}(t+1)\right)} - \frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)}, \tag{18}$$

where $s_{h1}$ represents the partial derivative of the skill function for an older agent with respect to hours worked when young. Equation 18 demonstrates that generally in the LBD model the government has an incentive to condition labor income taxes on age. Moreover, they will generally want to tax labor income at a relatively higher rate when agents are young.[18] This result contrasts with the exogenous model, in which the government has no incentive to condition labor income taxes on age (see equation 9). As Garriga (2001), Erosa and Gervais (2002), and Peterman (2013) demonstrate, if the government wants to condition labor income taxes on age but age-dependent taxes are not allowed then the government will typically use a non-zero capital tax to mimic this type of age-dependent tax policy.

### 2.2.2 LBD Enhances Motive for Positive Tax on Capital

In order to get a sense of why the government wants to tax labor income when an agent is young at a relatively higher rate, I examine the intertemporal Euler equation (determined by combining equations 15, 16 and 17),

$$s_2(h_{1,t})\frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1+r(1-\tau_k))\frac{1-\tau_{h,1}}{1-\tau_{h,2}} + \beta h_{2,t+1}s_{h1}(t+1). \tag{19}$$

Including LBD causes the intertemporal Euler equation to have an extra positive term on the right hand side (see equation 8 and equation 19). Therefore, holding all else equal and setting $\varepsilon_2 = s_2$, the tax on young labor income would need to be relatively higher in order to induce the same wedge on the marginal rate of substitution in the LBD model.

Examining the Frisch elasticities in the exogenous and LBD models, provides the intuition why adding

[18] In particular, the relative tax on young labor is higher than the tax on old labor income as long as $\lambda$ is positive.

LBD increases the optimal relative tax on young labor income or tax on capital. Since the functional forms of these elasticities extend to a model where agents live for more than two periods, I denote an agent's age with $i$. In the exogenous model, the Frisch elasticity simplifies to $\Xi_{\text{exog}} = \sigma_2$. The Frisch elasticity in the LBD model is, $\Xi_{\text{LBD}} = \frac{\sigma_2}{1-\frac{h_{i+1,t+1}w_{t+1}(h_{i,t}\sigma_2 s_{hi,hi}(t+1)-s_{hi}(t+1))}{s_{i,t}(1+r_t(1-\tau_k))w_t}}$.[19]

The Frisch elasticity in the exogenous model is constant and valued at $\sigma_2$. In the LBD model, the extra terms in $\Xi_{\text{LBD}}$ increase the size of the denominator, thus holding hours and consumption constant between the two models, $\Xi_{\text{exog}} > \Xi_{\text{LBD}}$. Intuitively, the inclusion of the human capital benefit makes workers less responsive to a one-period change in wages since the wage benefit is only part of their total compensation for working in the LBD model. Moreover, the human capital benefit does not have a constant effect on an agent's Frisch elasticity over his lifetime. The relative importance of the human capital benefit decreases as an agent ages because he has fewer periods to use his human capital.[20] Therefore, adding LBD causes an agent to supply labor relatively less elastically when they are young than when they are old. This shift in relative elasticities creates an incentive for the government to tax the labor income when agents are younger at a relatively higher rate. Thus, if the government cannot condition labor income taxes on age, then the optimal capital tax will be higher in the LBD model to mimic this age-dependent tax. I use the term "elasticity channel" to describe the effect on optimal tax policy caused by a change in the Frisch elasticity from including endogenous human capital. The elasticity channel is responsible for the change in optimal tax policy from including LBD.[21]

## 2.3 Learning-or-Doing

### 2.3.1 Including LOD Creates Motive for Age-Dependent Taxes on Labor Income

Next, I examine the LOD model to demonstrate that this form of endogenous age-specific human capital accumulation also creates a motive for the government to condition labor income taxes on age. Similar to the other models, age-specific human capital for a young agent is normalized to one. Age-specific human capital for an old agent is determined by the function $s_2(n_{1,t})$ which is a positive and concave function

[19] Since this is the Frisch elasticity with respect to a temporary increase in the wage, one must distinguish between $w_t$ and $w_{t+1}$.

[20] For the human capital benefit to decline over the lifetime, it is sufficient to assume agents work for a finite number of periods.

[21] Alternative intuition for this result can be demonstrated in the commodity tax framework of Corlett and Hague (1953). In their static framework, the government wants to tax leisure. However, if they cannot directly tax leisure, the government will tax commodities that are more complementary to leisure at a higher rate. Viewing this simple two generation model in that framework, adding LBD raises the relative opportunity cost of leisure when agents are young so young labor is less of substitute (more of a complement) with leisure. This change leads the government to want to increase the tax on young labor. Moreover, if the government cannot use age-dependent taxes then they will increase the tax on capital to implicitly tax consumption from the old at a relatively higher rate since LBD makes consumption and leisure more complementary for the older agents than the younger agents.

of the hours spent training when an agent is young ($n_{1,t}$). In the LOD model, I need a utility function that incorporates training. I alter the benchmark utility specification so that it consistently incorporates the disutility of non-leisure activities, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h+n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$. In this model agents maximize their utility function subject to,

$$c_{1,t} + a_{1,t} = (1-\tau_{h,1})h_{1,t}w, \tag{20}$$

and

$$c_{2,t+1} = (1+r(1-\tau_k))a_{1,t} + (1-\tau_{h,2})s_2(n_{1,t})h_{2,t+1}w. \tag{21}$$

The agent's first order conditions are given by,

$$\frac{U_{h1}(t)}{U_{c1}(t)} = -[w(1-\tau_{h,1})], \tag{22}$$

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -ws_2(n_{1,t})(1-\tau_{h,2}), \tag{23}$$

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1+r(1-\tau_k)), \tag{24}$$

and

$$\frac{U_{n1}(t)}{U_{c2}(t+1)} = -\beta w(1-\tau_{h,2})s_{n1}(n_{1,t})h_{2,t+1}. \tag{25}$$

The first order conditions with respect to $h_1$, $h_2$, and $a_1$ are similar in the LOD model (equations 22, 23, and 24) and the exogenous model (equations 5, 6, and 7). However, since agents have the additional choice variable $n_1$ in the LOD model, there is an additional first order condition (equation 25).

The formulation of the government's problem and resulting first order conditions are provided in appendix A.4.[22] Combing the first order conditions yields the following relationship for optimal taxes on labor income,

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)}. \tag{26}$$

Equation 26 demonstrates that the government generally has an incentive to condition labor income taxes on age when LOD is introduced into the model.

Although equation 26 shows that including LOD creates an incentive for the government to condition labor income taxes on age, it is unclear at which age the government wants to impose a relatively higher labor income tax. Comparing equations 9 and 26, there are two channels through which introducing LOD

[22] $\eta_t$ is the Lagrange multiplier on an additional constraint that is included to ensure that in the optimal allocations both 23 and 25 are respected.

changes the optimal tax policy. The first channel results from using a utility function that is non separable in training and labor. The non separability affects the optimal tax policy through the elasticity channel since it causes LOD to alter the Frisch elasticity. This channel causes the numerator of the ratio to include the additional term $\frac{h_{1,t}}{h_{1,t}+n_{1,t}}$. As a result of this new term, the expression decreases.

The second channel results from the intertemporal link created because agents can save not only with ordinary savings in this model but also can save via training. I refer to this channel as the savings channel. This second channel causes the inclusion of the additional terms $-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)$ and $\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}$ in the denominator and numerator, respectively.[23] Assuming that $\eta_t$ is positive, these additional terms cause the expression to increase.[24] Thus, the two channels may have opposing effects on the optimal tax policy, and the overall effect is unclear.

Examining the Frisch labor supply elasticities provides intuition for how the first channel affects the optimal tax policy. Since the altered utility function is not additively separable in time spent working and training, the Frisch labor supply elasticity is not constant in the LOD model. The Frisch elasticity for the altered utility function is $\Xi_{\text{LOD}} = \frac{\sigma_2(h+n)}{h}$. This functional form implies that an agent supplies labor relatively more elastically with LOD than with exogenous human capital accumulation because the agent has a substitute for working in the form of training. Additionally, the effect on the Frisch elasticity is larger when he spends a larger proportion of his non-leisure time training (or when training is a better substitute for generating lifetime income). Therefore, if an agent spends less time training as he ages, then he will supply labor relatively more elastically when he is young, and the government would want to tax the labor income from agents when they are young at a relatively lower rate. Decreasing the tax on capital mimics this type of age-dependent tax. Thus, the elasticity channel from LOD can causes a decrease in the optimal capital tax.

Examining an agent's first order condition with respect to training demonstrates how the savings channel affects the optimal tax policy. An agent optimizes his choices such that the marginal disutility of training when he is young equals the marginal benefit of training ($U_{n1}(t) = \frac{U_{c1}(t)w(1-\tau_{h,2})h_{2,t+1}s_{n1}(t+1)}{1+r(1-\tau_k)}$). The marginal benefit is increased by raising the tax on capital or by decreasing the tax on older labor income. By adopting either of these changes, the government makes it relatively more beneficial for the agent to use training to save as opposed to ordinary capital.[25]

[23] The term in the numerator comes from both the intertemporal link and the nonseparability of the utility function. However, I group both terms in the savings channel because the impact on the optimal tax policy will be in the same direction as the other term.

[24] The sign of $\eta$ will depend on whether the government wants to increase the relative incentive to save with training or capital. If $\eta$ is positive, it implies that the government wants to increase the relative incentive to save with training. I generally find in the computational simulations that $\eta$ is positive and therefore treat it as positive in the exposition.

[25] An additional reason that the government wants to increase the capital tax is to unwind the distortion to savings behavior that are induced by a positive labor income tax.

# 3 Computational Model

Next, I determine the direction and magnitude of the effect of how human capital is accumulated on optimal capital tax policy in a less parsimonious model. I solve for the optimal tax policy in separate versions of the model with exogenous human capital accumulation, LBD and LOD. CKK and Peterman (2015) find that idiosyncratic earnings risk and heterogenous ability types can affect the optimal progressivity of the labor tax but do not affect the optimal capital tax. Thus, I exclude these sources of heterogeneity in my model in order to focus on the mechanisms that may affect the optimal capital tax.

## 3.1 Demographics

In the computational model, time is assumed to be discrete. Agents enter the model when they start working at the age of 20, and can live to a maximum age of J. Thus, the model is populated with J-19 overlapping generations. Conditional on being alive at age $j$, $\Psi_j$ is the probability of an agent living to age $j+1$. If an agent dies with assets, the assets are confiscated by the government and distributed equally to all the living agents as transfers ($Tr_t$). All agents are required to retire at an exogenously set age $j_r$.

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

$$\mu_j = \frac{\Psi_{j-1}}{1+n}\mu_{j-1}, \text{for } i = 2, ...., J, \tag{27}$$

where $\mu_1$ is normalized such that

$$\sum_{j=20}^{J} \mu_j = 1. \tag{28}$$

## 3.2 Individual

An individual is endowed with one unit of productive time per period that he divides between leisure and non-leisure activities. In the exogenous and LBD models the non-leisure activity is providing labor. In the LOD model the non-leisure activities include training and working. An agent chooses consumption as well as how to spend his time endowment in order to maximize his lifetime utility,

$$u(c_j, h_j + n_j) + \sum_{s=20}^{J-j-1} \beta^s \prod_{q=1}^{s} (\Psi_q) u(c_{s+1}, h_{s+1} + n_{s+1}), \tag{29}$$

where $c_j$ is the consumption of an agent at age $j$, $h_j$ is the hours spent working, $n_j$ is the time spent training, and $\beta$ is the discount factor conditional on surviving.

In the exogenous model, an agent's age-specific human capital is $\varepsilon_j$. In the endogenous models, an agent's age-specific human capital, $s_j$, is endogenously determined. In the LBD model, $s_j$ is a function of a skill accumulation parameter ($\Omega_{j-1}$), previous age-specific human capital ($s_{j-1}$), and time worked in the previous periods($h_{j-1}$), denoted by $s_j = S_{\text{LBD}}(\Omega_{j-1}, s_{j-1}, h_{j-1})$. In the LOD model, $s_j$ is a function of a skill accumulation parameter ($\Omega_{j-1}$), previous age-specific human capital ($s_{j-1}$), and time spent training ($n_{j-1}$), denoted by $s_j = S_{\text{LOD}}(\Omega_{j-1}, s_{j-1}, n_{j-1})$. The sequence of skill accumulation parameters $\{\Omega_j\}_{j=20}^{j_r-1}$ are calibration parameters set so that in the endogenous model, under the baseline-fitted U.S. tax policy, the agent's choices result in an agent having the same age-specific human capital as in the exogenous model. Individuals command a labor income of $h_j \varepsilon_j w_t$ in the exogenous model and $h_j s_j w_t$ in the endogenous model. Agents split their income between consumption and saving using a risk-free asset. An agent's level of assets is denoted $a_j$, and the asset pays a pre-tax net return of $r_t$.

### 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,

$$C_t + K_{t+1} - (1-\delta)K_t + G_t \leq K_t^{\alpha} N_t^{1-\alpha}, \tag{30}$$

where $K_t$, $C_t$, and $N_t$ represent the aggregate capital stock, aggregate consumption, and aggregate labor (measured in efficiency units), respectively. Additionally, $\alpha$ is the capital share and $\delta$ is the depreciation rate for physical capital. Unlike the analytically tractable model, I do not assume a linear production function in the computational model, so prices are determined endogenously and fluctuate with regard to the aggregate capital and labor.

### 3.4 Government Policy

The government has two fiscal instruments to finance its unproductive consumption, $G_t$.[26] First, the government taxes capital income, $y_k \equiv r_t(a + Tr_t)$, according to a capital income tax schedule $T^K[y_k]$. Second, the government taxes 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

[26] As opposed to assuming $G_t$ is unproductive, including $G_t$ such that it enters the agent's utility function in an additively separable manner will result in the same optimal tax policies.

$T^l[y_l]$. I impose four restrictions on the labor and capital income tax policies. First, I assume human capital is unobservable, meaning that the government cannot tax human capital accumulation. Second, I assume the rates cannot be age-dependent. Third, both of the taxes are solely functions of the individual's relevant taxable income in the current period. Finally, I exclude the use of lump sum taxes.

In addition to raising resources for consumption in the unproductive sector, the government runs a pay-as-you-go (PAYGO) social security system. I include a simplified social security program in the model because Peterman (2013) demonstrates that excluding this type of program in a model with exogenously determined retirement causes unrealistic life cycle profiles and can alter the optimal tax policy. In this reduced-form social security program, the government pays $SS_t$ to all individuals that are retired. Social security benefits are determined such that retired agents receive an exogenously set fraction, $b_t$, of the average income of all working individuals.[27] 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 assure that 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.

### 3.5 Definition of Stationary Competitive Equilibrium

In this section I define the competitive equilibrium for the exogenous model. See appendix B for the definition of the competitive equilibriums in the endogenous models.

Given a social security replacement rate $b$, a sequence of exogenous age-specific human capital $\{\varepsilon_j\}_{j=20}^{j_r-1}$, government expenditures $G$, and a sequence of population shares $\{\mu_j\}_{j=20}^{J}$, a stationary competitive equilibrium in the exogenous model consists of the following: a sequence of agent allocations, $\{c_j, a_{j+1}, h_j\}_{j=20}^{J}$, a production plan for the firm $(N, K)$, a government labor tax function $T^l : \mathbb{R}_+ \to \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \to \mathbb{R}_+$, a social security tax rate $\tau_{ss}$, a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, social security benefits $SS$, prices $(w, r)$, and transfers $Tr$ such that:

1. Given prices, policies, transfers, and benefits, the agent maximizes equation 29 subject to
$$c_j + a_{j+1} = w\varepsilon_j h_j - \tau_{ss} w\varepsilon_j h_j, +(1+r)(a_j + Tr) - T^l[w\varepsilon_j h_j(1-.5\tau_{ss})] - T^k[r(a_j + Tr)], \tag{31}$$
for $j < j_r$, and
$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k[r(a_j + Tr)], \tag{32}$$
for $j \geq j_r$.
Additionally,
$$c \geq 0, 0 \leq h \leq 1, a_j \geq 0, a_{20} = 0. \tag{33}$$

[27] Although an agent's social security benefits are a function of the average income of all workers, since all agents are homogenous within a cohort, the benefits are directly related to an individual's personal earnings history.

2. Prices $w$ and $r$ satisfy

$$r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \tag{34}$$

and

$$w = (1-\alpha)\left(\frac{K}{N}\right)^{\alpha}. \tag{35}$$

3. The social security policies satisfy

$$SS = b\frac{wN}{\sum_{j=20}^{j_r-1}\mu_j} \tag{36}$$

and

$$\tau_{ss} = \frac{ss\sum_{j=j_r}^{J}\mu_j}{w\sum_{j=20}^{j_r-1}\mu_j}. \tag{37}$$

4. Transfers are given by

$$Tr = \sum_{j=20}^{J}\mu_j(1-\Psi_j)a_{j+1}. \tag{38}$$

5. Government balances its budget

$$G = \sum_{j=20}^{J}\mu_j T^k[r(a_j+Tr)] + \sum_{j=20}^{j_r-1}\mu_j T^l[w\varepsilon_j h_j(1-.5\tau_{ss})]. \tag{39}$$

6. The market clears

$$K = \sum_{j=20}^{J}\mu_j a_j, \tag{40}$$

$$N = \sum_{j=20}^{J}\mu_j \varepsilon_j h_j, \tag{41}$$

and

$$\sum_{j=20}^{J}\mu_j c_j + \sum_{j=20}^{J}\mu_j a_{j+1} + G = K^{\alpha}N^{1-\alpha} + (1-\delta)K. \tag{42}$$

# 4 Calibration and Functional Forms

To determine the optimal tax policy, it is necessary to choose functional forms and calibrate the model's parameters. Calibrating the models involves a two-step process. The first step is choosing parameter values for which there are direct estimates in the data. These parameter values are in Table 1. Second, to calibrate the remaining parameters, values are chosen so that under the baseline-fitted U.S. tax policy certain targets in the model match the values observed in the U.S. economy.[28] These values are in Table 2.

[28] Since these are general equilibrium models, changing one parameter will alter all the values in the model that are used as targets. However, I present targets with the parameter that they most directly correspond to.

Table 1: **Calibration Parameters**

| Parameter | Value | Target |
|---|---|---|
| **Demographics** | | |
| Retire Age: $j_r$ | 65 | By Assumption |
| Max Age: $J$ | 100 | By Assumption |
| Surv. Prob: $\Psi_j$ | Bell and Miller (2002) | Data |
| Pop. Growth: $n$ | 1.1% | Data |
| **Firm Parameters** | | |
| $\alpha$ | .36 | Data |
| $\delta$ | 8.33% | $\frac{I}{Y} = 25.5\%$ |
| A | 1 | Normalization |

Adding endogenous human capital accumulation to the model fundamentally changes the model. Accordingly, if the calibration parameters are the same, then the value of the targets will be different in the endogenous and exogenous models. Thus, I calibrate the set of parameters based on targets separately in the three models. This calibration implies that these parameters are different in the three models.

### 4.1 Demographics

Agents enter the model at age of 20 when they begin to work and are exogenously forced to retire at a real world age of 65. If an individual survives until the age of 100, they die the next period. I set the conditional survival probabilities in accordance with the estimates in Bell and Miller (2002) and assume a population growth rate of 1.1 percent.

Table 2: **Calibration Parameters**

| Parameter | Exog. | LBD | LOD | Target |
|---|---|---|---|---|
| **Calibration Parameters** | | | | |
| Conditional Discount: $\beta$ | 0.995 | 0.993 | 0.997 | $K/Y = 2.7$ |
| Unconditional Discount: $\Psi_j\beta$ | 0.982 | 0.980 | 0.984 | $K/Y = 2.7$ |
| Risk aversion: $\sigma_1$ | 2 | 2 | 2 | CKK |
| Frisch Elasticity: $\sigma_2$ | 0.5 | 0.73 | 0.47 | Frisch= $\frac{1}{2}$ |
| Disutility of Labor: $\chi$ | 61 | 46 | 80 | Avg. $h_j + n_j = \frac{1}{3}$ |
| **Government Parameters** | | | | |
| $\Upsilon_0$ | .258 | .258 | .258 | Gouveia and Strauss (1994) |
| $\Upsilon_1$ | .768 | .768 | .768 | Gouveia and Strauss (1994) |
| G | 0.137 | 0.136 | 0.13 | 17% of Y |
| b | 0.5 | 0.5 | 0.5 | CKK |

## 4.2 Preferences

Agents have time-separable preferences over consumption and labor services. I use the benchmark utility function for the exogenous and LBD models, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$, and an altered form of this utility function for the LOD model, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h+n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$.[29]

I determine $\beta$ such that the capital-to-output ratio matches U.S. data of 2.7.[30] One reason that the reduced form social security program is included is to capture the relevant savings motives that affect the capital to output ratio. I determine $\chi$ such that under the baseline-fitted U.S. tax policy, agents spend on average one third of their time endowment in non-leisure activities. Following CKK, I set $\sigma_1 = 2$, which controls the relative risk aversion.[31] Past micro-econometric studies (such as Altonji (1986), MaCurdy (1981), and Domeij and Flodén (2006)) estimate the Frisch elasticity to be between 0 and 0.5. However, more recent research has shown that these estimates may be biased downward. Reasons for this bias include: utilizing weak instruments; not accounting for borrowing constraints; disregarding the life cycle effect of endogenous-age specific human capital; omitting correlated variables such as wage uncertainty; ignoring secondary earners; and not accounting for labor market frictions.[32] Therefore, I set $\sigma_2$ such that the Frisch elasticity is at the upper bound of the range (0.5). The preference parameters are summarized in table 2.

## 4.3 Age-Specific Human Capital

The age-specific human capital parameters are different in the three models. In the exogenous model, I set $\{\varepsilon_j\}_{j=20}^{j_r-1}$ so that the sequence matches a smoothed version of the relative hourly earnings estimated by age in Hansen (1993). In the endogenous models, I use the functional forms from Hansen and İmrohoroğlu (2009). Specifically, in the LBD model, agents accumulate age-specific human capital according to the following process,

$$s_{j+1} = \Omega_j s_j^{\Phi_1} h_j^{\Phi_2}, \tag{43}$$

where $s_j$ is the age-specific human capital for an agent at age $j$, $\Omega_j$ is an age-specific calibration parameter, $\Phi_1$ controls the importance of an agent's current human capital on LBD, and $\Phi_2$ controls the importance of time worked on LBD. In the LOD model, agents accumulate human capital according to the following

[29] Using this benchmark utility function for the exogenous, which is homothetic and separable, implies that the Frisch labor supply elasticity is constant as opposed to being a function of the level of labor supply. This flexibility allows me to isolate the effects of each of the channels on the optimal tax policy.

[30] This is the ratio of fixed assets and consumer durable goods, less government fixed assets to GDP (CKK).

[31] Even though CKK use a different utility specification, their specification has a parameter that corresponds to $\sigma_1$.

[32] Some of these studies include Imai and Keane (2004), Domeij and Flodén (2006), Pistaferri (2003), Chetty (2009), Peterman (2016), and Contreras and Sinclair (2008).

process,

$$s_{j+1} = \Omega_j s_{j,t}^{\kappa_1} n_j^{\kappa_2}, \tag{44}$$

where $n_j$ is the percent of an agent's time endowment he spends training. In this formulation, $\kappa_1$ controls the importance of an agent's current human capital on LOD and $\kappa_2$ controls the importance of time training on LOD. In the endogenous models, I calibrate the sequence $\{\Omega_j\}_{j=20}^{j_r-1}$ such that the agent's equilibrium labor or training choices cause $\{s_j\}_{j=20}^{j_r-1}$ under the baseline-fitted U.S. tax code to match the age-specific human capital calibrated in the exogenous model ($\{\varepsilon_j\}_{j=20}^{j_r-1}$).[33]

To calibrate the rest of the LBD parameters, I rely on the estimates in Chang et al. (2002), setting $\Phi_1 = 0.407$ and $\Phi_2 = 0.326$. Following Hansen and İmrohoroğlu (2009), I set $\kappa_1 = 1$ and $\kappa_2 = 0.004$ in the LOD model. Both functional forms imply full depreciation of skills if individuals choose not to work or train at all in the LBD and LOD models, respectively. In the case of the LBD model, full depreciation will never be binding because agents choose to work large quantities in all periods in the exogenous model which does not include the additional human capital incentive for working . In the LOD model, I find that if I include skill accumulation with a function form that is separable in past skills and training time, so as to not imply full depreciation when agents do not train, then the life-cycle profiles are more consistent with formal education as opposed to training.[34] Therefore, I use this nonseparable functional form with the value of $\kappa_1 = 1$ which implies that there is little depreciation of human capital as long as agents use just a small amount of their time endowment for training.[35] The values of $\kappa_2$ and $\{\Omega_j\}_{j=20}^{j_r-1}$ imply that at the start of an agent's career the ratio of time spent training to working is approximately 10 percent and declines steadily until retirement. Through the agent's entire working life, the ratio of the average time spent training to market hours is about 6.25 percent. This average value is in line with the calibration target in Hansen and İmrohoroğlu (2009).[36]

## 4.4 Firm

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

[33] I calibrate these sets of parameters such that they are smooth over the life cycle.

[34] Guvenen et al. (2009) use an alternative LOD accumulation specification that is additively separable in past skills and training. I find that when I use this specification an agent does not accumulate any assets for the first 10-15 years of their working life, and instead tends to save using skill accumulation. In addition, during this time agents work only the necessary hours to finance consumption causing their labor supply profile to be low and flat (see Figure 5 in Guvenen et al. (2009)). Since the shape of these life cycle profiles does not match the data, I choose not to use this functional form.

[35] See Kuruscu (2006) and Heckman et al. (1998) for other examples of quantitative studies that assume little depreciation.

[36] Mulligan (1995) provides empirical estimates of hours spent on employer financed training that are similar to the calibration target.

### 4.5 Government Policies and Tax Functions

Before calibrating the parameters so that the model matches targets in the data, I need to set a baseline tax function that mimics the U.S. tax code. I use the estimates of the U.S. tax code in Gouveia and Strauss (1994) for this tax policy, which I refer to as the baseline-fitted U.S. tax policy. The authors match the U.S. tax code to the data using a three parameter functional form,

$$T(y;\Upsilon_0,\Upsilon_1,\Upsilon_2) = \Upsilon_0(y - (y^{-\Upsilon_1} + \Upsilon_2)^{-\frac{1}{\Upsilon_1}}), \tag{45}$$

where $y$ represents the sum of labor and capital income. The average tax rate is principally controlled by $\Upsilon_0$, and $\Upsilon_1$ governs the progressivity of the tax policy. To ensure that taxes satisfy the budget constraint, $\Upsilon_2$ is left free. Gouveia and Strauss (1994) estimate that $\Upsilon_0 = .258$ and $\Upsilon_1 = .768$ when fitting the data. The authors do not fit separate tax functions for labor and capital income. Accordingly, I use a uniform tax system on the sum of both sources of income for the baseline-fitted U.S. tax policy. I calibrate government consumption, $G$, so that it equals 17 percent of output under the baseline-fitted U.S. tax policy, consistent with CKK. In particular, $\Upsilon_2$ is determined as the value that equates government spending to 17 percent of GDP.

When searching for the optimal tax policy, I restrict my attention to revenue neutral changes that imply that government consumption is equal under the baseline-fitted U.S. tax policy and the optimal tax policy. However, when searching for the optimal tax policy, I allow the tax rates on capital and labor to differ.

The government also runs a balanced-budget social security program. Social security benefits are set so that the replacement rate, $b$, is 50 percent.[37] The payroll tax, $\tau_{ss}$, is determined so that the social security system is balanced each period.

## 5 Computational Experiment

The computational experiment is designed to determine the tax policy that maximizes a given social welfare function holding government revenue constant. I choose a social welfare function (SWF) that corresponds to a Rawlsian veil of ignorance (Rawls (1971)). When searching for the optimal tax policy, I search over both flat and progressive tax policies. However, I determine that the optimal tax policy consists of separate flat taxes on capital and labor income. For notational convenience, I present the computational experiment as

[37]The replacement rate matches the rate in CKK and Conesa and Krueger (2006). The Social Security Administration estimates that the replacement ratio for the median individual is 40 percent (see table VI.F10 in the 2006 Social Security Trustees Report; available at www.ssa.gov/OACT/TR/TR06/). This estimate is lower than the replacement rate I use; however, if one also includes the benefits paid by Medicare, then the observed replacement ratio would be higher.

choosing the optimal flat tax rates on capital and on labor.[38] Since living agents face no earnings uncertainty, the social welfare is equivalent to maximizing the expected lifetime utility of a newborn,

$$SWF(\tau_h,\tau_k) = u(c_j,h_j) + \sum_{s=20}^{J-j-1} \beta^s \prod_{q=20}^{s} (\Psi_q) u(c_{s+1},h_{s+1}), \tag{46}$$

where $\tau_h$ is the flat tax rate on labor income and $\tau_k$ is the flat tax rate on capital income.

# 6 Results

In this section, I quantitatively assess the effects on the optimal capital tax policy of how age-specific human capital is accumulated in my benchmark life cycle model. I determine the optimal tax policies in the exogenous, LBD, and LOD models and then highlight the channels that cause the differences. To fully understand the effects of human capital accumulation, I analyze the aggregate economic variables and life cycle profiles in all three models under the baseline-fitted U.S. tax policy, as well as the changes induced by implementing the optimal tax policies.

## 6.1 Optimal Tax Policies in Exogenous, LBD, and LOD Models

Table 3 describes the optimal tax policies in the three models. The optimal tax policy in the exogenous model is an 18.2 percent flat capital income tax ($\tau_k = 18.2\%$) and a 23.7 percent flat labor income tax ($\tau_h = 23.7\%$). Unlike the analytically tractable model, the optimal capital tax in the computational exogenous model is not zero due to: the inability of the government to borrow; agents being liquidity constrained, the government not being able to tax transfers at a separate rate from ordinary capital income, and exogenous retirement coupled with social security.[39]

Including LBD, I find that the optimal capital tax increases 7.3 percentage points (forty percent) to 25.5% and the optimal labor tax decreases to 22.1%.[40] The alteration in the Frisch labor supply elasticity

[38] This finding is similar to CKK who find that the optimal tax policies are flat in their model without within cohort heterogeneity. However, in contrast, Gervais (2010) finds that the government prefers to use both a tax on capital and a progressive tax on labor income to mimic an age-dependent tax.

[39] I include some of these features that motivate a positive capital tax so that incentives in the model correspond to the incentives in the U.S. economy. For example, the reduced form social security program is necessary so that the level of individual savings are realistic. Other of these features are included to close the model in a tractable manner. See Peterman (2013) for a thorough discussion of the relative strengths of each of these motives in a model similar to the exogenous model.

[40] Although their are large differences between the optimal tax policy in the LBD and exogenous models the welfare losses in the LBD from adopting the optimal tax policy solved for in the exogenous model causes a welfare loss that is equivalent to only 0.1% of lifetime consumption. The small welfare effects from adopting the wrong optimal tax policy are not surprising since Peterman (2015) shows that the welfare losses from adopting sub-optimal levels of the capital and labor tax are much smaller than adopting a sub-optimal level of progressivity.

Table 3: **Optimal Tax Policies in Benchmark Models**

| Tax Rate | Exog | LBD | LOD |
|---|---|---|---|
| $\tau_k$ | 18.2% | 25.5% | 18.9% |
| $\tau_h$ | 23.7% | 22.1% | 23.6% |
| $\frac{\tau_k}{\tau_h}$ | 0.77 | 1.16 | 0.8 |

profile is the principal reason that the optimal capital tax increases in the LBD model. The left panel of Figure 1 plots the lifetime Frisch labor supply elasticities in the LBD model and the exogenous model.[41] The lifetime labor supply elasticity is flat in the exogenous model and upward sloping in the LBD model. Adding LBD causes agents to supply labor relatively more elastically as they age because the human capital benefit decreases. The optimal capital tax is higher in the LBD model in order to implicitly tax agents when they are younger, and supply labor less elastically, at a higher rate.

To confirm that the elasticity channel is responsible for the change in the optimal tax policy in the LBD model, I vary $\sigma_2$ by age in a counterfactual exogenous model (LBD elasticity) so that the shape of the lifetime Frisch labor supply elasticity profile is the same as it is in the LBD model under the optimal tax policy. I find that the optimal tax policy in this counterfactual exogenous model (LBD elasticity), $\tau_k = 25.6\%$ and $\tau_h = 22.1\%$, is almost identical to the optimal tax policy in the LBD model. Thus, the elasticity channel is primarily responsible for the change in the optimal capital tax in the LBD model.

The optimal tax policy in the LOD model is $\tau_k = 18.9\%$ and $\tau_h = 23.6\%$. The optimal capital tax in the LOD model is only 0.7 percentage point larger (approximately five percent) compared to the exogenous model and 6.6 percentage points smaller (approximately twenty five percent) compared to the LBD model. In section 2.3.1, I show analytically that both the elasticity channel and the savings channel affect the optimal capital tax in the LOD model. The right panel of Figure 1 plots the Frisch elasticity profile in the exogenous and LOD models. Compared to the exogenous model, adding LOD to the model causes agents to supply labor relatively more elastically when they are young versus when they are old. The elasticity channel causes a decrease in the optimal capital tax so that agents are implicitly taxed at a lower rate when they are young. Additionally, the inclusion of LOD allows individuals to use training to save, which activates the savings channel.

To quantify the effect of the channels, I solve for the optimal tax policy in an alternative version of the

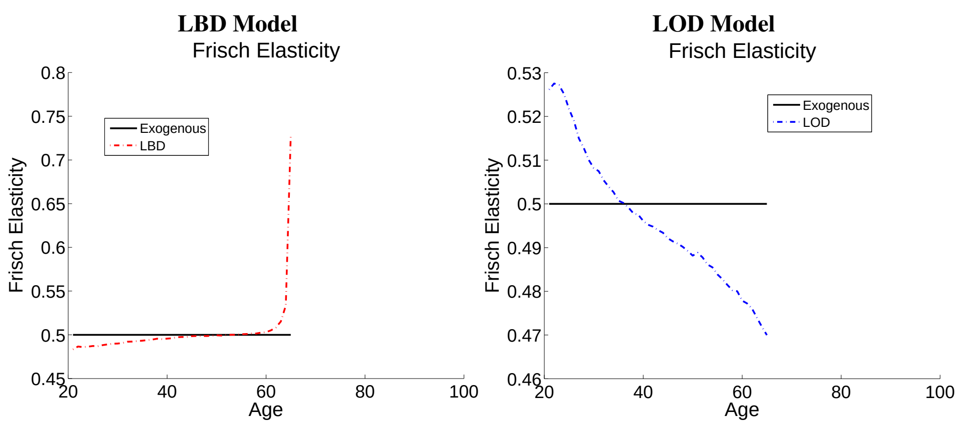
[41] The profiles are determined under the optimal tax policies.

LOD model that uses an alternative utility function,

$$\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi_1 \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} - \chi_2 \frac{(n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}, \tag{47}$$

which is separable in training and hours worked. Using this alternative utility function eliminates the elasticity channel since the Frisch labor supply elasticity in this LOD is no longer a function of time spent training.[42] Eliminating the elasticity channel means that only the savings channel remains in this alternative LOD model (separable utility) with this alternative utility function.[43] The optimal tax policy in this alternative LOD model (separable utility) is $\tau_k = 19.9\%$ and $\tau_h = 23.3\%$. These results indicate that the savings channel causes a 1.7 percentage point increase in the optimal capital tax in order to encourage agents to save via human capital as opposed to physical capital. The total increase in the optimal capital tax is just 0.7 percentage points when both channels are included in the benchmark LOD model and the optimal capital tax increases 1.7 percentage points in the alternative LOD model (separable utility) with just the saving channel. These increases imply that the elasticity channel causes the optimal capital tax to decrease 1 percentage point, canceling just over half of the savings channel's effect.

Figure 1: **Life Cycle Frisch Labor Supply Elasticity in Endogenous Model**




From the ages of 20 to 63, adding LOD has an opposite, but similarly sized, effect on the Frisch elasticity profile as adding LBD. However, the elasticity channel has a much smaller effect on the optimal tax policy in the LOD model. There are two potential reasons that the elasticity channel may have a larger effect in the LBD model. First, the Frisch elasticity increases rapidly over the last two working years in the LBD model making the overall magnitude of the slope much steeper in the LBD model compared to the LOD model.

[42]In particular the Frisch elasticity si $\sigma_2$ since the utility function is separable in all three arguments.

[43]This alternative utility function also eliminates part of the impact of the savings channel so these results are a lower bound on the impact of both the savings and elasticity channel. See the section 2.3.1 for more details.

To confirm the slope over the last few years is not the reason for the elasticity channel having a larger effect on the optimal capital tax in the LBD model versus the LOD model, I determine the optimal tax in another counterfactual exogenous model (smoothed LBD elasticity) which matches the slope of the Frisch elasticity in the LBD model only from ages 20 - 63.[44] Again I find that the optimal tax policy in this counterfactual exogenous model (smoothed LBD elasticity) is almost identical to the optimal tax in the LBD model. Thus the magnitude of the effect of the elasticity channel is not stronger in the LBD model versus the LOD model because the Frisch elasticity in the LBD model increases rapidly over the last few years of the working lifetime.

Instead, the reason that the elasticity channel in the LOD model has a smaller effect on the optimal tax policy is that it causes young agents to be more liquidity constrained. Figure 2 plots the labor supply profile for a young agent. The solid line represents the labor supply profile in the exogenous model under the optimal tax policy. The dashed line represents the labor supply profile, under the same tax policy, but in a counterfactual exogenous model (LOD elasticity) calibrated such that the labor supply elasticity profile matches the LOD model.[45] In all the models, young agents tend to work less compared to middle aged agents since their lower human capital implies their total wage is lower. In this counterfactual exogenous model (LOD elasticity), the labor supply elasticity is higher for younger agents compared to the labor supply elasticity in the unaltered exogenous model. Therefore, in this counterfactual exogenous model (LOD elasticity) agents tend to work relatively less hours when they are young compared to the benchmark model. Because they supply less labor, these agents are more liquidity constrained when they are young in the counterfactual exogenous model (LOD elasticity). The optimal tax policy includes a larger capital tax in order to alleviate these binding liquidity constraints by shifting some of the tax burden to later in an agent's life, when he is no longer liquidity constrained.[46] Thus, the decrease in the optimal capital tax in the LOD model due to the downward sloping elasticity profile is somewhat offset by the desire to increase the capital tax due to liquidity constraints being exacerbated in the LOD model.[47]

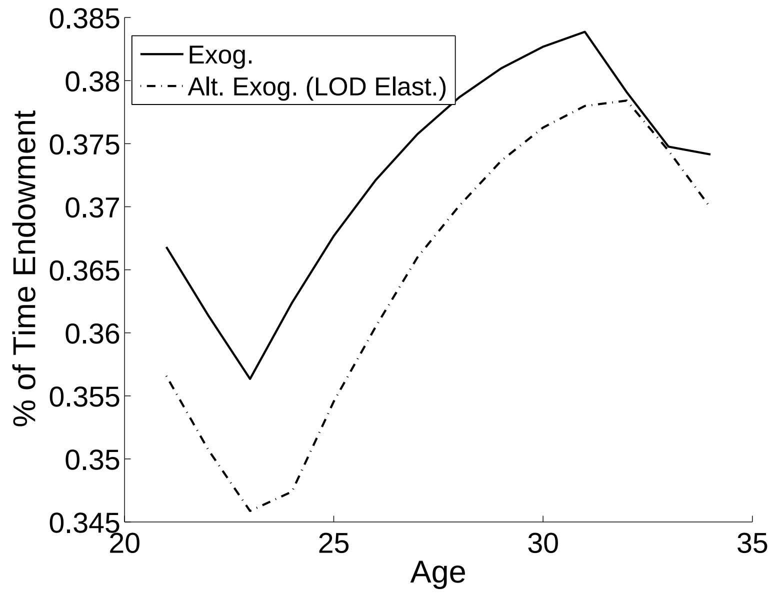
[44]In this counterfactual exogenous model (smoothed LBD elasticity), I set the Frisch elasticity at ages 64 and 65 equal to age 63 in the LBD model.

[45]For the most relevant comparison, I choose to match the labor supply elasticity profile in the LOD model under the optimal tax policy.

[46]Furthermore, the motive to shift the tax burden away from these earlier years when agents are liquidity constrained is enhanced because, in the LOD model, these younger liquidity constrained agents provide labor more elastically which enhances the distortions from binding liquidity constraints.

[47]For a detailed discussion of magnitude of the relationship between liquidity constraints and the optimal capital tax, see Peterman (2013) and CKK.

Figure 2: **Affect of LOD Elasticity on Young Labor Supply**




### 6.2 Comparison of Model to Data

In this section, I examine how well the exogenous model fits compared to the observed data.[48] Figure 3 plots the life cycle profiles from the exogenous model under the baseline-fitted U.S. tax policy and in the actual data.[49] Overall, the model does a decent job matching the data; the general shapes of the profiles are similar. However, there are some discrepancies between the profiles predicted by the model and the life cycle profiles from the data.

The upper left panel compares the average percent of the time endowment that is spent working over the lifetime and the upper right compares the labor income. The actual profiles are constructed from the 1967 - 1999 waves of the Panel Survey of Income Dynamics (PSID). I focus my attention on the earnings for the head of the household between ages 20 and 80.[50]

Starting by focusing on the labor supply profiles, the model generated profiles have a similar hump shape as the profiles from the data. Additionally, both profiles decline rapidly after the age of sixty. Despite the general shapes being similar, there are two main differences. First, in the data, households work 30 percent of their total labor endowment at age 20, which grows rapidly over the next few years until it peaks at around 35 percent of the time endowment. In contrast, in the model, agents work 35 percent of the total

[48]The differences in the life cycle profiles between the three models are not large, thus I choose to only compare the data to the exogenous model (see section 6.3)

[49]Earnings, consumption, and savings from the model are converted to real 2012 dollars by equating the average earnings in the model and the data.

[50]I find that the data for individuals older than 80 were extremely volatile.

labor endowment at age 20. Although the model continues to over predict labor supply, the increase in labor supply over the next few subsequent years is more gradual than in the data. This difference in the labor supply of young households is primarily due to liquidity constraints. In the model agents cannot borrow against future earnings. Thus, agents tend to work more early in their lifetime in the model because wages rise rapidly in the beginning of the life cycle. In contrast, in the data, young households may have a means to borrow, decreasing the relative wealth effect for young households from working. The second major difference between the profiles is that the model generated profile starts to slope downwards just before the age of forty while the profile from the data does not start to slope downwards until households are in their fifties. However, I find in section 7.1 that a more rapidly declining labor supply profile does not materially affect the optimal tax policy.[51]

Focusing on the upper right panel, the earnings profile in the data is similar to the one generated by the model. Both profiles are humped shaped with a peak around forty years old. However, since in the model agents are forced to retire at 65, but in the U.S. economy some households retire after the age of 65, the earnings profile for these older households is higher in the data.

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

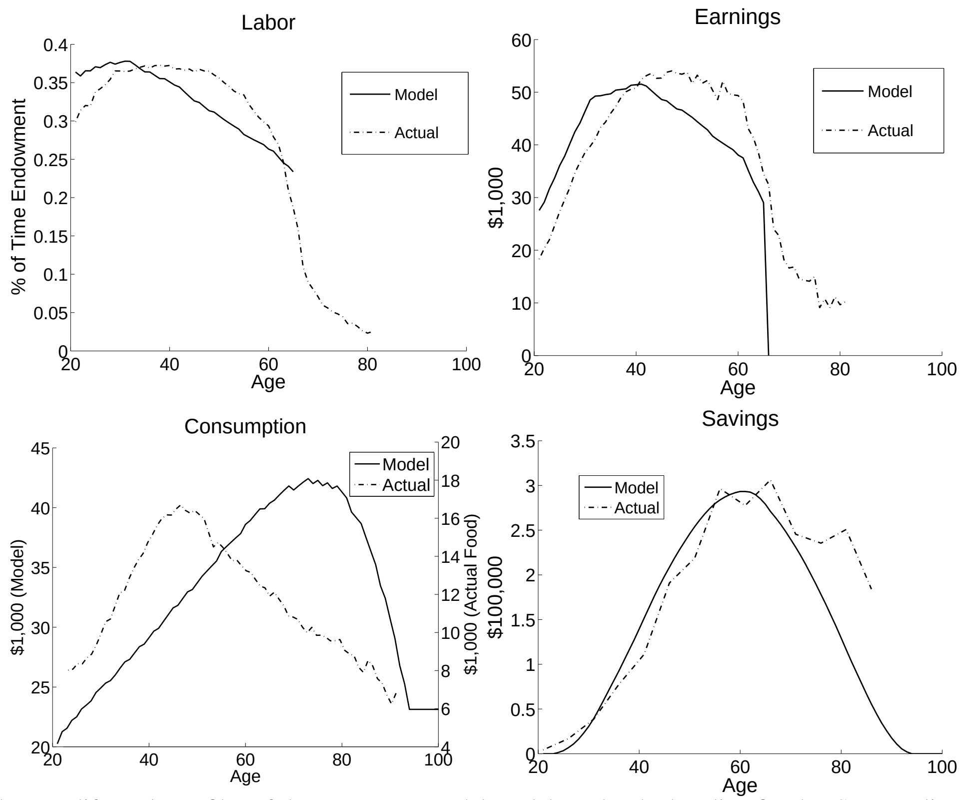
Finally, the lower right panel examines savings in the model and median total wealth in the 2007 Survey of Consumer Finances (SCF) for individuals between the ages of 20 and 80.[52] I smooth through some of the volatility in the wealth data by using five year age bins. Even after smoothing, the data for individuals after age 80 is not included because there are few observations in the sample leading the estimates to be extremely volatile. I find that the profiles are similar in the model and the data. Both are hump-shaped, peaking around \$300,000 at the age of 60. One discrepancy between the two profiles is that the model predicts that agents will deplete their savings more quickly than they do in the data. This discrepancy could arise because the

[51]The lack of relationship between the labor supply profile and optimal tax policy is not surprising. Peterman (2013), Garriga (2001), Erosa and Gervais (2002) all demonstrate, when using a utility function that is homothetic and separable in labor and consumption, such as the one in the benchmark model, that regardless of the labor supply profile the government does not want to condition taxes on age. However, if the utility function is not homothetic and separable then the government wants to condition labor income taxes on age and in the absence of the ability to use age-dependent taxes a downward sloping labor supply will lead to a positive optimal tax on capital.

[52]In order to prevent the upper tail of the wealth distribution from skewing the statistic for comparison, I choose to focus on the median level of wealth as opposed to the average.

model does not include any motive for individuals leaving a bequest for younger generations. Overall, I find that the model does fair job matching the data.

Figure 3: **Actual and Exogenous Life Cycle Profiles**




**Note:** These plots are life cycle profiles of the exogenous model models under the baseline-fitted U.S. tax policy and the actual profiles in the data. The units of the consumption, earnings, and capital profiles are converted to real dollars by matching the average earnings in the model and in the data.

## 6.3 The Effects of Adding Endogenous Age-Specific Human Capital

This section analyzes the effect on the aggregate economic variables and life cycle profiles of changing from exogenous human capital accumulation to either LBD or LOD under the baseline-fitted U.S. tax policy. Figure 4 plots the life cycle profiles of hours, consumption, assets, and age-specific human capital in all three models. Table 4 describes the optimal tax policies and summarizes the aggregate economic variables under both the baseline-fitted U.S. tax policy and optimal tax policies. The first, fourth, and seventh columns are the aggregate economic variables under the baseline-fitted U.S. tax policy in the exogenous, LBD, and LOD models, respectively. The second, fifth, and eighth columns are the aggregate economic variables under the optimal tax policies. The third, sixth, and ninth columns are the percentage changes in the aggregate economic variables induced from adopting the optimal tax policies.

Table 4: **Aggregate Economic Variables**

| Aggregate | Exogenous Baseline | Exogenous Optimal | Exogenous *% Change from Baseline to Optimal* | LBD Baseline | LBD Optimal | LBD *% Change from Baseline to Optimal* | LOD Baseline | LOD Optimal | LOD *% Change from Baseline to Optimal* |
|---|---|---|---|---|---|---|---|---|---|
| Y | 0.81 | 0.82 | *1.8%* | 0.80 | 0.81 | *1.0%* | 0.76 | 0.77 | *1.6%* |
| K | 2.17 | 2.25 | *3.6%* | 2.17 | 2.17 | *0.2%* | 2.06 | 2.12 | *3.0%* |
| N | 0.46 | 0.46 | *0.8%* | 0.46 | 0.46 | *1.5%* | 0.44 | 0.44 | *0.8%* |
| Avg Hours | 0.33 | 0.34 | *0.7%* | 0.33 | 0.34 | *0.8%* | 0.33 | 0.34 | *0.7%* |
| w | 1.12 | 1.13 | *1.0%* | 1.12 | 1.12 | *-0.5%* | 1.12 | 1.13 | *0.8%* |
| r | 0.05 | 0.05 | *-4.6%* | 0.05 | 0.05 | *2.2%* | 0.05 | 0.05 | *-3.6%* |
| tr | 0.03 | 0.03 | *4.2%* | 0.02 | 0.02 | *2.8%* | 0.03 | 0.03 | *3.8%* |
| Value | -139.26 | -138.46 | *0.6%* | -159.01 | -158.10 | *0.6%* | -155.14 | -154.36 | *0.5%* |
| CEV | | | *0.7%* | | | *0.9%* | | | *0.6%* |
| **Average Tax Rate** | **Baseline** | **Optimal** | | **Baseline** | **Optimal** | | **Baseline** | **Optimal** | |
| Capital | 15.5% | 18.2% | | 15.6% | 25.5% | | 15.3% | 18.9% | |
| Labor | 23.7% | 23.7% | | 23.7% | 22.1% | | 23.7% | 23.6% | |
| Ratio | 0.65 | 0.77 | | 0.66 | 1.16 | | 0.65 | 0.80 | |
| **Marginal Tax Rate** | **Baseline** | **Optimal** | | **Baseline** | **Optimal** | | **Baseline** | **Optimal** | |
| Capital | 19.4% | 18.2% | | 19.6% | 25.5% | | 19.1% | 18.9% | |
| Labor | 25.5% | 23.7% | | 25.5% | 22.1% | | 25.5% | 23.6% | |
| Ratio | 0.76 | 0.77 | | 0.77 | 1.16 | | 0.75 | 0.80 | |

**Note:** The average hours refers to the average percent of time endowment worked in the productive labor sector. Both the marginal and average tax rates vary with income under the baseline-fitted U.S. tax policy. The marginal tax rates are the population weighted average marginal tax rates for each agent.

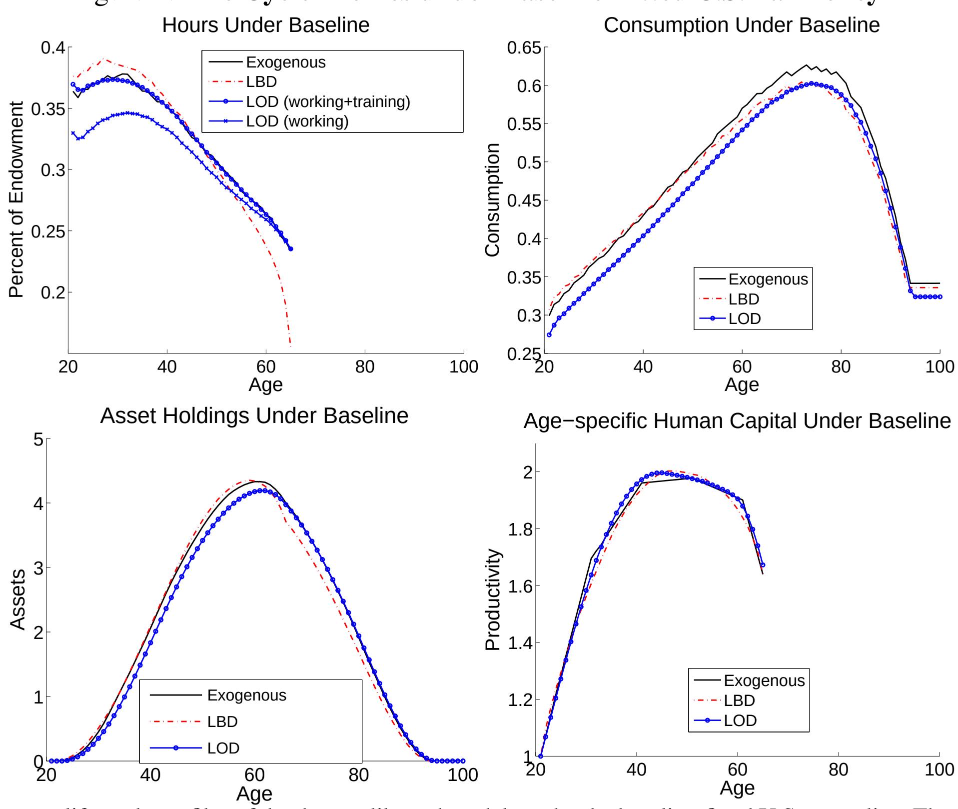
Starting by comparing the exogenous and LBD models, the first and fourth columns of table 4 demonstrate that the aggregate level of hours, labor, and capital are similar in the two models. The calibrated parameters are determined so that under the baseline-fitted U.S. tax policy the models match certain targets from the data. Since many of the aggregate economic variables are targets and these calibration parameters are determined separately in each model, the aggregates are similar in the exogenous and LBD models.

Although adding LBD does not have a large effect on the aggregate economic variables, it does cause changes in the life cycle profiles. Adding LBD causes agents to work relatively more at the beginning of their working life when the human capital benefit is larger, and less later when the benefit is smaller (see the solid black and dashed red lines in the upper-left panel of Figure 4). The upper-right panel shows that the lifetime consumption profile is steeper in the exogenous model compared to the LBD model. The intertemporal Euler equation controls the slope of consumption profile over an agent's lifetime. The relationship is

$$\left(\frac{c_{j+1}}{c_j}\right)^{\sigma_1} = \Psi_j \beta \widetilde{r}_t, \tag{48}$$

where $\widetilde{r}_t$ is the marginal after-tax return on capital. In order to induce the same capital to output ratio in the LBD model, $\beta$ is lower which is the primary cause of the flatter consumption profile. The lower value of

Figure 4: **Life Cycle Profiles under Baseline-Fitted U.S. Tax Policy**




**Note:** These plots are life cycle profiles of the three calibrated models under the baseline-fitted U.S. tax policy. There are two labor lines for the LOD model, one solely for hours worked and the other for hours worked plus hours spent training.

$\beta$ in the LBD model decreases the value an agent places on their consumption in future periods so agents' savings are also relatively smaller for the second half of their lifetime (see the lower-left panel). The lifetime age-specific human capital profiles are similar in the two models since the sequence of parameters $\{\Omega j\}_{j=20}^{j_r-1}$ is calibrated so that age-specific human capital matches (see the lower-right panel of Figure 4).

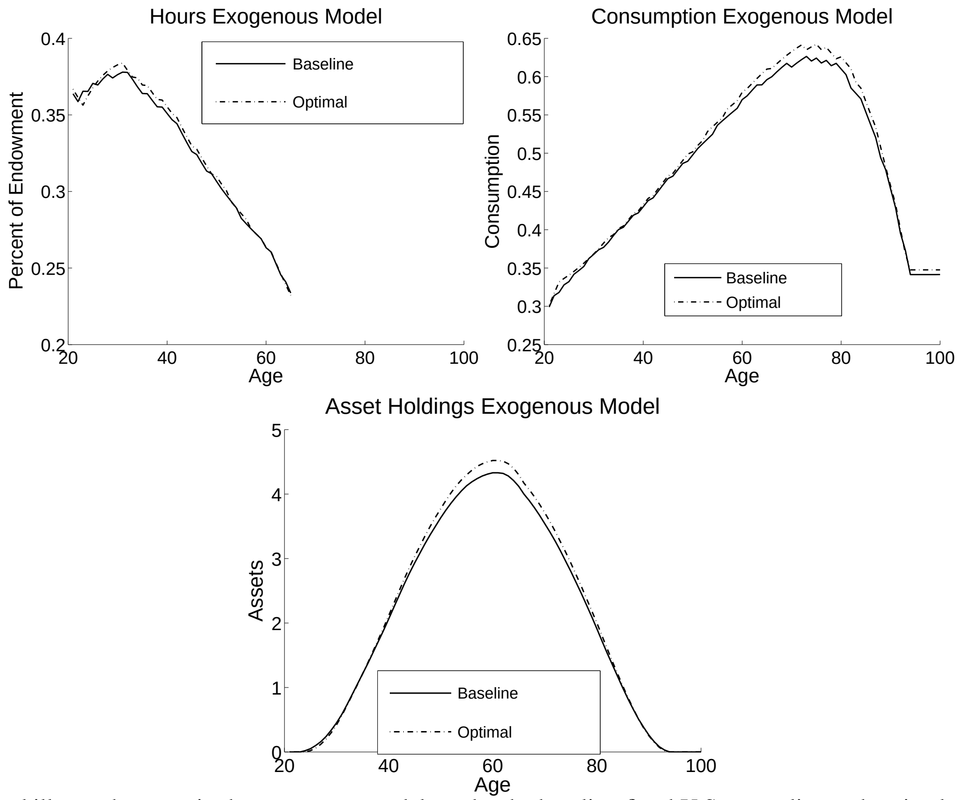
Next, comparing the exogenous and LOD models, although the parameters values are calibrated such that the targets match, the size of the economy is smaller in the LOD model because agents must spend part of their time endowment training. Comparing the first and seventh columns of table 4, aggregate labor supply, and physical capital are smaller in the LOD model compared with the exogenous model, however, the relative ratios of the aggregates are similar.

Adding LOD also affects the life cycle profiles. Figure 4 plots two labor supply profiles for the LOD model — the first is solely hours spent working, and the second is the sum of hours spent working and training (see the blue lines in the upper-left panel). The LOD labor supply profile, including training, is similar to the labor supply profile in the exogenous model; however the profile that excludes training is smaller. The difference between the two LOD profiles is the amount of time spent training. This gap shrinks as an agent ages, representing a decrease in the amount of time spent training. Agents spend less time training as they age because the benefits decrease since they have fewer periods to take advantage of their human capital. Adding LOD causes the size of the economy to decrease, causing a shift down in the life cycle profile for consumption. In the LOD model, agents can use their time endowment to accumulate human capital, which acts as an alternative form of savings from assets. Therefore, during their working lives, agents hold less ordinary capital and opt to use human capital to supplement their savings. As an agent approaches retirement the value of the human capital decreases and the ordinary savings profile in the LOD model converges to the profile in the exogenous model. Finally, similar to LBD, the lifetime age-specific human capital profiles are similar in the exogenous and LOD models since the profiles are a calibration target.

### 6.4 The Effects of the Optimal Tax Policy in the Exogenous Model

This section examines the effects on the economy of adopting the optimal tax policy in the exogenous model. In the exogenous model, the optimal capital tax is smaller than the average marginal tax under the baseline-fitted U.S. tax policy, so adopting the optimal tax policy causes an increase in aggregate capital (see columns one and two of table 4). The average marginal labor tax is also less under the optimal tax policy than the baseline so the labor supply increases.[53] The increase in labor supply is relatively smaller than the increase in capital so the rental rate on capital decreases and the wage rate increases.

Figure 5: **Life Cycle Profiles in the Exogenous Model**




**Note:** Since the skills are the same in the exogenous models under the baseline-fitted U.S. tax policy and optimal tax policy, they are not plotted.

Figure 5 plots the life cycle profiles for time worked, consumption, and assets in the exogenous model under the baseline-fitted U.S. tax policies and the optimal tax policies. The solid lines are the profiles under the baseline-fitted U.S. tax policies, and the dashed lines are the profiles under the optimal tax policies. Although the changes in the profiles from adopting the optimal tax policies are small, I still interpret them in order to provide the reader with a better understanding of the dynamics in the model. Adopting the optimal tax policy in the exogenous model causes changes in all three life cycle profiles: (i) early in their life, agents work relatively more; (ii) agents save more, especially during periods when they are wealthier; and (iii) the lifetime consumption profile steepens. The first change, agents working more early in their life, is a result of the lower implicit tax on young labor income due to a decrease in the tax rate on capital income.

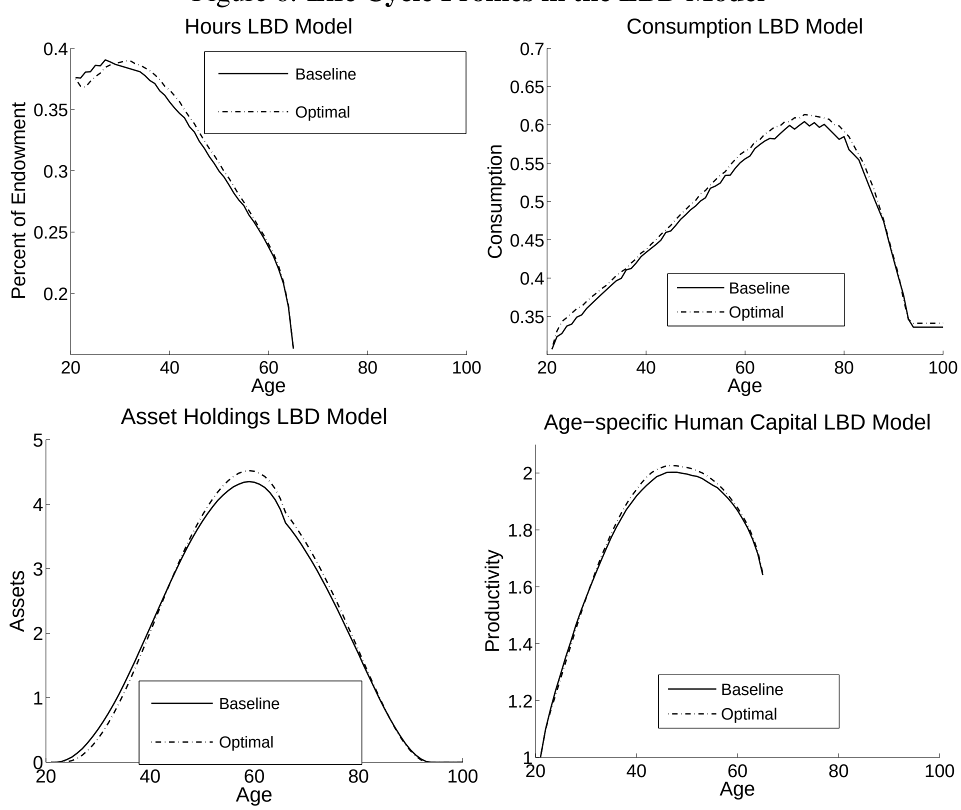
Implementing the optimal tax policy causes a decrease in both the capital tax and the rental rate on

[53] A revenue neutral tax change can include a decrease in both the average marginal tax rate on labor and capital since the baseline is progressive and the optimal is flat. Additionally, agents generally work longer under the optimal tax policy so the tax base is larger.

capital leading to shifts in both the capital and savings profiles. These declines have competing effects on the marginal after-tax return on capital. Furthermore, the change in the tax rate has an uneven effect on agent's net return over his lifetime since the baseline-fitted US tax on capital is progressive and the optimal tax is flat. The decrease in the tax rates is larger for agents who hold more savings since their marginal tax rate was relatively higher under the progressive baseline-fitted US tax policy. Overall, the change in the tax rate dominates for these middle-aged agents and the after tax return increases. The converse is true for younger agents who experience a decrease in the after tax return when the optimal tax policy is adopted. In response to these changes, middle-aged individuals increase their savings under the optimal tax policy, while younger and older agents decrease their savings (see the lower left panel of Figure 5). In addition, since the after tax return to capital controls the slope of the consumption profile, adopting the optimal tax policy causes a steeper consumption profile for middle-aged agents (Figure 5, upper-right panel).

## 6.5 The Effects of Optimal Tax Policy in the LBD Model

Figure 6: **Life Cycle Profiles in the LBD Model**




Adopting the optimal tax policy in the LBD model causes an increase in the capital tax and a decrease in the labor tax (see column four, five, and six of table 4). The changes in the tax policy cause a small increase in the capital stock and a large increase in aggregate labor supply in the LBD model. The relatively larger rise in labor translates into a decrease in the wage rate and an increase in the rental rate on capital.

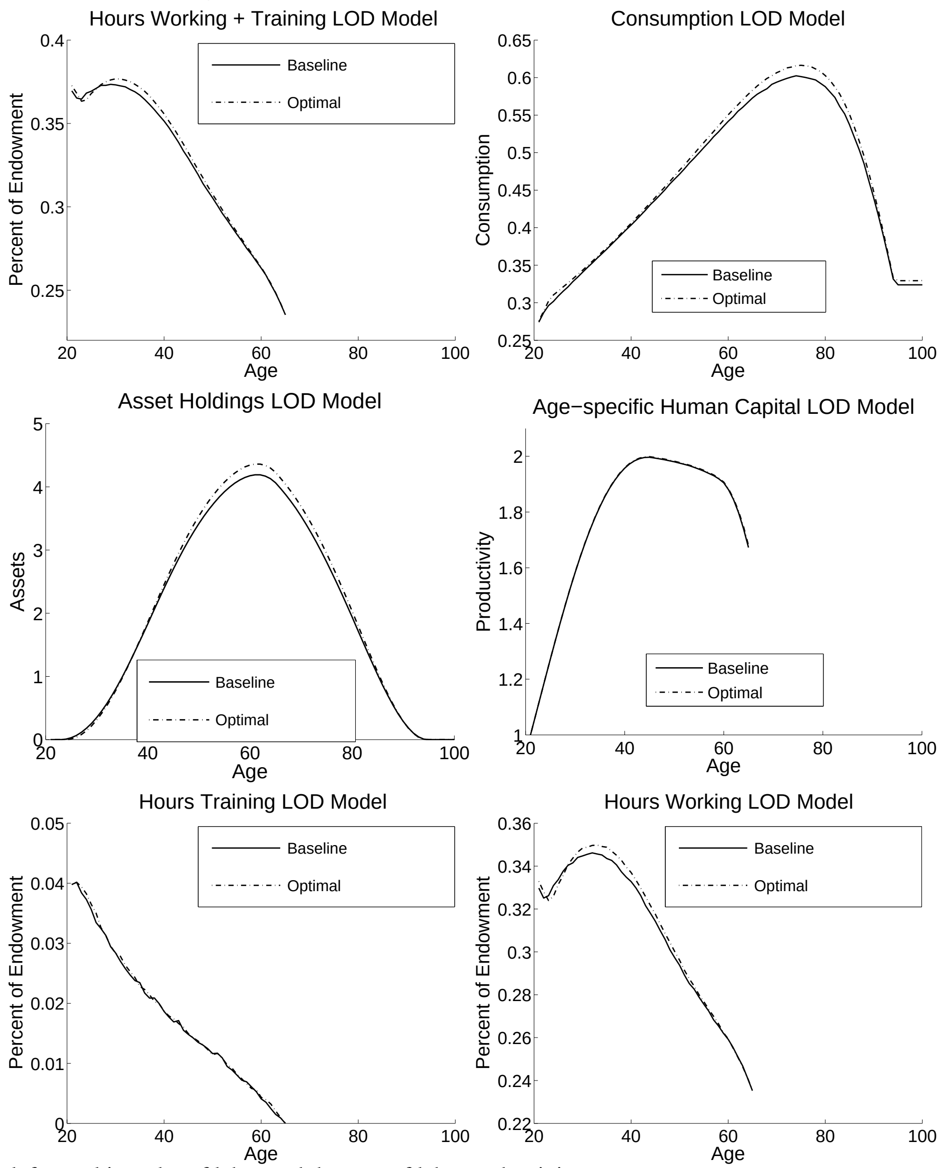
Implementing the optimal tax policies in the LBD model causes the life cycle profiles to change somewhat differently than in the exogenous model (see Figure 6). Agents shift time worked from earlier to later years in response to the larger capital tax, which implicitly taxes labor income from early years at a higher rate (upper-right panel of Figure 6). Because agents work more in their middle years, age-specific human capital is also higher for middle aged agents (see the lower-right panel). Applying the optimal tax policy introduces two opposing effects on the agent's lifetime asset profile. First, agents increase their savings under the optimal tax policy because the economy is larger. Second, the larger capital tax under the optimal tax policy decreases the average marginal after-tax return on capital, causing agents to hold fewer assets. The first effect is constant for all agents. The second effect is not constant for all agents, but it is negatively proportional to an agent's capital income because the baseline-fitted U.S. tax policy is progressive and the optimal tax policy is flat. This second effect dominates for younger and older agents and they tend to save less under the optimal tax policy. Conversely, the first effect dominates for middle-aged agents and they tend to save more. I find that the first effect has the dominate impact on consumption leading the consumption profile to uniformly shift upwards (see the upper-right panel).[54]

### 6.6 The Effects of Optimal Tax Policy in the LOD Model

Although the optimal capital tax is larger in the LOD model than in the exogenous model, the direction of the changes in the tax rates from adopting the optimal tax policy are similar in the two models: a decrease in the average marginal tax on capital and labor. Therefore, the aggregate economic variables respond to adopting the optimal tax policy in a similar fashion in both models: capital increases, labor increases, wages increase, and the rental rate decreases. Adopting the optimal tax policy in the LOD also induces changes in the life cycle profiles much like those in the exogenous model (see Figures 5 and 7): (i) agents work more earlier in their life, (ii) agents increase their savings during the middle of their lifetime, and (iii) agents increase their consumption at a faster rate throughout their life. One additional feature of the LOD model is that agents choose how much to train. I find that adopting the optimal tax policy has minimal effects on the training profile (see the lower-left panel of Figure 7).

[54] Although adopting the optimal tax policy does not cause a uniform change in the after-tax return to capital in the LBD model, liquidity constraints cancel out their effect on the slope of the consumption profile.

Figure 7: **Life Cycle Profiles in the LOD Model**




**Note:** The upper-left panel is a plot of labor and the sum of labor and training.

# 7 Sensitivity Analysis

This section examines the sensitivity of two different aspects of the model. First, I demonstrate that the general shape of the labor supply profile does not affect the optimal tax policy in the exogenous model. Second, I determine that using a different utility function does not weaken the relationship between how human capital is accumulated and the optimal capital tax.

## 7.1 The Effect of Shape of Labor Supply Profile on Optimal Tax Policy

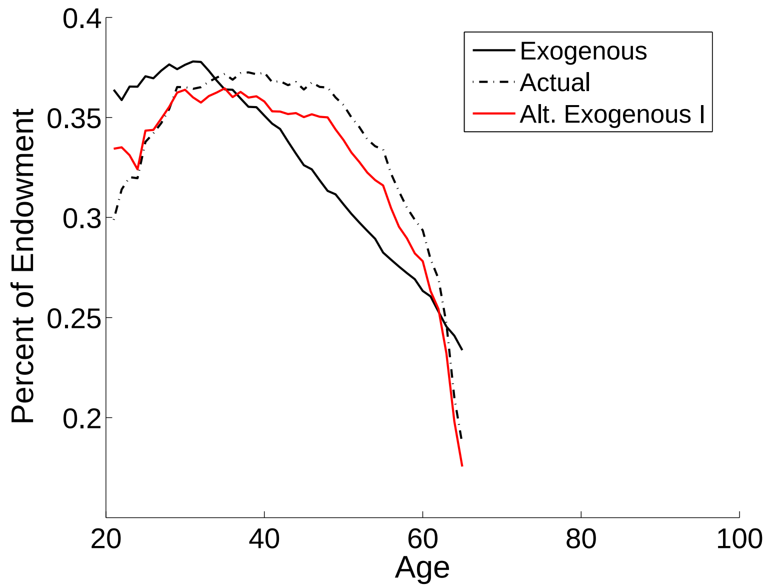
In this section, I test the relationship between the shape of the labor supply profile and the optimal tax. I examine this relationship because there are differences between the profile in the data and the models. Moreover, there are differences in the labor supply profiles between the three models. For example, comparing the labor supply profile in the actual data and the exogenous model (Figure 3), the exogenous model predicts that the labor supply profile will be downward sloping over a majority of the lifetime while the actual profile from the data tends to be much flatter. Moreover, comparing the labor supply profile predicted by the exogenous and LBD models (Figure 4), the labor supply profile in the LBD model declines much more rapidly over the second half of the working lifetime than it does in the exogenous model. In order to test whether the shape of the labor supply profile affects the optimal tax policy, I find the optimal tax policies in two alternative exogenous models in which I vary the values of $\chi$ over the lifetime such that the labor supply profile matches either the data or the profile predicted by the LBD model.

First, I determine whether the flatter labor supply profile in the actual data has an affect on the optimal tax policy. Figure 8 plots the labor supply profile generated in the exogenous benchmark model (solid black line) and the average hours worked in the data (dashed black line). Additionally, the solid red line plots the labor supply generated in an alternative exogenous model (labor supply match data) which is calibrated to more closely match the data.[55] I find that the optimal tax policy in this alternative exogenous model (labor supply match data), $\tau_h = 23.8\%$ $\tau_k = 17.9\%$, is almost identical to the optimal tax policy in the benchmark exogenous model ($\tau_h = 23.7\%$ and $\tau_k = 18.2\%$). This result indicates that the steeper labor supply profile predicted by the model has only a negligible affect on the optimal tax policy.

Next, I examine whether the more rapid decline in the labor supply profile over the end of the working lifetime in the LBD model affects the optimal tax policy. In this experiment, I calibrate an alternative exogenous model (match LBD labor) such that the labor supply profile more closely matches the profile in the LBD model. Figure 9 plots the labor supply profiles in the benchmark exogenous model (solid black),

[55]The labor supply profiles are all under the baseline-fitted U.S. tax policy.

Figure 8: **Flatter Labor Supply Profile**

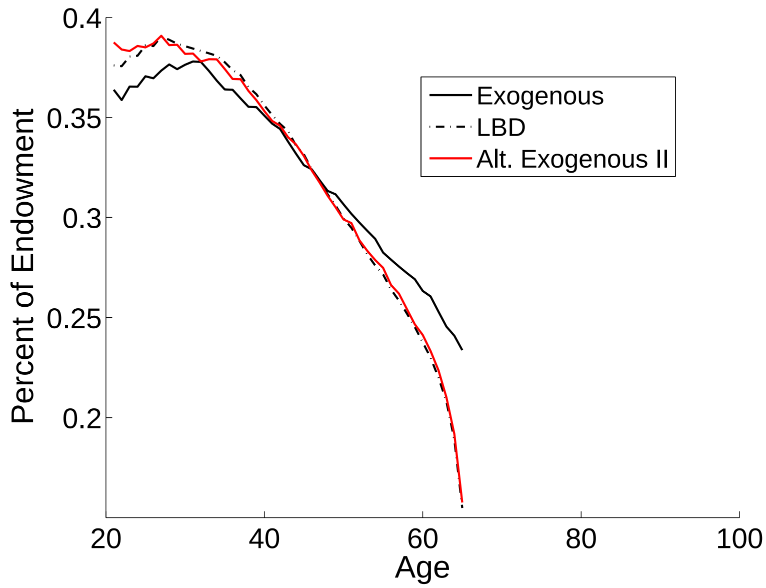



the LBD model (dashed black), and the new alternative exogenous model (match LBD labor) (solid red). Comparing the dashed black line and the red line, the labor supply profile over the second half of the working lifetime in the alternative exogenous model (match LBD labor) matches the rapid decline predicted in the LBD model. I find that the optimal tax policy in this alternative exogenous model (match LBD labor) is $\tau_h = 23.6\%$ and $\tau_k = 18.9\%$. Again, the optimal tax policy in this altered exogenous model (match LBD labor) is almost identical to the optimal tax policy in the benchmark exogenous model.

These results indicate that the optimal tax policy in my benchmark exogenous model is not related to the general shape of the labor supply profile.[56] These results are not surprising since the benchmark utility function is homothetic and separable in labor and consumption. Therefore, labor supply is not related to the Frisch labor supply elasticity. This utility function eliminates the most active channel by which the labor supply profile affects the optimal tax policy. Some previous works, such as Peterman (2013) and CKK, use a utility function in which labor supply affects the Frisch labor supply elasticity. The next section examines whether the relationship between endogenous human capital accumulation and optimal taxation changes when this type of utility function is used.

[56]One exception to this result is described in section 7.1, where I demonstrate that a lower labor supply in the first few years of working leads agents to have binding liquidity constraints for more years and can alter the optimal tax policy.

Figure 9: **Labor Supply in Alternative Exogenous**




## 7.2 Utility Function

In this section, I determine the effect of human capital accumulation on the optimal capital tax in a model with an alternative utility function, $U(c_{1,t}, 1-h_{1,t}) = \frac{(c_{1,t}^{\gamma}(1-h_{1,t})^{1-\gamma})^{1-\varsigma}}{1-\varsigma}$. This utility function is the benchmark specification in CKK. I refer to this utility function as the nonseparable utility function. This function includes additional motives for a positive capital tax since it is no longer both separable and homothetic in each consumption and labor.

### 7.2.1 Calibration

The nonseparable utility function requires calibrating two new parameters. The new parameters are $\gamma$, which determines the comparative importance of consumption and leisure, and $\varsigma$, which controls risk aversion. It is no longer possible to separately target the Frisch elasticity and average time worked since $\gamma$ controls both of these values. Therefore, I calibrate $\gamma$ to match the percentage of the time endowment worked and no longer use the Frisch elasticity as a target.

Table 5 lists the calibration parameters for the nonseparable utility parameters. The Frisch elasticity in the exogenous model for this utility function is $\frac{(1-h)}{h}\frac{1-\gamma(\varsigma-1)}{\varsigma}$. The average Frisch elasticity implied by the calibration in the exogenous model is 1.13, which is more than twice as large as with the benchmark utility specification in the exogenous model.

Table 5: **Preference Parameters**

| Parameter | Exog | LBD | LOD | Target |
|---|---|---|---|---|
| $\beta$ | 1.012 | 1.009 | 1.013 | $K/Y = 2.7$ |
| $\Psi_j\beta$ | 0.998 | 0.996 | 1.000 | $K/Y = 2.7$ |
| $\gamma$ | 0.35 | 0.27 | 0.34 | Avg. $h_j + n_j = \frac{1}{3}$ |
| $\varsigma$ | 4 | 4 | 4 | CKK |

### 7.2.2 Optimal Tax Policies in Nonseparable Models

There is a larger motive for a positive capital tax in all the models with nonseparable preferences for two reasons. First, the nonseparable utility implies that the Frisch elasticity profile is negatively related to the labor supply profile. Since the labor supply profile is downward sloping over a majority of the life, the Frisch elasticity profile is upward sloping in all the models. The upward sloping Frisch elasticity profile motivates a large positive capital tax.[57] Second, there are less degrees of freedom when calibrating the nonseparable model so the Frisch elasticity is larger in the nonseparable model. Therefore, the government would prefer to rely on a capital tax, as opposed to a labor income tax.

Table 6 lists the optimal tax policies for the nonseparable models. Even with the nonseparable utility function — which motivates a large capital tax on its own — there is still a large range of optimal capital tax rates depending on how human capital is accumulated. Compared to the exogenous model, adding LBD causes the optimal capital tax to increase by 14.5 percentage points (approximately forty five percent). Moreover, adding LOD causes a 4.7 percentage point (approximately a fifteen percent) increase in the optimal capital tax compared to the exogenous model and a 9.8 percentage point decrease compared to the LBD model (approximately thirty percent). The range of the optimal capital taxes is even larger in this model indicating that the importance of how human capital is accumulated on optimal capital taxation is robust to this change in the utility specification.

Table 6: **Optimal Tax Policies in Nonseparable Models**

| Tax Rate | Exog | LBD | LOD |
|---|---|---|---|
| $\tau_k$ | 31.8% | 46.3% | 36.5% |
| $\tau_h$ | 20.2% | 15.0% | 18.7% |
| $\frac{\tau_k}{\tau_h}$ | 1.57 | 3.09 | 1.95 |

[57] I find that adding LBD causes the Frisch elasticity to be even steeper and further enhances this motive for a positive tax on capital. In contrast, I find that the Frisch elasticity is still upward sloping when I add LOD, however it is less steep.

# 8 Conclusion

In this paper, I characterize the optimal capital and labor tax rates in three separate life cycle models in which age-specific human capital is accumulated exogenously, endogenously through LBD, and endogenously through LOD. Analytically, I demonstrate that compared to the exogenous model, including either form of endogenous human capital accumulation creates a motive for the government to condition labor income taxes on age and in their absence, it will use a non-zero capital tax to mimic these age-dependent taxes. Quantitatively, I find large variation in the optimal capital tax depending on whether human capital is accumulated endogenously or exogenously. Moreover, I find that the form of endogenous human capital accumulation matters, the optimal tax rate is between 6.6 and 9.8 percentage points larger with LBD compared to LOD depending on the utility function. These findings demonstrate that the form by which human capital is assumed to accumulate has large impacts on the optimal capital tax.

LBD increases the motive for a capital tax since it alters the lifetime labor supply elasticity profile. Adding LBD to the model causes younger agents to supply labor relatively less elastically since the human capital benefit decreases over an agent's lifetime. A larger capital tax is optimal because it implicitly taxes younger labor supply income, which is supplied less elastically, at a higher rate. Adding LOD to the model has two counteracting affects on the optimal tax policy. Including LOD causes younger agents to supply labor relatively more elastically because training is an imperfect substitute for working. This change in the elasticity motivates the government to decrease the capital tax and raise the labor tax. However, a tax on labor in the LOD model decreases the incentive for agents to save with human capital as opposed to physical capital. Therefore, the government has an incentive to increase the tax on capital in order to promote more training. Overall, I find that this second effect dominates and adding LOD also causes the optimal capital tax to increase in numerical simulations.

In a standard life cycle model, I find a large bound on the estimates of the optimal capital tax depending on the model's assumptions with regard to how human capital is accumulated. Moreover, the way in which human capital is accumulated effects the shape of the lifetime Frisch labor supply elasticity. For economists to reach more precise conclusions from life cycle models, they must determine the process by which agents acquire age-specific human capital once they start working. Determining the shape of the labor supply elasticity profile could provide helpful guidance as to which form of human capital accumulation is consistent with the data.

# A Analytical Derivations

**For Online Publication**

## A.1 Primal Approach

I use the primal approach to determine the optimal tax policy.[58] I use a social welfare function that maximizes the expected utility of a newborn and discounts future generations with social discount factor $\theta$ (see section 5 for more details),

$$[U(c_{2,0},h_{2,0})/\theta]+\sum_{t=0}^{\infty}\theta^{t}[U(c_{1,t},h_{1,t})+\beta U(c_{2,t+1},h_{2,t+1})]. \tag{49}$$

The government maximizes this objective function with respect to two constraints: the implementability constraint and the resource constraint.[59] The implementability constraint is the agent's intertemporal budget constraint, with prices and taxes replaced by his first order conditions (equations 5, 6, and 7)

$$c_{1,t}U_{c1}(t)+\beta c_{2,t+1}U_{c2}(t+1)+h_{1,t}U_{h1}(t)+\beta h_{2,t+1}U_{h2}(t+1)=0. \tag{50}$$

Including this constraint ensures that any allocation the government chooses can be supported by a competitive equilibrium. The resource constraint is

$$c_{1,t}+c_{2,t}+K_{t+1}-K_{t}+G_{t}=rK_{t}+w(h_{1,t}+h_{2,t}\varepsilon_{2}). \tag{51}$$

Including the benchmark utility specification, the Lagrangian the government maximizes is

$$\begin{aligned}\mathscr{L}=&\frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1}-\chi\frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}+\beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1}-\chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}\\&-\rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}\varepsilon_2))\\&-\rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}\varepsilon_2))\\&+\lambda_t(c_{1,t}^{1-\sigma_1}+\beta c_{2,t+1}^{1-\sigma_1}-\chi h_{1,t}^{1+\frac{1}{\sigma_2}}-\beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}})\end{aligned} \tag{52}$$

where $\rho$ is the Lagrange multiplier on the resource constraint and $\lambda$ is the Lagrange multiplier on the implementability constraint.

[58]See Lucas and Stokey (1983) or Erosa and Gervais (2002) for a full description of the primal approach.

[59]The government budget constraint is a third constraint. Due to Walras' Law, I only need to include two of three constraints in the Lagrangian and leave out the government budget constraint.

## A.2 Exogenous

The Lagrangian for this specification is

$$\begin{aligned}\mathscr{L} =&\frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\ &-\rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}\varepsilon_2)) \\ &-\rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}\varepsilon_2)) \\ &+\lambda_t(c_{1,t}^{1-\sigma_1}+\beta c_{2,t+1}^{1-\sigma_1}-\chi h_{1,t}^{1+\frac{1}{\sigma_2}}-\beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}})\end{aligned} \tag{53}$$

where $\rho$ is the Lagrange multiplier on the resource constraint and $\lambda$ is the Lagrange multiplier on the implementability constraint. The first order conditions with respect to labor, capital and consumption are

$$w\rho_t = \chi h_{1,t}^{\frac{1}{\sigma_2}}(1+\lambda_t(1+\frac{1}{\sigma_2})) \tag{54}$$

$$w\rho_{t+1}\theta\varepsilon_2 = \beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}(1+\lambda_t(1+\frac{1}{\sigma_2})) \tag{55}$$

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

$$\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \tag{57}$$

and

$$\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1}. \tag{58}$$

Combining the first order equations for the governments problem with respect to capital and consumption yields

$$\Big(\frac{c_{2,t+1}}{c_{1,t}}\Big)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta} \tag{59}$$

where $\rho$ is the Lagrange multiplier on the resource constraint and $\lambda$ is the Lagrange multiplier on the implementability constraint. Taking the ratio of the agent's first order conditions, equations 5 and 6 under the benchmark utility specification gives

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1}{\varepsilon_2}\Big(\frac{c_{1,t}}{c_{2,t+1}}\Big)^{-\sigma_1}\Big(\frac{h_{2,t+1}}{h_{1,t}}\Big)^{\frac{1}{\sigma_2}}. \tag{60}$$

Combining equation 59 and 60 yields

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1}{\varepsilon_2}\Big(\frac{\beta\rho_t}{\rho_{t+1}\theta}\Big)\Big(\frac{h_{2,t+1}}{h_{1,t}}\Big)^{\frac{1}{\sigma_2}}. \tag{61}$$

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

$$\frac{\rho_t\beta}{\varepsilon_2\rho_{t+1}\theta}\Big(\frac{h_{2,t+1}}{h_{1,t}}\Big)^{\frac{1}{\sigma_2}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})}. \tag{62}$$

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

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})} = 1. \tag{63}$$

## A.3 LBD

The Lagrangian for this LBD specification is modified from the exogenous model. In particular,human capital benefit alters the implementability constraint. Suppressing the arguments of the skills function, the implementability constraint in the LBD model is

$$c_{1,t}U_{c1}(t) + \beta c_{2,t+1}U_{c2}(t+1) + h_{1,t}U_{h1}(t) - \frac{\beta h_{1,t}U_{h2}(t+1)h_2 s_{h1}(t+1)}{s_2} + \beta h_{2,t+1}U_{h2}(t+1) = 0, \tag{64}$$

where $s_{h1}(t+1)$ represents the partial derivative of the skill function for an older agent with respect to hours worked when young. Thus the Lagrangian for the LBD model is,

$$\begin{aligned}\mathscr{L} =& \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\ & - \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}s_{2,t})) \\ & - \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}s_{2,t+1})) \\ & + \lambda_t(c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi h_{1,t}^{1+\frac{1}{\sigma_2}} + \frac{\chi\beta h_{2,t+1}^{1+\frac{1}{\sigma_2}} h_{1,t}s_{h1}(t+1)}{s_2} - \beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}})\end{aligned} \tag{65}$$

where $\rho$ is the Lagrange multiplier on the resource constraint and $\lambda$ is the Lagrange multiplier on the implementability constraint. The first order conditions with respect to labor, capital and consumption are

$$\begin{aligned} w\rho_t =& \chi h_{1,t}^{\frac{1}{\sigma_2}}(1+\lambda_t(1+\frac{1}{\sigma_2})) - \theta\rho_{t+1}h_{2,t+1}s_{h1}(t+1) \\ & + \lambda_t\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}\beta h_{1,t}\left[\frac{s_{h1}(t+1)^2}{s_2^2} - \frac{s_{h1,h1}(t+1)}{s_2}\right]\end{aligned} \tag{66}$$

$$w\rho_{t+1}\theta s_2 = \beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}\left[1+\lambda_t(1+\frac{1}{\sigma_2}) + (1+\frac{1}{\sigma_2})\frac{h_{1,t}s_{h1}(t+1)\lambda_t}{s_2}\right] \tag{67}$$

$$\rho_t = \theta(1+r)\rho_{t+1} \tag{68}$$

$$\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \tag{69}$$

and

$$\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1}. \tag{70}$$

The first order conditions with respect to capital and consumption are the same in the exogenous (56, 57, and 58) and LBD models (68, 69, and 70). Therefore equation 10 still holds for this model and therefore the optimal tax on capital is still zero when the government can condition labor income taxes on age.

Combining the first order equations for the governments problem with respect to capital and consumption yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta} \tag{71}$$

Taking the ratio of the agent's first order conditions, equations 15 and 16 and combining with equation 71 yields

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}} = \left(\frac{h_{1,t}}{h_{2,t+1}}\right)^{\frac{1}{\sigma_2}} \left(\frac{\rho_{t+1}\theta s_2}{\beta\rho_t}\right) - \frac{h_{2,t+1}s_{h1}(t+1)}{1+r(1-\tau_k)}. \tag{72}$$

Combining equations 72, 66 and 67 the ratio of the optimal taxes on labor is,

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}} = \frac{\left(1+\lambda_t(1+\frac{1}{\sigma_2})-\lambda_t(1+\frac{1}{\sigma_2})\frac{h_{1,t}s_{h1}(t+1)}{s_2}\right)\left(1+\frac{h_{2,t+1}s_2}{1+r(1-\tau_k)}\right)}{1+\lambda_t(1+\frac{1}{\sigma_2})+h_{2,t+1}^{1+\frac{1}{\sigma_2}}h_{1,t}^{1+\frac{-1}{\sigma_2}}\frac{\lambda_t}{s_2}\left(\frac{s_{h1}(t+1)}{s_2}-s_{h1,h1}(t+1)\right)} - \frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)}. \tag{73}$$

## A.4 LOD

Since agents have the additional choice variable $n_1$ in the LOD model, the have an additional first order condition with respect to this variable (equation 25). This new first order condition requires an additional constraint in the government's Lagrange that ensures that the allocation the government chooses properly equates an individual's disutility of training when young and working when old (see equations 23 and 25). This constraint simplifies to $U_{n1}(t)s_2 = \beta U_{h2}(t+1)h_{2,t+1}s_n(t+1)$. I use $\eta_t$ as the Lagrange multiplier on this new constraint.

The Lagrangian for the LOD model is

$$\begin{aligned}
\mathscr{L} =& \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{(h_{1,t}+n_{1,t})^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\
& - \rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}s_2)) \\
& - \rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}s_2)) \\
& + \lambda_t(c_{1,t}^{1-\sigma_1}+\beta c_{2,t+1}^{1-\sigma_1}-\chi h_{1,t}^{1+\frac{1}{\sigma_2}}-\beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}) \\
& + \eta_t(\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}s_{n1}(t+1)-\chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}}s_2)
\end{aligned} \tag{74}$$

where $\rho$ is the Lagrange multiplier on the resource constraint, $\lambda$ is the Lagrange multiplier on the implementability constraint and $\eta$ is the Lagrange multiplier on the constraint equating the first order conditions with respect to training and work. The first order conditions with respect to labor, capital, consumption and training are,

$$w\rho_t = \chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}}\left[1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}\right] \tag{75}$$

$$w\rho_{t+1}\theta s_2 = \beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}\left[1+\lambda_2\left(1+\frac{1}{\sigma_2}\right)-\eta_t\left(1+\frac{1}{\sigma_2}\right)s_{n1}(t+1)\right] \tag{76}$$

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

$$\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \tag{78}$$

$$\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1} \tag{79}$$

and

$$\theta\rho_{t+1}h_{2,t+1}s_{n2}(t+1) = \tag{80}$$

$$\frac{\chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}}\Big(\lambda_t h_{1,t}+\eta_t s_2+\sigma_2(h_{1,t}+n_{1,t})(1+\eta_t s_{n2}(t+1))\Big)-\beta\chi\eta_t\sigma_2 h_{2,t+1}^{1+\frac{1}{\sigma_2}}(h_{1,t}+n_{1,t})s_{n2,n2}(t+1)}{\sigma_2(h_{1,t}+n_{1,t})} \tag{81}$$

The first order conditions with respect to capital and consumption are the same in the exogenous (56, 57, and 58) and LOD models (77, 78, and 79). Therefore equation 10 still holds for this model and therefore the optimal tax on capital is still zero when the government can condition labor income taxes on age.

Combining the first order equations for the governments problem with respect to capital and consumption yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta} \tag{82}$$

Taking the ratio of the agent's first order conditions, equations 22 and 23 and combining with equation 82 yields

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \left(\frac{h_{2,t+1}}{h_{1,t}+n_{1,t}}\right)^{\frac{1}{\sigma_2}}\left(\frac{\beta\rho_t}{\rho_{t+1}\theta s_2}\right). \tag{83}$$

Taking the ratio of equations 75 and 76 yields,

$$\left(\frac{h_{2,t+1}}{h_{1,t}+n_{1,t}}\right)^{\frac{1}{\sigma_2}}\left(\frac{\beta\rho_t}{\rho_{t+1}\theta s_2}\right) = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)}. \tag{84}$$

Combining equations 83 and 84 generates the following expression for the ratio of the optimal labor taxes,

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)}. \tag{85}$$

# B Competitive Equilibrium

**For Online Publication**

## B.1 LBD Model

Given a social security replacement rate $b$, a sequence of skill accumulations parameters $\{\Omega_j\}_{j=20}^{j_r-1}$, government expenditures $G$, and a sequence of population shares $\{\mu_j\}_{j=20}^{J}$, a stationary competitive equilibrium in the LBD model is a sequence of agent allocations, $\{c_j, a_{j+1}, h_j\}_{j=20}^{J}$, a production plan for the firm $(N,K)$, a government labor tax function $T^l : \mathbb{R}_+ \to \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \to \mathbb{R}_+$, a social security tax rate $\tau_{ss}$, a age-specific human capital accumulation function $S : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, social security benefits $SS$, prices $(w,r)$, and transfers $Tr$ such that:

1. Given prices, policies, transfers, and benefits, the agent maximizes equation 29 subject to

$$c_j + a_{j+1} = ws_jh_j - \tau_{ss}ws_jh_j, +(1+r)(a_j+Tr) - T^l[ws_jh_j(1-.5\tau_{ss})] - T^k[r(a_j+Tr)], \tag{86}$$

$$s_{j+1} = S_{\text{LBD}}(\Omega_j, s_j, h_j), \tag{87}$$

   for $j < j_r$, and

$$c_j + a_{j+1} = SS + (1+r)(a_j+Tr) - T^k[r(a_j+Tr)], \tag{88}$$

   for $j \geq j_r$.
   Additionally,

$$c \geq 0, 0 \leq h \leq 1, a_j \geq 0, a_{20} = 0. \tag{89}$$

2. Prices $w$ and $r$ satisfy

$$r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \tag{90}$$

$$w = (1-\alpha)\left(\frac{K}{N}\right)^{\alpha} \tag{91}$$

3. The social security policies satisfy

$$SS = b\frac{wN}{\sum_{j=20}^{j_r-1}\mu_j} \tag{92}$$

$$\tau_{ss} = \frac{ss\sum_{j=j_r}^{J}\mu_j}{w\sum_{j=20}^{j_r-1}\mu_j} \tag{93}$$

4. Transfers are given by

$$Tr = \sum_{j=20}^{J}\mu_j(1-\Psi_j)a_{j+1} \tag{94}$$

5. Government budget balance:

$$G = \sum_{j=20}^{J}\mu_jT^k[r(a_j+Tr)] + \sum_{j=20}^{j_r-1}\mu_jT^l[ws_jh_j(1-.5\tau_{ss})] \tag{95}$$

6. Market clearing:

$$K = \sum_{j=20}^{J} \mu_j a_j \tag{96}$$

$$N = \sum_{j=20}^{J} \mu_j s_j h_j \tag{97}$$

$$\sum_{j=20}^{J} \mu_j c_j + \sum_{j=20}^{J} \mu_j a_{j+1} + G = K^{\alpha} N^{1-\alpha} + (1-\delta)K \tag{98}$$

### B.1.1 LOD Model

Given a social security replacement rate $b$, a sequence of skill accumulations parameters $\{\Omega_j\}_{j=20}^{j_r-1}$, government expenditures $G$, and a sequence of population shares $\{\mu_j\}_{j=20}^{J}$, a stationary competitive equilibrium in the LBD model is a sequence of agent allocations, $\{c_j, a_{j+1}, h_j\}_{j=20}^{J}$, a production plan for the firm $(N,K)$, a government labor tax function $T^l : \mathbb{R}_+ \to \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \to \mathbb{R}_+$, a social security tax rate $\tau_{ss}$, a age-specific human capital accumulation function $S : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, social security benefits $SS$, prices $(w,r)$, and transfers $Tr$ such that:

1. Given prices, policies, transfers, and benefits, the agent maximizes equation 29 subject to

$$c_j + a_{j+1} = ws_jh_j - \tau_{ss}ws_jh_j, +(1+r)(a_j+Tr) - T^l[ws_jh_j(1-.5\tau_{ss})] - T^k[r(a_j+Tr)], \tag{99}$$

$$s_{j+1} = S_{\text{LOD}}(\Omega_j, n_j, h_j), \tag{100}$$

for $j < j_r$, and

$$c_j + a_{j+1} = SS + (1+r)(a_j+Tr) - T^k[r(a_j+Tr)], \tag{101}$$

for $j \geq j_r$.
Additionally,

$$c \geq 0, 0 \leq h \leq 1, a_j \geq 0, a_{20} = 0. \tag{102}$$

2. Prices $w$ and $r$ satisfy

$$r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \tag{103}$$

$$w = (1-\alpha)\left(\frac{K}{N}\right)^{\alpha} \tag{104}$$

3. The social security policies satisfy

$$SS = b\frac{wN}{\sum_{j=20}^{j_r-1}\mu_j} \tag{105}$$

$$\tau_{ss} = \frac{ss\sum_{j=j_r}^{J}\mu_j}{w\sum_{j=20}^{j_r-1}\mu_j} \tag{106}$$

4. Transfers are given by

$$Tr = \sum_{j=20}^{J} \mu_j(1-\Psi_j)a_{j+1} \tag{107}$$

5. Government budget balance:

$$G = \sum_{j=20}^{J} \mu_j T^k [r(a_j + Tr)] + \sum_{j=20}^{j_r - 1} \mu_j T^l [w s_j h_j (1 - .5\tau_{ss})] \tag{108}$$

6. Market clearing:

$$K = \sum_{j=20}^{J} \mu_j a_j \tag{109}$$

$$N = \sum_{j=20}^{J} \mu_j s_j h_j \tag{110}$$

$$\sum_{j=20}^{J} \mu_j c_j + \sum_{j=20}^{J} \mu_j a_{j+1} + G = K^{\alpha} N^{1-\alpha} + (1-\delta)K \tag{111}$$

## References

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

**Armenter, Roc and Stefania Albanesi**, "Understanding Capital Taxation in Ramsey Models," Working Paper September 2009.

**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.

**Atkinson, A.B. and J.E. Stiglitz**, "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 1976, *6*, 55–75.

**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.

**Best, Micahel and Henrick Kleven**, "Optimal Income Taxation with Career Effects of Work Effort," Working Paper April 2012.

**Cespedes, Nikita and Michael Kuklik**, "Optimal Taxation and Life Cycle Labor Supply Profile," Working Paper December 2012.

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

**Chang, Yongsun, Joao Gomes, and Frank Schorfheide**, "Learning-By-Doing as a Propagation Mechanism," *American Economic Review*, December 2002, *92* (5), 1498–1520.

**Chen, Been-Lon, Hung-Ju Chen, and Ping Wang**, "Taxing Capital is Not a Bad Idea Indeed: The Role of Human Capital and Labor-Market Frictions," Working Paper May 2010.

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

**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.

**Corlett, W. J. and D. C. Hague**, "Complementarity and the Excess Burden of Taxation," *The Review of Economic Studies*, 1953, *21* (1), pp. 21–30.

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

**da Costa, Carlos E and Marcelo R Santos**, "Age-dependent Taxes with Endogenous Human Capital Formation," Technical Report, Rede de Economia Aplicada 2015.

**Deaton, Angus**, "Optimally uniform commodity taxes," *Economics Letters*, 1979, *2* (4), 357 – 361.

**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.

**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**, "On the Optimality of Age-Dependent Taxes and the Progressive U.S. Tax System," 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," NBER Working Papers 15526, National Bureau of Economic Research, Inc 2009.

**Hansen, Gary D. and Selahattin İmrohoroğlu**, "Business Cycle Fluctuations and the Life Cycle: How Important Is On-the-Job Skill Accumulation?," *Journal of Economic Theory*, 2009, *144* (6), 2293 – 2309. Dynamic General Equilibrium.

**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.

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

**Huggett, Mark, Gustavo Ventura, and Amir Yaron**, "Sources of Lifetime Inequality," Working Paper 13224, NBER 2007.

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

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

**Jones, Larry E. and Rodolfo E. Manuelli**, "On the Taxation of Human Capital," Unpublished Manuscript, Northwestern University and University of Wisconsin November 1999.

**_ , _ , 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.

**Judd, Kenneth L.**, "Optimal Taxation and Spending In General Competitive Growth Models," *Journal of Public Economics*, 1999, *71* (1), 1 – 26.

**Kapicka, Marek**, "Optimal Income Taxation with Human Capital Accumulation and Limited Record Keeping," *Review of Economic Dynamics*, 2006, *9* (4), 612 – 639.

_ , "The Dynamics of Optimal Taxation when Human Capital is Endogenous," Working Paper August 2009.

**Karabarbounis, Marios**, "Heterogeneity in Labor Supply Elasticity and Optimal Taxation," Working Paper, Federal Reserve Bank of Richmond 2012.

**Kuruscu, Burhanettin**, "Training and Lifetime Income," *The American Economic Review*, 2006, *96* (3), 832–846.

**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.

**Mulligan, Casey**, "The Intertemporal Substitution of Work–What Does the Evidence Say?," University of Chicago - Population Research Center, Chicago - Population Research Center June 1995.

**Peterman, Willam**, "An Extensive Look at Taxes: How does Endogenous Retirement Affect Optimal Taxation?," Working Paper 2012-28, Federal Reserve Board of Governors April 2012.

_ , "The Effect of Endogenous Human Capital Accumulation on Optimal Taxation," Working Paper, Federal Reserve Board of Governors October 2015.

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

_ , "Reconciling Micro and Macro Estimates of the Frisch Labor Supply Elasticity," *Economic Inquiry*, 2016, *54* (1), 100–120.

**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.

**Reis, Catarina**, "Optimal Taxation with Unobservable Investment in Human Capital," Mimeo, MIT 2007.