Taxing Capital? The Importance of How Human Capital is Accumulated

William B Peterman*

June 21, 2016

Abstract

This paper explores both analytically and quantitatively how the optimal capital tax changes when human capital is accumulated in three different cases: (i) exogenously, (ii) endogenously through learning-*by*-doing (LBD), (iii) and endogenously through learning-*or*-doing (LOD). While Peterman (2016a) demonstrates that the optimal tax on capital increases when human capital is accumulated through LBD as opposed to exogenously, it is still unknown whether this result is robust to other forms of endogenous human capital accumulation. Analytically this paper demonstrates in a simple two period model that including either LBD or LOD create novel motives for the government to use age-dependent labor income taxes. If the government cannot condition taxes on age, then a capital tax can be optimal in order to mimic age-dependent taxes. However, the channel by which these two forms of endogenous human capital affect the optimal capital tax are quite different. LBD causes a change in the optimal tax because young agents supply labor relatively less elastically. In contrast, LOD changes the optimal tax because it causes young agents to supply labor relatively more elastically and provides the ability for agents to save both via training and ordinary capital. Turning to a quantitative model, these different channels lead to numerically different effects on the optimal capital tax from including either LBD or LOD. Including human capital accumulation with LOD, as opposed to exogenously, causes the optimal capital tax to increase by between 0.7 and 3.7 percentage points. In contrast, the optimal tax increase between 7.3 and 14.5 percentage points when human capital is accumulated with LBD, as opposed to exogenously. Taken as a whole, these findings indicate that not only whether human capital is accumulated exogenously or endogenously, but also the specific form of endogenous human capital accumulation can have important implications for 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. Given this potential interaction between human capital accumulation and optimal capital tax Peterman (2016a) examines how incorporating a specific form of endogenous human capital accumulation can affect the optimal capital tax and finds it can cause considerable effects. However, it is still an open question of how the effect on optimal capital tax varies between different forms of endogenous human capital accumulations. Thus, this paper both analytically and quantitatively assesses the effect of various different human capital accumulation processes on the optimal capital tax. Overall, these findings indicate that the optimal capital tax is affected not only by whether human capital is accumulated exogenously or endogenously, but also by the specific form of endogenous human capital accumulation.

Specifically, this paper determines 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] In contrast to exogenous human capital accumulation, 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 these different human capital assumptions is important.[3]

In order to demonstrate the channels by which human capital accumulation affects the optimal capital

[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).

tax, this paper analytically explores a simple overlapping generations model (OLG) model where the utility function is both separable and homothetic with respect to each consumption and hours worked. This model is ideal for isolating the additional channels due to LBD or LOD since 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.[5] 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. This channel, the elasticity channel, is transmitted through changes in the labor supply elasticity profile and is consistent with the elasticity channel described in Peterman (2016a).

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. In contrast to the elasticity channel in the LBD model, the elasticity channel in the LOD model causes the optimal capital tax to be lower in order to decrease the implicit taxes on labor income when agents are younger. The second channel, the savings channel, is unique to the LOD model. This channel arises because training is an alternative method of saving, as opposed to accumulating physical capital. When the government taxes

[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]Peterman (2016a) demonstrates that LBD can affect the optimal capital tax in a similar analytical model. However, the model includes other channels that motivate a non-zero tax on capital because it uses a utility function that is not homothetic in labor.

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 capital tax. Thus, adding either LOD or LBD could lead to a significant difference in the optimal capital tax compared to with exogenous human capital accumulation.

Next, I quantitatively assess the effect of the form by which age-specific human capital is accumulated on the optimal capital tax in a rigorous calibrated life cycle model. 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. Similar to Peterman (2016a), when LBD is included, I find that the optimal tax on capital increases between 7.3 and 14.5 percentage points, depending on the utility specification. 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. Moreover, the much larger increase with LBD versus LOD indicates that not only whether human capital is accumulated endogenously but the specific form by which it is accumulated can affect the optimal capital tax. 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.[6] 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 different 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,

[6] 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. Instead, CKK find that the within cohort heterogeneity from idiosyncratic productivity shocks affects the optimal progressivity of the labor tax. Therefore, I exclude this feature since this study focuses on the effects of human capital accumulation on the optimal capital tax and abstracts from the effect of endogenous human capital accumulation on the optimal progressivity.

but concludes that the restriction on the government from being able to tax accidental bequests at a different rate from ordinary capital income is also a large contribution to the positive optimal capital tax in these models.[7] 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 incorporating different 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.[8] 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.[9] However, this second strand of literature is unable to account for the effects of endogenous human capital accumulation through life cycle channels which CKK and Peterman (2013) demonstrate are quantitatively important for motivating a positive capital tax. 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.

Two papers that combine these two strands and examine the effects on optimal taxation of endogenous human capital accumulation in a life cycle model are Peterman (2016a) and da Costa and Santos (2015).

[7]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.

[8]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.

[9]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.

Peterman (2016a) finds that adding LBD causes the optimal tax on capital to increase. Unlike this paper, Peterman (2016a) does not consider LOD so no conclusion can be drawn on how the specific form of endogenous human capital may change its affect on the optimal capital tax. da Costa and Santos (2015) consider the effects of both LBD and LOD on optimal taxation. However, instead of focusing on how the optimal capital tax changes in order to mimic an age-dependent tax with different forms of human capital accumulation, da Costa and Santos (2015) focus on the effects when taxes can be conditioned on age. Since CKK and Peterman (2013) demonstrate that a major reason to tax capital is to mimic an age-dependent tax it is important to examine the effects without age-dependent taxation.[10]

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 and discusses the channels affecting the optimal tax. 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 the intuition for why adding different forms of endogenous human capital accumulation can overturn the result from Garriga (2001) that with a utility function that is separable and homothetic in each consumption and labor the government does not want to condition labor income taxes on age.[11] It is useful to determine if the government wants to use age-dependent taxes because 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 (see Garriga (2001) and Erosa and Gervais (2002)).

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 mortality risk. 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

[10] Both Peterman (2016a) and da Costa and Santos (2015) 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 changes with different forms of human capital is accumulation.

[11] A similar set of results for the exogenous and LBD model are in Peterman (2016a).

in the computational model.

2.1 General Set-up

I nest the analytically tractable exogenous, LBD, and LOD models in a general set up in which agents live with certainty for two periods, and their preferences over consumption and labor are represented by

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

where β is the discount rate, $c_{j,t}$ is the consumption of an age j agent at time t, $h_{j,t}$ is the percent of the time endowment the agent works, and $n_{1,t}$ is the percent of the time endowment that a young agent spends training in the LOD model.[12] The agent maximizes equation 1 with respect to consumption, hours, and training in the LOD model subject to the following constraints,

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

and

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

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, τ_k is the tax rate on capital income, w is the efficiency wage for labor services, and r is the rental rate on capital. Age-specific human capital is normalized to unity when the agent is young. At age two, age-specific human capital is $s(\cdot)_2$. In the exogenous model s_2 is exogenously determined. In the LBD model s_2 is a positive and concave function of the hours worked when young, $s(\cdot)_2=s_2(h_{1,t})$. In the LOD model s_2 is a positive and concave function of the hours spent training when an agent is young, $s(\cdot)_2=s_2(n_{1,t}$. I assume that the tax rate on labor income can be conditioned on age; however, the tax rate on capital income cannot. I combine equations 2 and 3 to form a joint intertemporal budget constraint,

$$c_{1,t}+\frac{c_{2,t+1}}{1+r(1-\tau_k)}=w(1-\tau_{h,1})h_{1,t}+\frac{w(1-\tau_{h,2})s_2(\cdot)h_{2,t+1}}{1+r(1-\tau_k)}. \tag{4}$$

The agent's first order conditions are,

$$\frac{U_{h1}(t)}{U_{c1}(t)}=-[w(1-\tau_{h,1})+\beta\frac{U_{c2}(t+1)}{U_{c1}(t)}w(1-\tau_{h,2})h_{2,t+1}s_{h1}(t+1)], \tag{5}$$

[12] 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.

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -ws_2(\cdot)(1-\tau_{h,2}), \tag{6}$$

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

and

$$\frac{U_{n1}(t)}{U_{c2}(t+1)} = -\beta w(1-\tau_{h,2})s_{n1}(n_{1,t})h_{2,t+1}. \tag{8}$$

where $U_{c1}(t) \equiv \frac{\partial U(c_{1,t},h_{1,t})}{\partial c_{1,t}}$. I highlight the pieces of the first order conditions that are unique to the LBD model in red and the pieces unique to the LOD model in blue. In particular, the first order condition with respect to h_1 is different with LBD (equations 5) because working has the additional human capital benefit.[13] Moreover, since agents have the additional choice variable n_1 in the LOD model, there is an additional first order condition (equation 8) which does not exist in the other two models. Given a social welfare function, prices, and taxes, these first order conditions, combined with the intertemporal budget constraint, determine the optimal allocation of $(c_{1,t}, h_{1,t}, n_{1,t}, c_{2,t+1}, h_{2,t+1})$.

2.2 Optimal Tax Policy

Next, I solve for the optimal tax policy, with a benchmark utility function that is homothetic with respect to each consumption and hours worked, $U(c,h,n) = \frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h+n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$. I focus on this utility function since Garriga (2001) demonstrates that it is not optimal to condition taxes on age with this utility function and exogenous human capital accumulation. I solve for the optimal tax policy using the primal approach in which I solve for the optimal allocation and use them to back out the optimal tax function (see Appendix A.1 for details of the approach).

Generally, when thinking about the optimal capital tax it useful to solve for the optimal age-dependent labor income taxes. The intertemporal Euler equation demonstrates the potential link between age-dependent labor taxes and a capital tax,

$$s_2(\cdot)\frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1+r(1-\tau_k))\frac{1-\tau_{h,1}}{1-\tau_{h,2}} + \beta h_{2,t+1}s_{h1}(t+1). \tag{9}$$

In particular, if the government wants to create a wedge on the marginal rate of substitution by varying the labor income tax rate by age, then τ_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

[13] In particular, in both the exogenous and LOD models $s_{h1}(t+1) = 0$ so this term cancels out.

income. Thus, throughout the analytical analysis, it will be important to determine whether age-dependent labor income taxes are optimal.

In my model, I find that the optimal allocations imply the following ratios for the optimal labor taxes in the exogenous, LBD, and LOD models, respectively, (see Appendix A for the formulation of the problem),

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})} = 1, \tag{10}$$

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}} = \frac{\left(1+\lambda_t(1+\frac{1}{\sigma_2})-\lambda_t(1+\frac{1}{\sigma_2})\frac{h_{1,t}s_{h1}(t+1)}{s_2}\right)\left(1+\frac{h_{2,t+1}s_2}{1+r(1-\tau_k)}\right)}{1+\lambda_t(1+\frac{1}{\sigma_2})+h_{2,t+1}^{1+\frac{1}{\sigma_2}}h_{1,t}^{1+\frac{-1}{\sigma_2}}\frac{\lambda_t}{s_2}\left(\frac{s_{h1}(t+1)}{s_2}-s_{h1,h1}(t+1)\right)} - \frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)} < 1, \tag{11}$$

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)} \neq 1. \tag{12}$$

Equation 10 demonstrates that in the exogenous model the government has no incentive to condition labor taxes on age.[14] Equation 11 demonstrate that adding LBD creates an incentive to tax labor income at a relatively higher when an agent is young.[15] Although equation 12 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. The results in both of the endogenous models contrast with the exogenous model, in which the government has not incentive to condition labor income taxes on age. 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. Thus, adding either form of endogenous human capital accumulation will create a motive for the government to tax capital if age-dependent taxes are allowed.

2.2.1 Effect of Learning-by-Doing

As Peterman (2016a) demonstrates, examining the Frisch elasticities in the exogenous and LBD models, provides the intuition why adding LBD increases the optimal relative tax on young labor income or tax on capital.[16] In the exogenous model, the Frisch elasticity simplifies to $\Xi_{\text{exog}} = \sigma_2$. The Frisch elasticity in

[14] λ is the Lagrange multiplier on the implementability constraint. See Appendix A.1 for more details. This result is specific to this utility function.

[15] In particular, the relative tax on young labor is higher than the tax on old labor income as long as λ is positive.

[16] Although Peterman (2016a) provides the intuition for why adding LBD alters the optimal tax policy, I summarize the effects in this paper to provide a complete set of results which help distinguish the different effects of LOD. Moreover, Peterman (2016a)

the LBD model is, $\Xi_{\text{LBD}} = \frac{\sigma_2}{1-\frac{h_{i+1,t+1}w_{t+1}(h_{i,t}\sigma_2 s_{hi,hi}(t+1)-s_{hi}(t+1))}{s_{i,t}(1+r_t(1-\tau_k))w_t}}$. [17] As opposed to the constant Frisch elasticity in the exogenous model, adding LBD causes the Frisch elasticity to vary with hours worked. In particular, the extra terms in Ξ_{LBD} increase the size of the denominator, thus holding hours and consumption constant between the two models, $\Xi_{\text{exog}} > \Xi_{\text{LBD}}$.

The change in the Frisch elasticity is due to the inclusion of the human capital benefit in the LBD model which is absent in the exogenous model. 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 relative importance of the human capital benefit decreases as an agent ages because he has fewer periods to use his human capital. Thus, LBD causes an agent to supply labor relatively less elastically when they are young than when they are old creating an incentive for the government to tax the labor income when agents are younger at a relatively higher rate. 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.

2.2.2 Effect of Learning-or-Doing

Comparing equations 10 and 12, there are two channels through which introducing LOD changes the optimal tax policy. The first channel, the elasticity channel, is due to LOD altering 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.[18] Assuming that η_t is positive, these additional terms cause the expression to increase.[19] Thus, the two channels may have opposing effects on the

examines the effects with a utility function that is not homothetic with respect to labor making it harder to isolate the effects of adding LBD.

[17] 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. Moreover, 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}.

[18] The additional term in the numerator comes from both the intertemporal link and the non-separability 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.

[19] The sign of η will depend on whether the government wants to increase the relative incentive to save with training or capital. If η 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 η is positive and therefore treat it as positive in the exposition.

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. The Frisch elasticity in the LOD model 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 from adding LOD 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, since an agent will tend to spend less time training as he ages, he will supply labor relatively more elastically when he is young. Thus, the government would want to tax the labor income from agents when they are young at a relatively lower rate which can be mimicked by decreasing the tax on capital. Overall, the elasticity channel from LOD 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.[20]

Overall, adding either form of endogenous human capital alters the optimal tax policy such that age-dependent taxes are optimal. If age-dependent taxes are disallowed then a positive tax on capital will be optimal in the LBD model. In contrast, in the LOD model the optimal tax on capital will depend on the relative strengths of the elasticity and savings channels. Thus, in order to determine the importance of the human capital accumulation process when examining the optimal capital tax it is necessary to turn to a quantitative model.

3 Computational Model

Next, I contrast the optimal capital tax policies in less parsimonious models that each include one of the forms of human capital accumulation. CKK and Peterman (2013) 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

[20] 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.

that may affect the optimal capital tax.[21]

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. Overall, the model is populated with J-19 overlapping generations. Conditional on being alive at age j, Ψ_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. Thus, the cohort shares, $\{\mu_j\}_{j=1}^{J}$, are given by,

$$\mu_j = \frac{\Psi_{j-1}}{1+n}\mu_{j-1}, \text{for } j = 20, ..., J, \tag{13}$$

where μ_{20} is normalized such that

$$\sum_{j=20}^{J} \mu_j = 1. \tag{14}$$

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. In addition to choosing how much time to spend in non-leisure activities, an agent chooses consumption in order to maximize his lifetime utility,

$$u(c_j, h_j + n_j) + \sum_{s=20}^{J-j-1} \beta^s \prod_{q=1}^{s} (\Psi_q) u(c_{s+1}, h_{s+1} + n_{s+1}), \tag{15}$$

where c_j is consumption, h_j is the hours spent working, n_j is the time spent training, and β is the discount factor conditional on surviving.

In the exogenous model, an agent's age-specific human capital s_j is exogenous and pre-determined. 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 (Ω_{j-1}), previous age-specific human capital (s_{j-1}),

[21] This model is similar to the quantitative model in Peterman (2016a). However in addition to excluding idiosyncratic productivity shocks, I use a utility function that is separable and homothetic in each consumption and labor. Using this specific utility function eliminates some channels for an age-dependent tax (see Peterman (2013) for more details). Moreover, using this type of utility function allows me to decompose the strength of the elasticity channel.

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 (Ω_{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 $h_j s_j w_t$ and they 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 a Cobb-Douglas production function. Thus, the aggregate resource constraint is,

$$C_t + K_{t+1} - (1-\delta)K_t + G_t \leq K_t^{\alpha} N_t^{1-\alpha}, \tag{16}$$

where K_t, C_t, and N_t represent the aggregate capital stock, aggregate consumption, and aggregate labor (measured in efficiency units), respectively. Additionally, α is the capital share and δ is the depreciation rate for physical capital. Assuming a non-linear production function in the computational model implies prices are determined endogenously and fluctuate with regard to the aggregate capital and labor unlike in the analytically tractable model where they are fixed.

3.4 Government Policy

The government has two fiscal instruments to finance its unproductive consumption, G_t.[22] 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 $T^l[y_l]$. 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.

[22] 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.

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 which includes retirement can cause 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.[23] 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 models.

Given a social security replacement rate b, a sequence of skill accumulation parameters $\Omega_{j=20}^{jr-1}$ (endogenous models) or age-specific human capital $\{s_j\}_{j=20}^{j_r-1}$ (exogenous model), 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 $T^l : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a social security tax rate τ_{ss}, a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, social security benefits SS, prices (w, r), and transfers Tr such that:

1. Given prices, policies, transfers, and benefits, the agent maximizes equation 15 subject to
$$c_j + a_{j+1} = ws_jh_j - \tau_{ss}ws(\cdot)_jh_j, +(1+r)(a_j + Tr) - T^l[ws(\cdot)_jh_j(1-.5\tau_{ss})] - T^k[r(a_j + Tr)], \quad (17)$$
for $j < j_r$, and
$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k[r(a_j + Tr)], \quad (18)$$
for $j \geq j_r$.
Additionally,
$$c \geq 0, 0 \leq h \leq 1, a_j \geq 0, a_{20} = 0. \quad (19)$$

2. Prices w and r satisfy
$$r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \quad (20)$$
and
$$w = (1-\alpha)\left(\frac{K}{N}\right)^{\alpha}. \quad (21)$$

[23] 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.

3. The social security policies satisfy
$$SS = b\frac{wN}{\sum_{j=20}^{j_r-1}\mu_j} \tag{22}$$
and
$$\tau_{ss} = \frac{SS}{w\sum_{j=20}^{j_r-1}\mu_j}. \tag{23}$$
4. Transfers are given by
$$Tr = \sum_{j=20}^{J}\mu_j(1-\Psi_j)a_{j+1}. \tag{24}$$
5. Government balances its budget
$$G = \sum_{j=20}^{J}\mu_j T^k[r(a_j + Tr)] + \sum_{j=20}^{j_r-1}\mu_j T^l[ws(\cdot)_j h_j(1-.5\tau_{ss})]. \tag{25}$$
6. The market clears
$$K = \sum_{j=20}^{J}\mu_j a_j, \tag{26}$$
$$N = \sum_{j=20}^{J}\mu_j s(\cdot)_j h_j, \tag{27}$$
and
$$\sum_{j=20}^{J}\mu_j c_j + \sum_{j=20}^{J}\mu_j a_{j+1} + G = K^{\alpha}N^{1-\alpha} + (1-\delta)K. \tag{28}$$

4 Calibration and Functional Forms

In order to calibrate the model first I choose the 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.[24] These values are in Table 2. Since changing the human capital accumulation process fundamentally changes the model, I calibrate the second set of parameters separately in each of the 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.

[24]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: Ψ_j	Bell and Miller (2002)	Data
Pop. Growth: n	1.1%	Data
Firm Parameters		
α	.36	Data
δ	8.33%	$\frac{I}{Y} = 25.5\%$
A	1	Normalization

Table 2: **Calibration Parameters**

Parameter	Exog.	LBD	LOD	Target
Calibration Parameters				
Conditional Discount: β	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: σ_1	2	2	2	CKK
Frisch Elasticity: σ_2	0.5	0.73	0.47	Frisch= $\frac{1}{2}$
Disutility of Labor: χ	61	46	80	Avg. $h_j + n_j = \frac{1}{3}$
Government Parameters				
Υ_0	.258	.258	.258	Gouveia and Strauss (1994)
Υ_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, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{(h+n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$. Using this utility function implies that for the exogenous model 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.

Following CKK, I determine β such that the capital-to-output ratio matches U.S. data of 2.7.[25] I determine χ 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, the parameter σ_1 which controls risk aversion is set $\sigma_1 = 2$. 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.[26] Therefore, I set σ_2 such that the Frisch elasticity is at the upper bound of the range (0.5).

4.3 Age-Specific Human Capital

The age-specific human capital parameters are different in the three models. In the exogenous model, I set $\{s_j\}_{j=20}^{j_r-1}$ 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,

$$s_{j+1} = \Omega_j s_j^{\Phi_1} h_j^{\Phi_2}, \tag{29}$$

where s_j is the age-specific human capital for an agent at age j, Ω_j is an age-specific calibration parameter, Φ_1 controls the importance of an agent's current human capital on LBD, and Φ_2 controls the importance of time worked on LBD. In the LOD model, agents accumulate human capital according to the following process,

$$s_{j+1} = \Omega_j s_{j,t}^{\kappa_1} n_j^{\kappa_2}, \tag{30}$$

where n_j is the percent of an agent's time endowment he spends training. In this formulation, κ_1 controls the importance of an agent's current human capital on LOD and κ_2 controls the importance of time training on

[25] One reason that the reduced form social security program is included is to capture the relevant savings motives that can affect the capital to output ratio.

[26] Reasons for this downward 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. See Imai and Keane (2004), Domeij and Flodén (2006), Pistaferri (2003), Chetty (2009), Peterman (2016b), and Contreras and Sinclair (2008) for more details.

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 ($\{s_j\}_{j=20}^{j_r-1}$).[27]

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.[28] The values of κ_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).[29]

4.4 Firm

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

4.5 Government Policies and Tax Functions

As the baseline tax function I use the U.S. tax code estimates from Gouveia and Strauss (1994) which estimate the tax code using a three parameter functional form,

$$T(y;\Upsilon_0,\Upsilon_1,\Upsilon_2) = \Upsilon_0(y - (y^{-\Upsilon_1} + \Upsilon_2)^{-\frac{1}{\Upsilon_1}}), \tag{31}$$

where y represents the sum of labor and capital income. The average tax rate is principally controlled by Υ_0, and Υ_1 governs the progressivity of the tax policy. To ensure that taxes satisfy the budget constraint, Υ_2 is left free. Gouveia and Strauss (1994) estimate that $\Upsilon_0 = .258$ and $\Upsilon_1 = .768$ when fitting the data.[30] I refer to this calibrated tax policy as the baseline-fitted U.S. tax policy.

[27] In order to reduce the dimensionality when calibrating these sets of parameters, I restrict my attention to polynomials in age.

[28] 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. However, in the case of the LBD model, full depreciation will never bind since 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 eliminate full depreciation by assuming a skill accumulation formulation that is separable in training and past skills then the life-cycle profiles are more consistent with formal education as opposed to training (this is consistent with the profiles in Guvenen et al. (2009)). However, by setting $\kappa_1 = 1$ there is very little depreciation as long as agents spend just a small amount of time training. See Kuruscu (2006) and Heckman et al. (1998) for other examples of quantitative studies that assume little depreciation.

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

[30] 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 when calibrating the model (but allow for separate tax policies on the two sources of income when solving for the optimal).

I calibrate government consumption, G, so that it equals 17 percent of output under the baseline-fitted U.S. tax policy, consistent with CKK. 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.

The government also runs a balanced-budget social security program. Social security benefits are set so that the replacement rate, b, is 50 percent.[31] The payroll tax, τ_{ss}, is determined so that the social security system is balanced each period.

5 Computational Experiment

A social welfare function is necessary in order to determine the optimal tax policy. I use the standard social welfare function that maximizes the expected utility of an agent prior to entering the economy in the steady state. When determining the optimal tax policy I examine both flat and progressive tax policies. However, I find that for all the models the optimal tax policy consists of separate flat taxes on capital and labor income. For notational convenience, I present the computational experiment as choosing the optimal flat tax rates on capital and on labor.[32] Thus, the social welfare is function is,

$$SWF(\tau_h, \tau_k) = u(c_j, h_j + n_j) + \sum_{s=20}^{J-j-1} \beta^s \prod_{q=20}^{s} (\Psi_q) u(c_{s+1}, h_{s+1} + n_{s+1}), \tag{32}$$

where τ_h is the flat tax rate on labor income and τ_k is the flat tax rate on capital income.

6 Results

In this section, I compare the models to the data and then compare the optimal tax policies in each of the models (for a discussion of the effects on the steady state economies of both adopting endogenous human capital accumulation and adopting the optimal tax policies see Appendix B).

[31] The replacement rate matches the rate in Conesa and Krueger (2006).

[32] 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.

6.1 Comparison of Model to Data

Figure 1 plots the life cycle profiles from each of the models under the baseline-fitted U.S. tax policy and in the actual data.[33] Although there are some discrepancies between the profiles predicted by the models and the data, overall the models do a decent job matching the data.

Focusing on the labor supply profiles (the upper left panel), the model generated profiles have a similar hump shape as the data.[34] Although the models predict a labor supply profile that is generally similar to the data, there is one main difference. 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.

The labor earnings data is similar to the profiles generated by the models (see upper right panel). Both the data and model generated 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 than in the models' profiles.

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. One possible reason for this difference 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.[35] I smooth through some of the volatility in the wealth data by using five year age bins.[36] 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 data and model generated profiles is that the models predicts that agents will deplete their savings a bit more quickly than they do in the data. One potential explanation for this discrepancy could be that the model does not include any motive for individuals leaving a bequest for younger generations.

Overall, I find that the models do a fair job matching the data.

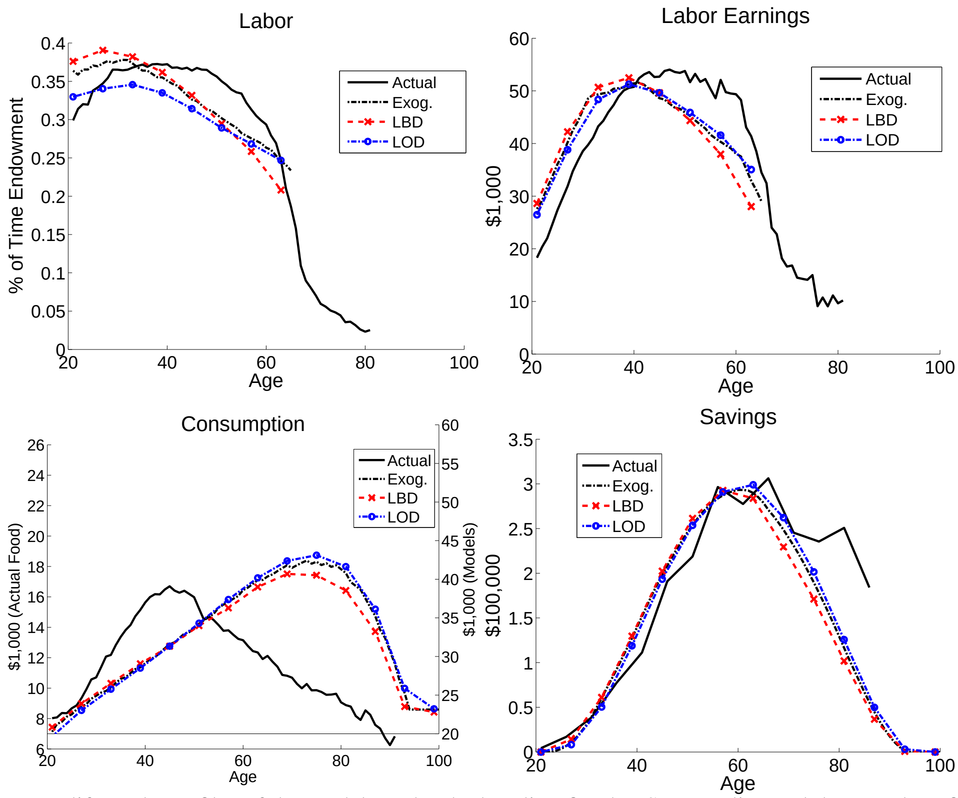
[33] Earnings, consumption, and savings from the model are converted to real 2012 dollars by equating the average earnings in the model and the data.

[34] The profiles from the data are constructed from the 1967 - 1999 waves of the Panel Survey of Income Dynamics (PSID).

[35] 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.

[36] Data after the age 80 is excluded because there are few observations in the sample leading the averages to be extremely volatile.

Figure 1: **Actual and Exogenous Life Cycle Profiles**



Note: These plots are life cycle profiles of the 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.

6.2 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 ($\tau_h = 23.7\%$). 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.[37] Similar to Peterman (2016a), I find including LBD causes the optimal capital tax to increase. In particular, the optimal capital tax increases 7.3 percentage points (forty percent) to 25.5% and the optimal labor tax decreases to 22.1%. In contrast, 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. Thus, there is large variation in the optimal capital tax depending on how human capital is accumulated.[38]

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

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

Next, I decompose the channels that cause the optimal tax policy to vary across the three models. The alteration in the Frisch labor supply elasticity profile is the principal reason for the larger optimal capital tax in the LBD model. The left panel of Figure 2 plots the lifetime Frisch labor supply elasticities in the LBD model and the exogenous model.[39] 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

[37] 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.

[38] Although their are large differences between the optimal tax policy in the LBD and exogenous models the welfare losses in the LBD model 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. Moreover, I find that the welfare losses in the LOD model from adopting the optimal tax policy solved for in the exogenous model causes a trivial welfare loss. The small welfare effects from adopting the wrong optimal tax policy are not surprising since without idiosyncratic risk there is no reason for a progressive tax policy and Peterman (2016a) 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.

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

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 σ_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 confirming that the elasticity channel is primarily responsible for the change in the optimal capital tax in the LBD model.

In the case of the LOD model, both the elasticity channel and the savings channel affect the optimal capital tax. The right panel of Figure 2 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 affecting the optimal capital tax policy.

To quantify the effect of the channels, I solve for the optimal tax policy in an alternative version of the LOD model that uses an alternative utility function,

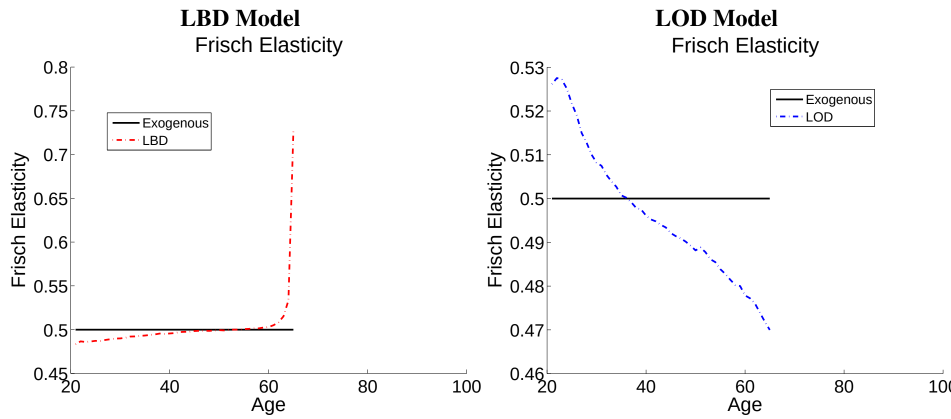
$$\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi_1 \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} - \chi_2 \frac{(n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}, \tag{33}$$

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 alternative LOD model (separable utility) is no longer a function of time spent training.[40] Eliminating the elasticity channel means that only the savings channel remains in this alternative LOD model (separable utility) with this alternative utility function.[41] The optimal tax policy in this alternative LOD model (separable utility) is $\tau_k = 19.9\%$ and $\tau_h = 23.3\%$ indicating that the savings channel causes a 1.7 percentage point increase in the optimal capital tax, relative to the exogenous model, in order to encourage agents to save via human capital as opposed to physical capital. Since 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 increase is 1.7 percentage points in the alternative LOD model (separable utility) with just the saving channel, the elasticity channel is responsible for a decrease in the optimal capital tax of 1 percentage point (canceling just over half of the savings channel's effect).

[40]In particular the Frisch elasticity is σ_2 since the utility function is separable in all three arguments.

[41]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 2.2.2 for more details.

Figure 2: **Life Cycle Frisch Labor Supply Elasticity in Endogenous Model**



Interestingly, 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 magnitude of the effect of the elasticity channel on the optimal capital tax is much larger in the LBD model. There are two potential reasons for the larger magnitude of the effect in the LBD model. First, the Frisch elasticity increases rapidly over the last two working years in the LBD model making the size of the increase in the slope much larger when LBD is included compared to the decrease in the slope caused by including LOD. Second, adding LOD to the model causes young agents to be more liquidity constrained.

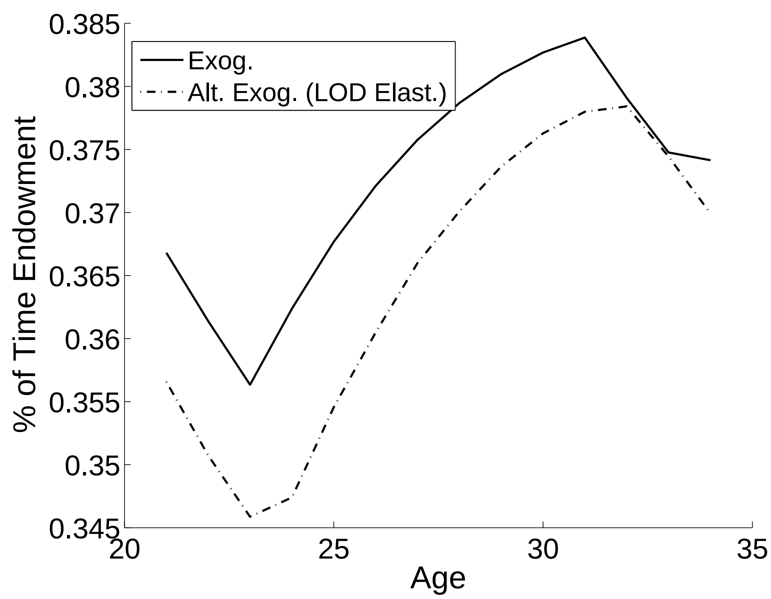
In order to determine the effect of the slope of the Frisch elasticity profile over the last few years on the optimal capital tax, I determine the optimal tax in another counterfactual exogenous model (smoothed LBD elasticity) which matches the slope of the Frisch elasticity in the LBD model from ages 20 - 63 and eliminates the rapid increase over the last few working years.[42] Even after eliminating this increase 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 larger magnitude of the effect of the elasticity channel in the LBD model compared to the LOD model is not due to the rapid increase in the Frisch elasticity in the LBD model over the last few years of the working lifetime.

Instead, including LOD causes young agents to be more liquidity constrained minimizing the effect of the elasticity channel in this model. In order to isolate this effect, I solve for a counterfactual exogenous model (LOD elasticity) which is calibrated such that the Frisch labor supply elasticity profile matches the LOD model. Generally, younger agents face lower wages due to age-specific human capital and a higher potential for binding liquidity constraints. Including LOD causes younger agents to supply labor more elastically than older agents leading them to supply even less labor early in their life compared to later in

[42]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.

life. Thus, the change in the elasticity which causes young agents to supply relatively less labor also leads young agents to be even more liquidity constrained. In response, the optimal capital tax increases in order to help 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.[43] Figure 3 confirms this intuition by demonstrating that when the Frisch elasticity profile from the LOD model is included in the counterfactual exogenous model (LOD elasticity) young agents work relatively less. Overall, 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.[44]

Figure 3: **Affect of LOD Elasticity on Young Labor Supply**



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.

[43]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.

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

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

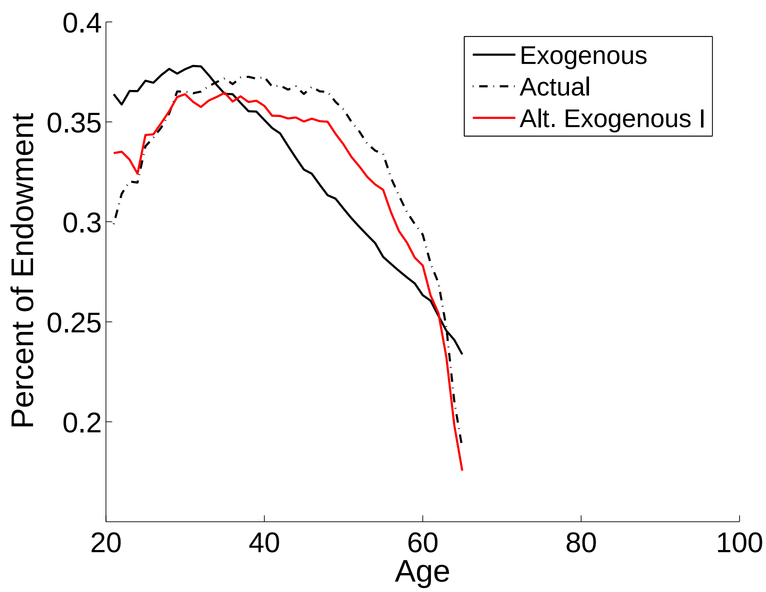
In this section, I test the relationship between the shape of the labor supply profile and the optimal tax. I examine this relationship because there are differences between the data and the labor supply profiles predicted by the models. Moreover, there are differences in the labor supply profiles between the three models. For example, comparing the labor supply profile in the actual data and the exogenous model (Figure 1), 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 1), the labor supply profile in the LBD model declines a bit more rapidly over the second half of the working lifetime than it does in the exogenous and LOD models. 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 χ over the lifetime such that the labor supply profile matches either the data or the profile predicted by the LBD model.

First, I determine the optimal tax policy in an alternative exogenous model calibrated such that it has a flatter labor supply profile in similar to the data. Figure 4 plots the labor supply profile generated in the exogenous benchmark model (solid black line), the profile in the counterfactual exogenous model calibrated to more closely match the shape of the data (red line), and the average hours worked in the data (dashed black line).[45] 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 benchmark 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 5 plots the labor supply profiles in the benchmark exogenous model (solid black), 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. Overall, these results indicate that the optimal tax policy in my benchmark exogenous model is not related to the general shape of

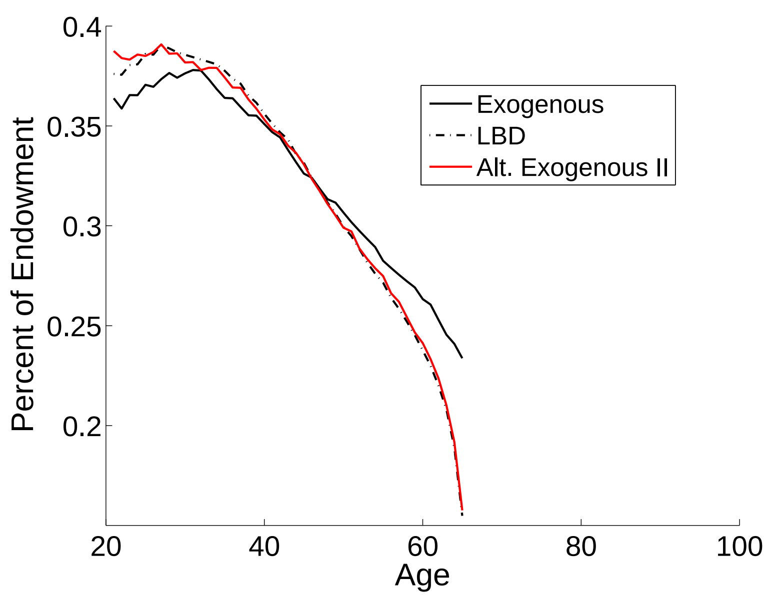
[45] The labor supply profiles are all under the baseline-fitted U.S. tax policy.

Figure 4: **Flatter Labor Supply Profile**



the labor supply profile.[46]

Figure 5: **Labor Supply in Alternative Exogenous**



[46]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.

7.2 Utility Function

In this section, I determine the effect of how human capital is accumulated 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. This utility function generates additional motives for a positive capital tax regardless of how human capital is accumulated 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 γ, which determines the relative importance of consumption and leisure, and ς, which controls risk aversion. I calibrate γ to match the percentage of the time endowment worked. Table 4 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.[47]

Table 4: **Preference Parameters**

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

7.2.2 Optimal Tax Policies in Nonseparable Models

Table 5 lists the optimal tax policies for the 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.[48] Second, since the Frisch elasticity is larger in the nonseparable model the government would prefer to rely

[47] Since this utility function has one less degree of freedom it is no longer possible to use all of the targets from the benchmark model. Consistent with CKK, I choose to eliminate the Frisch elasticity as a target.

[48] 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.

more on a capital tax, as opposed to a labor income tax in order to move some of the tax burden away from a factor that is supplied relatively more elastically.

Turning to the effect of endogenous human capital accumulation, in the model with the nonseparable utility function 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 5: **Optimal Tax Policies in Nonseparable Models**

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

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, a non-zero capital tax can be used 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.

Overall, 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. However, how exactly human capital is accumulated is still an open question. Interestingly, 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

A.1 Primal Approach

In this section I describe the general set up for the primal approach which I use to determine the optimal tax policy.[49] I use a social welfare function that maximizes the expected utility of a newborn and discounts future generations with social discount factor θ (see section 5 for more details),

$$[U(c_{2,0},h_{2,0})/\theta]+\sum_{t=0}^{\infty}\theta^t[U(c_{1,t},h_{1,t})+\beta U(c_{2,t+1},h_{2,t+1})]. \tag{34}$$

The government maximizes this objective function with respect to two constraints: the implementability constraint and the resource constraint.[50] 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)

$$c_{1,t}U_{c1}(t)+\beta c_{2,t+1}U_{c2}(t+1)+h_{1,t}U_{h1}(t)+\beta h_{2,t+1}U_{h2}(t+1)=0. \tag{35}$$

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

$$c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t=rK_t+w(h_{1,t}+h_{2,t}s_2). \tag{36}$$

A.2 Exogenous

The Lagrangian for this specification is

$$\begin{aligned}\mathscr{L}=&\frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1}-\chi\frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}+\beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1}-\chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\ &-\rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}s_2)) \\ &-\rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}s_2)) \\ &+\lambda_t(c_{1,t}^{1-\sigma_1}+\beta c_{2,t+1}^{1-\sigma_1}-\chi h_{1,t}^{1+\frac{1}{\sigma_2}}-\beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}})\end{aligned} \tag{37}$$

where ρ is the Lagrange multiplier on the resource constraint and λ is the Lagrange multiplier on the implementability constraint. The first order conditions with respect to labor, capital and consumption are

$$w\rho_t=\chi h_{1,t}^{\frac{1}{\sigma_2}}(1+\lambda_t(1+\frac{1}{\sigma_2})) \tag{38}$$

$$w\rho_{t+1}\theta\, s_2=\beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}(1+\lambda_t(1+\frac{1}{\sigma_2})) \tag{39}$$

$$\rho_t=\theta(1+r)\rho_{t+1} \tag{40}$$

$$\rho_t=c_{1,t}^{-\sigma_1}+\lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \tag{41}$$

and

$$\theta\rho_{t+1}=\beta c_{2,t+1}^{-\sigma_1}+\beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1}. \tag{42}$$

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

[50] 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.

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

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

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

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1}{s_2}\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\sigma_2}}. \tag{44}$$

Combining equation 43 and 44 yields

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1}{s_2}\left(\frac{\beta\rho_t}{\rho_{t+1}\theta}\right)\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\sigma_2}}. \tag{45}$$

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

$$\frac{\rho_t\beta}{s_2\rho_{t+1}\theta}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\sigma_2}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})}. \tag{46}$$

Combining equation 46 and 45 generates the following expression for labor taxes

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})} = 1. \tag{47}$$

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

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r). \tag{48}$$

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

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r(1-\tau_k)). \tag{49}$$

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.[51]

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

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

$$s_2\frac{U_{h1}(t)}{U_{c1}(t)} = \frac{U_{h2}(t+1)}{U_{c2}(t+1)}. \tag{50}$$

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.

implementability constraint in the LBD model is

$$c_{1,t}U_{c1}(t)+\beta c_{2,t+1}U_{c2}(t+1)+h_{1,t}U_{h1}(t)-\frac{\beta h_{1,t}U_{h2}(t+1)h_2 s_{h1}(t+1)}{s_2}+\beta h_{2,t+1}U_{h2}(t+1)=0, \tag{51}$$

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,

$$\begin{aligned}\mathscr{L}=&\frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1}-\chi\frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}+\beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1}-\chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}\\ &-\rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}s_{2,t}))\\ &-\rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}s_{2,t+1}))\\ &+\lambda_t(c_{1,t}^{1-\sigma_1}+\beta c_{2,t+1}^{1-\sigma_1}-\chi h_{1,t}^{1+\frac{1}{\sigma_2}}+\frac{\chi\beta h_{2,t+1}^{1+\frac{1}{\sigma_2}}h_{1,t}s_{h1}(t+1)}{s_2}-\beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}).\end{aligned} \tag{52}$$

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

$$\begin{aligned}w\rho_t=&\chi h_{1,t}^{\frac{1}{\sigma_2}}(1+\lambda_t(1+\frac{1}{\sigma_2}))-\theta\rho_{t+1}h_{2,t+1}s_{h1}(t+1)\\ &+\lambda_t\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}\beta h_{1,t}\left[\frac{s_{h1}(t+1)^2}{s_2^2}-\frac{s_{h1,h1}(t+1)}{s_2}\right]\end{aligned} \tag{53}$$

$$w\rho_{t+1}\theta s_2=\beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}\left[1+\lambda_t(1+\frac{1}{\sigma_2})+(1+\frac{1}{\sigma_2})\frac{h_{1,t}s_{h1}(t+1)\lambda_t}{s_2}\right] \tag{54}$$

$$\rho_t=\theta(1+r)\rho_{t+1} \tag{55}$$

$$\rho_t=c_{1,t}^{-\sigma_1}+\lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \tag{56}$$

and

$$\theta\rho_{t+1}=\beta c_{2,t+1}^{-\sigma_1}+\beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1}. \tag{57}$$

The first order conditions with respect to capital and consumption are the same in the exogenous (40, 41, and 42) and LBD models (55, 56, and 57). Therefore equation 48 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

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1}=\frac{\beta\rho_t}{\rho_{t+1}\theta} \tag{58}$$

Taking the ratio of the agent's first order conditions, equations 5 and 6 and combining with equation 58 yields

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}}=\left(\frac{h_{1,t}}{h_{2,t+1}}\right)^{\frac{1}{\sigma_2}}\left(\frac{\rho_{t+1}\theta s_2}{\beta\rho_t}\right)-\frac{h_{2,t+1}s_{h1}(t+1)}{1+r(1-\tau_k)}. \tag{59}$$

Combining equations 59, 53 and 54 the ratio of the optimal taxes on labor is,

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}} = \frac{\left(1+\lambda_t(1+\frac{1}{\sigma_2})-\lambda_t(1+\frac{1}{\sigma_2})\frac{h_{1,t}s_{h1}(t+1)}{s_2}\right)\left(1+\frac{h_{2,t+1}s_2}{1+r(1-\tau_k)}\right)}{1+\lambda_t(1+\frac{1}{\sigma_2})+h_{2,t+1}^{1+\frac{1}{\sigma_2}}h_{1,t}^{1+\frac{-1}{\sigma_2}}\frac{\lambda_t}{s_2}\left(\frac{s_{h1}(t+1)}{s_2}-s_{h1,h1}(t+1)\right)} - \frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)}. \tag{60}$$

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 8). 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 6 and 8). This constraint simplifies to $U_{n1}(t)s_2 = \beta U_{h2}(t+1)h_{2,t+1}s_n(t+1)$. I use η_t as the Lagrange multiplier on this new constraint.

The Lagrangian for the LOD model is

$$\begin{aligned}
\mathscr{L} = & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{(h_{1,t}+n_{1,t})^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\
& - \rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}s_2)) \\
& - \rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}s_2)) \\
& + \lambda_t(c_{1,t}^{1-\sigma_1}+\beta c_{2,t+1}^{1-\sigma_1}-\chi h_{1,t}^{1+\frac{1}{\sigma_2}}-\beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}) \\
& + \eta_t(\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}s_{n1}(t+1)-\chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}}s_2).
\end{aligned} \tag{61}$$

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

$$w\rho_t = \chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}}\left[1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}\right] \tag{62}$$

$$w\rho_{t+1}\theta s_2 = \beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}\left[1+\lambda_2\left(1+\frac{1}{\sigma_2}\right)-\eta_t\left(1+\frac{1}{\sigma_2}\right)s_{n1}(t+1)\right] \tag{63}$$

$$\rho_t = \theta(1+r)\rho_{t+1} \tag{64}$$

$$\rho_t = c_{1,t}^{-\sigma_1}+\lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \tag{65}$$

$$\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1}+\beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1} \tag{66}$$

and

$$\theta\rho_{t+1}h_{2,t+1}s_{n2}(t+1) = \tag{67}$$

$$\frac{\chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}}\left(\lambda_t h_{1,t}+\eta_t s_2+\sigma_2(h_{1,t}+n_{1,t})(1+\eta_t s_{n2}(t+1))\right)-\beta\chi\eta_t\sigma_2 h_{2,t+1}^{1+\frac{1}{\sigma_2}}(h_{1,t}+n_{1,t})s_{n2,n2}(t+1)}{\sigma_2(h_{1,t}+n_{1,t})} \tag{68}$$

The first order conditions with respect to capital and consumption are the same in the exogenous (40, 41, and 42) and LOD models (64, 65, and 66). Therefore equation 48 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

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta} \tag{69}$$

Taking the ratio of the agent's first order conditions, equations 5 and 6 and combining with equation 69 yields

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \left(\frac{h_{2,t+1}}{h_{1,t}+n_{1,t}}\right)^{\frac{1}{\sigma_2}}\left(\frac{\beta\rho_t}{\rho_{t+1}\theta s_2}\right). \tag{70}$$

Taking the ratio of equations 62 and 63 yields,

$$\left(\frac{h_{2,t+1}}{h_{1,t}+n_{1,t}}\right)^{\frac{1}{\sigma_2}}\left(\frac{\beta\rho_t}{\rho_{t+1}\theta s_2}\right) = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)}. \tag{71}$$

Combining equations 70 and 71 generates the following expression for the ratio of the optimal labor taxes,

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)}. \tag{72}$$

B Effects on Economy

B.1 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 6 plots the life cycle profiles of hours, consumption, assets, and age-specific human capital in all three models. Table 6 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 6: **Aggregate Economic Variables**

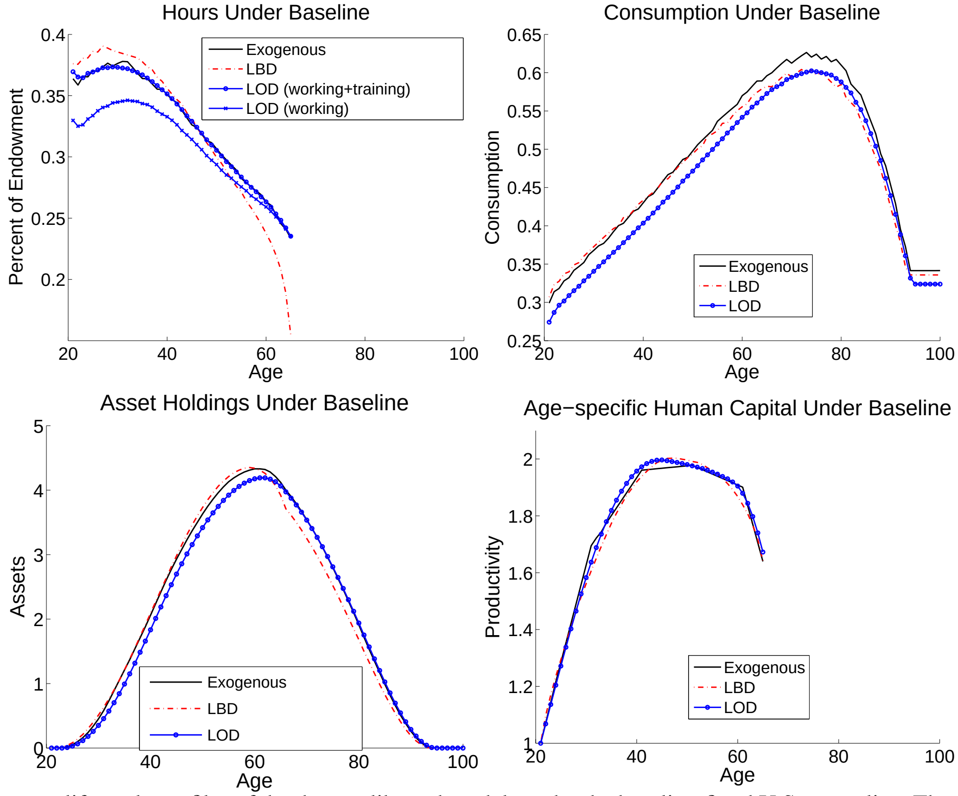
	Exogenous			**LBD**			**LOD**		
Aggregate	**Baseline**	**Optimal**	***% Change from Baseline to Optimal***	**Baseline**	**Optimal**	***% Change from Baseline to Optimal***	**Baseline**	**Optimal**	***% Change from Baseline to Optimal***
Y	0.81	0.82	*1.8%*	0.80	0.81	*1.0%*	0.76	0.77	*1.6%*
K	2.17	2.25	*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 6 demonstrate that the aggregate level of hours, labor, and capital are similar in the two models because many of the aggregate economic variables are used as calibration targets.

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 6). The upper-right panel shows that the lifetime consumption profile is steeper in the exogenous model compared to the LBD model primarily because β is lower in the LBD model. The lower value of β 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

Figure 6: **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.

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 6).

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. Adding LOD also affects the life cycle profiles. Figure 6 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.

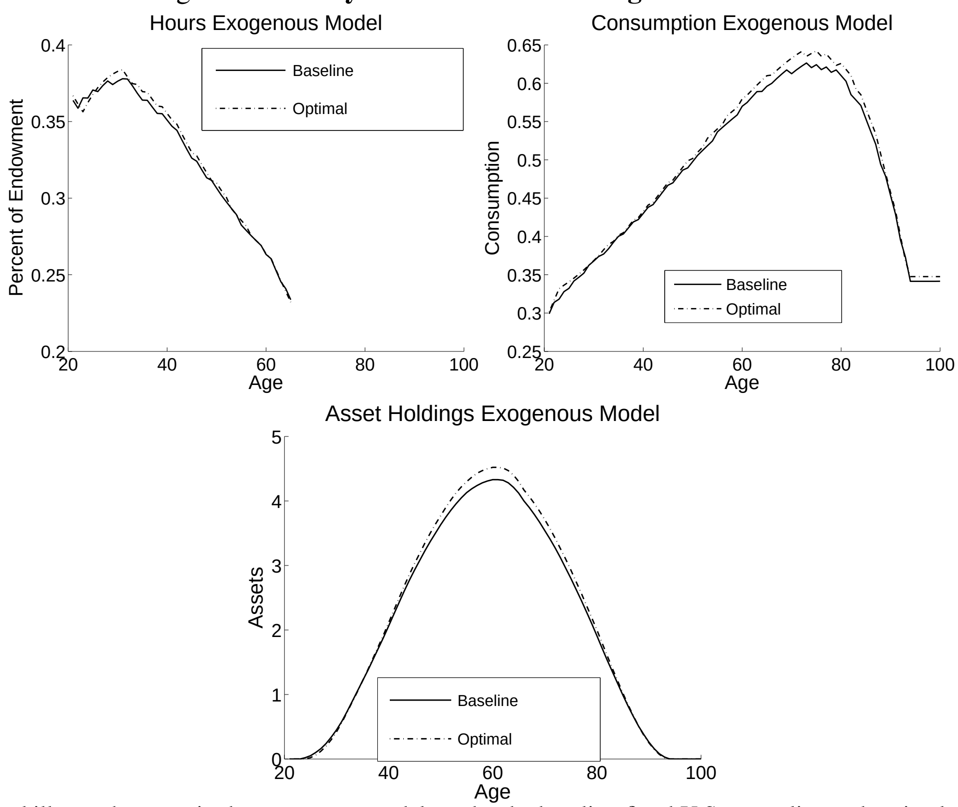
B.2 The Effects of the Optimal Tax Policy in the Exogenous Model

This section examines the effects on the economy of adopting the optimal tax policy in the exogenous model (for more detailed discussion see **?**). 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 6). The average marginal labor tax is also less under the optimal tax policy than the baseline so the labor supply increases.[52] 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 7 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. Adopting the optimal tax policy in the exogenous model causes agents to work more early in their life because 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 capital. These declines have competing effects on the marginal after-tax return on capital and the effects are uneven since the baseline-fitted US tax on capital is progressive and the optimal tax is flat. In particular, 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. Thus, while the change in the tax rate dominates for these middle-aged agents and the after tax return increases, both younger and older agents experience a decrease in the after tax return. In response, 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 7). In addition, adopting the optimal tax policy causes a steeper consumption profile for middle-aged agents with the after tax return increases (Figure 7, upper-right panel).

[52]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.

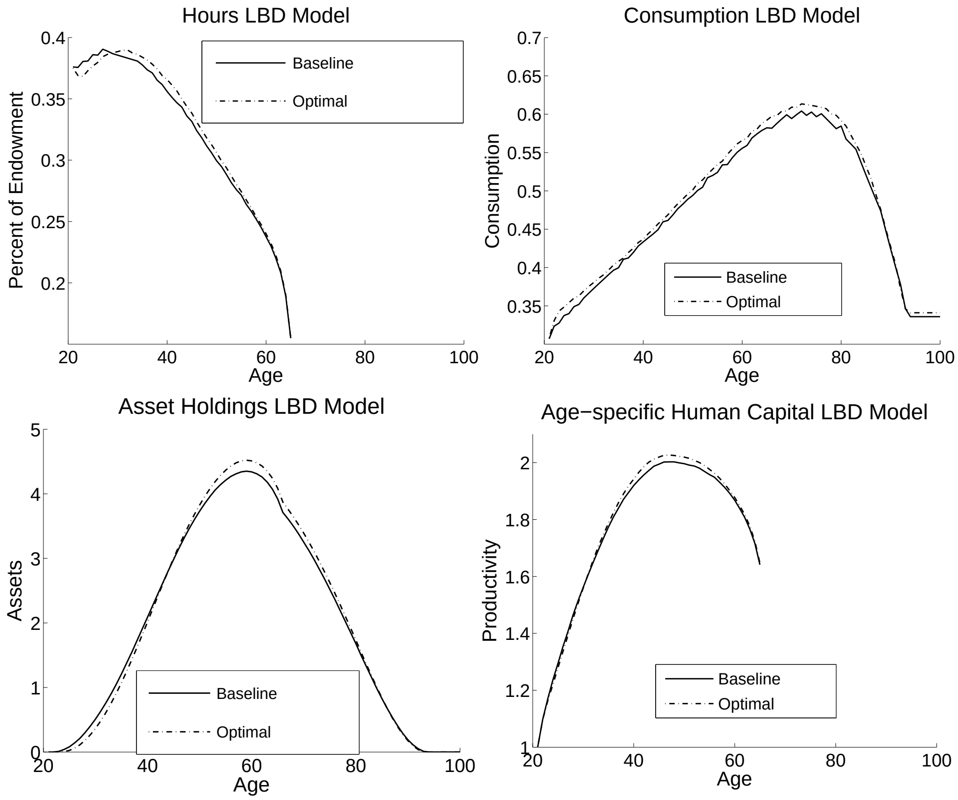
Figure 7: **Life Cycle Profiles in the Exogenous Model**



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.

B.3 The Effects of Optimal Tax Policy in the LBD Model

Figure 8: **Life Cycle Profiles in the LBD Model**



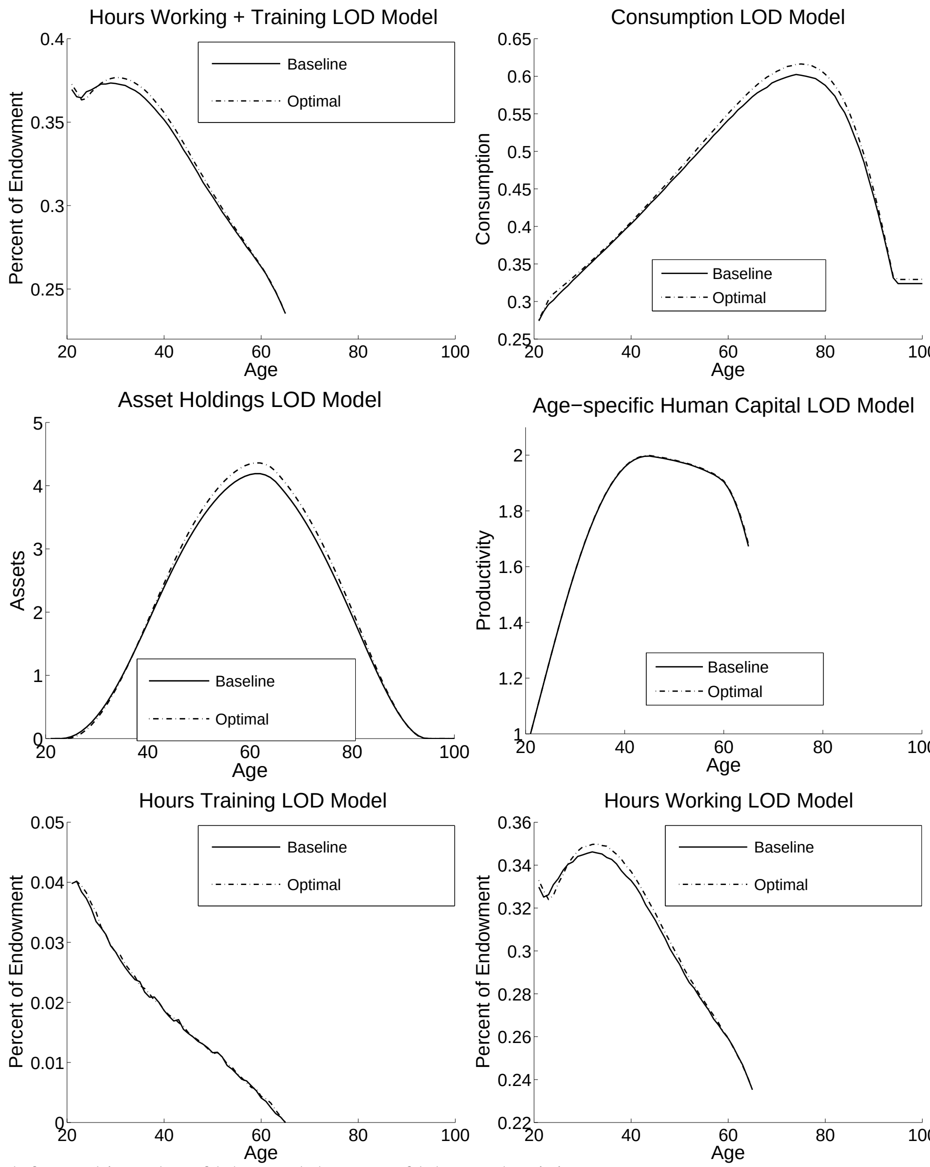
Adopting the optimal tax policy in the LBD model causes an increase in the capital tax and a decrease in the labor tax (for a more detailed discussion see Peterman (2016a)). The changes in the tax policy cause a small increase in the capital stock and a large increase in aggregate labor supply in the LBD model (see column four, five, and six of table 6). 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 8). 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 8). 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 due to the progressive baseline-fitted U.S. tax policy. Thus, 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).[53]

[53] 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.

B.4 The Effects of Optimal Tax Policy in the LOD Model

Figure 9: **Life Cycle Profiles in the LOD Model**



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

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 7 and 9): (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 9).

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.

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

_ , "The effect of endogenous human capital accumulation on optimal taxation," *Review of Economic Dynamics*, 2016, *21*, 46 – 71.

_ , "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.