The Effect of Endogenous Human Capital Accumulation on Optimal Taxation

William B. Peterman

The Effect of Endogenous Human Capital Accumulation on Optimal Taxation

William B Peterman[*]

January 18, 2012

Abstract

This paper considers the impact of endogenous human capital accumulation on optimal tax policy in a life cycle model. Including endogenous human capital accumulation, either through learning-*by*-doing or learning-*or*-doing, is analytically shown to create a motive for the government to use age-dependent labor income taxes. If the government cannot condition taxes on age, then it is optimal to use a tax on capital in order to mimic such taxes. Quantitatively, introducing learning-by-doing or learning-or-doing increases the optimal tax on capital by forty or four percent, respectively. Overall, the optimal tax on capital is thirty five percent higher in the model with learning-by-doing compared to the model with learning-or-doing implying that how human capital accumulates is of significant importance when determining the optimal tax policy.

JEL: E24, E62, H21.

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

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 such a model, taxing capital income is equivalent to an ever increasing tax on future consumption, thus implying an exponentially increasing distortion between the marginal rate of substitution and the marginal rate of transformation. In contrast, in a life cycle model agents live for a finite number of periods so the distortion imposed by a capital tax is bounded and may not necessarily be bigger than the distortions caused by other taxes. Atkeson et al. (1999), Erosa and Gervais (2002), and Garriga (2001) demonstrate in simplified life cycle models that if the government cannot condition labor income taxes on age, then it will generally tax capital in order to mimic an age-dependent tax.[1] The

---

[1]Atkeson et al. (1999) demonstrate a related result. They show conditions under which the optimal tax on capital is zero if age-dependent taxes on labor income are allowed. Gervais (2010) demonstrates that a progressive labor income tax can also mimicking age-dependent taxes on labor income.

government wants to condition taxes on age since agents vary their consumption and labor over the life cycle. Quantitative exercises, such as Conesa et al. (2009) and Peterman (2011), demonstrate that in a calibrated life cycle model the inability to condition taxes on age can be a strong motive for a positive tax on capital.

Variation in age-specific human capital causes an agent to vary his labor supply over his lifetime and hence the non-zero tax on capital result.[2] Even though age-specific human capital is a driving mechanism for the positive optimal tax on capital, it is typically incorporated in models exogenously through age-specific productivity levels. By including human capital accumulation exogenously, the models ignore any effect that endogenous accumulation may have on the optimal tax policy. This paper assesses, both analytically and quantitatively, the impact of including endogenous age-specific human capital accumulation in a life cycle model on the optimal capital tax.

Specifically, this paper explores the effect on optimal tax policy of including either of two different forms of endogenous age-specific human capital accumulation: learning-*by*-doing (LBD) and learning-*or*-doing (LOD). In LBD, an agent acquires human capital by working. 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.[3] 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 both forms, since there is empirical evidence that each form is responsible for age-specific human capital accumulation, and each is commonly employed in quantitative life cycle models.[4]

Including endogenous human capital accumulation changes the optimal capital income tax. Garriga (2001) demonstrates that in a simplified life cycle model with exogenous age-specific human capital accumulation and a utility function that is both separable and homothetic with respect to consumption and hours worked the optimal tax on capital is zero. In contrast, I analytically demonstrate that including either form of endogenous age-specific human capital accumulation in a similar model causes the optimal tax on capital to be non-zero. Including endogenous human capital accumulation creates an incentive for the government

---

[2]I define age-specific human capital as changes in an individual's labor productivity throughout his work life. Some other model features, including liquidity constraints and retirement, may also cause variations in consumption and labor over the life cycle and generate a rationale for taxes on capital income.

[3]This paper does not evaluate the effect of formal education on optimal tax policy since the paper focuses on training once an individuals begins working. Specifically, the quantitative model is calibrate to exclude time spent in school. However, the mechanisms by which LOD changes the optimal tax policy would be similar if individuals work while attending school.

[4]Examples of life cycle studies that include variations of LBD or LOD include 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). Topel (1990), Cossa et al. (1999), Altuğ and Miller (1998) provide empirical evidence of LBD. The authors show that past hours worked and length of current job tenure impact current wages. With regards to LOD, numerous studies provide evidence that individuals partake in training and that the training is responsible for wage growth. For examples see Mulligan (1995), Frazis and Loewenstein (2006), Kuruscu (2006), and Mincer (1989).

to condition labor income taxes on age. If age-dependent labor income taxes are not in the feasible policy set, then the a non-zero tax on capital is optimal in order to mimic the wedge created by conditioning labor income taxes on age. Specifically, a positive (negative) tax on capital imposes the same wedge on the marginal rate of substitution as a relatively larger (smaller) tax on young labor income. Therefore, including endogenous human capital overturns the zero capital tax result in Garriga (2001).

Furthermore, the form by which one includes endogenous human capital accumulation qualitatively changes the properties of the optimal capital income tax. Adding LBD to the model causes an agent to supply labor relatively less elastically early in his life, which alters the optimal tax policy.[5] 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 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 younger agents at a higher rate than older agents.[6] I refer to this channel as the elasticity channel since alterations to the labor supply elasticity profile is responsible for the change in the optimal tax on capital.

Adding LOD to the model also causes a non-zero tax on capital to be optimal if age-dependent taxes are unavailable. 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. The elasticity channel lowers the optimal tax on capital to implicity reduce the relative taxes on labor income of younger agents. The second channel, the savings channel, arises because training is an alternative method of saving, as opposed to accumulating physical capital. Therefore, the government can increase an agent's incentives to train by taxing capital (or taxing young labor income at relatively higher rates) since it makes training a relatively more desirable way to save. Since these two channels have counteracting effects, one cannot analytically determine the cumulative direction of their impact on the optimal tax policy.[7]

---

[5]Both Imai and Keane (2004) and Shaw (1989) point out that including LBD will alter the Frisch labor supply elasticity profile.

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

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

I quantitatively assess the effect of adding each form of endogenous age-specific human capital accumulation on optimal tax policy in a calibrated life cycle model using the specific utility function from Garriga (2001). The optimal tax rates in the model with exogenous age-specific human capital accumulation (exogenous model) are 18.2 percent on capital and 23.7 percent on labor. I find that adding either form of endogenous human capital increases the optimal tax on capital. In the model with LBD the optimal tax rates are 25.5 percent on capital and 22.1 percent on labor. The optimal tax rates in the model with the LOD framework are 18.9 percent on capital and 23.6 percent on labor. Adding endogenous age-specific human capital accumulation raises the optimal tax on capital by approximately forty percent in the LBD framework and approximately four percent in the LOD framework. Overall, the optimal tax on capital is 35 percent higher in the model with LBD compared to the model with LOD indicating that the form by which human capital endogenously accumulates is quantitatively important.

I test the sensitivity of these results with respect to the utility function. I find that using an alternative utility function that is neither separable nor homothetic with respect to consumption and hours worked implies that the optimal tax on capital is much larger in the exogenous model. The optimal tax on capital is larger because with this utility function the Frisch labor supply elasticity profile is upward sloping regardless of the form of human capital accumulation. I find that including either form of endogenous human capital accumulation with this utility function causes an even larger increase in the optimal tax on capital. The optimal tax on capital increases by approximately forty six percent and fifteen percent when LBD and LOD are included, respectively. The optimal tax on capital is approximately twenty seven percent higher in the the LBD model compared to the LOD model. Therefore, even in this set up which has a large motive for a tax on capital in the exogenous model, the form by which human capital endogenously accumulates has large impacts on the optimal capital tax.

Correia (1996), Armenter and Albanesi (2009), and Jones et al. (1997) demonstrate that in a model where the government has an incomplete set of tax instruments a non-zero tax on capital may be optimal in order to mimic the missing taxes. This paper combines two related strands of the literature that examine the optimal tax on capital in such a model where the government does not have a complete set of tax instruments. The first strand examines the optimal tax on capital in a calibrated life cycle model with exogenous human capital where the government cannot condition labor income taxes on age. Conesa et al. (2009), henceforth CKK, solve a calibrated life cycle model to determine the optimal tax on capital. They determine that the optimal tax policy is a flat 34 percent tax on capital and a flat 14 percent tax on labor income.[8] They

---

[8]This is model M4 in Conesa et al. (2009). I refer to CKK's model that abstracts from idiosyncratic earnings risk and within-cohort heterogeneity because they find that these features do not affect the level of the optimal tax on capital. Therefore, I also abstract from these features in my analysis.

state that a primary motive for imposing a high tax on capital income is to mimic a relatively larger labor income tax on younger agents when they supply labor relatively less elastically. An agent supplies labor more elastically as he ages because his labor supply is decreasing, and the authors use a utility specification in which the agent's Frisch labor supply elasticity is a negative function of hours worked. Peterman (2011) confirms that this is an economically significant motive for the positive tax on capital in a model similar to CKK's model, but concludes that the restriction on the government from being able to tax accidental bequests at a different rate from ordinary capital income is also a large contribution to the positive optimal tax on capital. This paper extends these previous life cycle studies of optimal tax policy by determining the effects of endogenous age-specific human capital accumulation on the optimal tax policy in a standard life cycle model.

This paper is related to a second strand of the literature that analyzes the trade off between labor and capital taxes in a model that includes endogenous age-specific human capital accumulation but not in a life cycle model.[9] 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 it is not optimal to distort either human or physical capital accumulation in the long run. Reis (2007) shows in a similar model that if the government cannot distinguish between consumption and human capital investment, then the optimal tax on capital 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 through both LBD and LOD causes the optimal tax on capital to increase because a higher tax on capital unravels the labor market frictions. The labor market frictions causes lower employment in the economy. In Chen et al. (2010), a tax on capital causes the wage discount to increase, thus causing firms to post more vacancies which in turn causes an increase in worker participation. This paper is related to this second strand of literature, however it differs in that it analyzes optimal tax policy in an overlapping generations (OLG) model as opposed to an infinitely lived agent model. Therefore, the second strand of literature does not account for the effects of endogenous human capital accumulation through life cycle channels. It is especially important to include the life cycle channel since Conesa et al. (2009) and Peterman (2011) demonstrate that this channel is

---

[9]Jacobs and Bovenberg (2009) analyze the trade off between a labor and capital tax in a life cycle model with pre-work education. The authors find that in a two-period model where agents acquire education in the first period and work in the second period the optimal tax on capital is generally positive if educational investment is not verifiable. The tax on capital reduces the tax on labor income, which in turn reduces the distortions on the benefit to education. Jacobs and Bovenberg (2009) is related to the current work; however they focus on human capital accumulation prior to working while this study examines the results of endogenous human capital accumulation once an agent begins to work.

quantitatively important for motivating a positive tax on capital. This paper combines both strands of the literature and determines the optimal tax policy in a life cycle model with endogenously determined human capital.

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

1 Analytical Model

In this section, I demonstrate that adding endogenous human capital accumulation overturns the result from Garriga (2001) that for a specific utility function that is separable and homothetic in both consumption and labor the government has no incentive to condition labor income taxes on age. I begin this section by setting up the agent's problem and demonstrating that a positive (negative) tax on capital induces a wedge on the marginal rate of substitution that is similar to a relatively larger tax on young (old) labor income. Using the primal approach, I then solve for the optimal tax policy in the exogenous model, with a benchmark utility function that is homothetic with respect to consumption and hours worked, $U(c, h) = \frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$. I confirm the Garriga (2001) result that the government has no incentive to condition labor income taxes on age and therefore the optimal tax on capital is zero. I show that adding endogenous human capital accumulation to this model causes the optimal tax policy to include age-dependent taxes and if age-dependent taxes are unavailable, then a non-zero tax on capital is optimal. I also demonstrate the channels by which the forms of endogenous human capital accumulation affect the optimal tax policy.

I derive these analytical results in a tractable two-period version of the computational model. For tractability purposes, the features I abstract from include: retirement, population growth, progressive tax policy, and conditional survivability. Additionally, I assume that the marginal products of capital and labor are constant. This assumption permits me to focus on the life cycle elements of the model, in that changes to the tax system do not affect the pre-tax wage or rate of return. Since the factor prices do not vary, I suppress their time subscripts in this section. All of these assumptions are relaxed in the computational model.

1.1 Exogenous Age-Specific Human Capital

1.1.1 General Set-up

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

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

where β is the discount rate, $c_{j,t}$ is the consumption of an age j agent at time t, and $h_{j,t}$ is the percent of the time endowment the agent works.[10] Age-specific human capital is normalized to unity when the agent is young. At age two, age-specific human capital is ε_2.[11] The agent maximizes equation 1 with respect to consumption and hours subject to the following constraints

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

and

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

where $a_{1,t}$ is the amount young agents save, $\tau_{h,j}$ is the tax rate on labor income for an agent of age j, 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. I assume that the tax rate on labor income can be conditioned on age; however, the tax rate on capital income cannot.[12] I combine equations 2 and 3 to form a joint intertemporal budget constraint:

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

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

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

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

---

[10]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.

[11]Since age-specific human capital is exogenous and predetermined in this model, it could also be considered age-specific productivity.

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

and

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

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

1.1.2 Tax on Capital Mimics Age-Dependent Tax on Labor

To demonstrate why a tax on capital has a similar effect to an age-dependent labor income tax, I derive the intertemporal Euler equation by combining equations 5, 6, and 7:

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

Equation 8 demonstrates that if the government wants to create a wedge on the marginal rate of substitution by varying the age-dependent labor income taxes, then τ_k is an alternative option. A positive tax on capital mimics the wedge imposed by a relatively higher tax rate on young labor income. Therefore, if the government has an incentive to condition taxes on age but age-dependent taxes are unavailable, then a non-zero tax on capital is optimal.

1.1.3 Primal Approach

I use the primal approach to determine the optimal tax policy.[13] I use a social welfare function that maximizes utility and discounts future generations with social discount factor θ,

$$[U(c_{2,0}, h_{2,0})/\theta] + \sum_{t=0}^{\infty} \theta^t [U(c_{1,t}, h_{1,t}) + \beta U(c_{2,t+1}, 1 - h_{2,t+1})]. \tag{9}$$

The government maximizes this objective function with respect to two constraints: the implementability constraint and the resource constraint.[14] 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)

$$c_{1,t}U_{c1}(t) + \beta c_{2,t+1}U_{c2}(t+1) + h_{1,t}U_{h1}(t) + \beta h_{2,t+1}U_{h2}(t+1) = 0. \tag{10}$$

---

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

[14] There is a third constraint: the government budget constraint. However, I assume that the government is allowed to hold constant levels of savings or debt. Therefore, due to Walras' Law, I only need to include two of three constraints in the Lagrangian and leave out the government budget constraint.

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

$$c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t = rK_t + w(h_{1,t} + h_{2,t}\varepsilon_2). \tag{11}$$

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

$$\begin{aligned}
\mathcal{L} = &\frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\
&- \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\varepsilon_2)) \\
&- \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\varepsilon_2)) \\
&+ \lambda_t(c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi h_{1,t}^{1+\frac{1}{\sigma_2}} - \beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}})
\end{aligned} \tag{12}$$

where ρ is the Lagrange multiplier on the resource constraint and λ is the Lagrange multiplier on the implementability constraint.

1.1.4 Optimal Tax Policy

I solve for the optimal tax policy in the analytically tractable exogenous model. The formulation of the government's problem and their first order conditions for this model can be found in appendix A.1. Combining the government's first order conditions generates the following expression for optimal labor income taxes:

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})} = 1. \tag{13}$$

Equation 13 demonstrates that the government has no incentive to condition labor income taxes on age when age-specific human capital is exogenous.[15]

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

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

---

[15]This result is specific to this utility function. See Garriga (2001) for further details.

Applying the benchmark utility function to equation 7 provides the following relationship:

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r(1-\tau_k)). \tag{15}$$

Equations 14 and 15 demonstrate that in order for the household to choose the optimal allocation indicated by the primal approach, the tax on capital must equal zero.[16]

1.2 Learning-by-Doing

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

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

$$c_{1,t} + a_{1,t} = (1-\tau_{h,1})h_{1,t}w \tag{17}$$

and

$$c_{2,t+1} = (1+r(1-\tau_k))a_{1,t} + (1-\tau_{h,2})s_2(h_{1,t})h_{2,t+1}w. \tag{18}$$

The agent's first order conditions are given by

$$\frac{U_{h1}(t)}{U_{c1}(t)} = -[w(1-\tau_{h,1}) + \beta\frac{U_{c2}(t+1)}{U_{c1}(t)}w(1-\tau_{h,2})h_{2,t+1}s_{h1}(t+1))], \tag{19}$$

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -ws_2(h_{1,t})(1-\tau_{h,2}), \tag{20}$$

and

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1+r(1-\tau_k)). \tag{21}$$

---

[16]Regardless of whether the government can condition labor income taxes on age, in this model they do not want to tax capital because there is no desire to mimic an age-dependent tax on labor income. When the government cannot condition labor income taxes on age then the Lagrangian includes an additional constraint:

$$\varepsilon_2\frac{U_{h1}(t)}{U_{c1}(t)} = \frac{U_{h2}(t+1)}{U_{c2}(t+1)}. \tag{16}$$

However, in the analytically tractable model with exogenous human capital accumulation, this constraint is not binding and thus the Lagrange multiplier equals zero.

The first order conditions with respect to h_2 and a_1 are similar in the LBD and exogenous models (equations 20 and 21 are similar to equations 6 and 7). However, the first order condition with respect to h_1 is different in the two models (equation 19 is different from equation 5) because working has the additional human capital benefit in the LBD model. This human capital benefit also alters the implementability constraint. Suppressing the arguments of the skills function, the implementability constraint in the LBD model is

$$c_{1,t}U_{c1}(t) + \beta c_{2,t+1}U_{c2}(t+1) + h_{1,t}U_{h1}(t) - \frac{\beta h_{1,t}U_{h2}(t+1)h_2 s_{h1}(t+1)}{s_2} + \beta h_{2,t+1}U_{h2}(t+1) = 0, \quad (22)$$

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

The formulation for the government's problem and the resulting first order conditions (utilizing the benchmark utility function) are in appendix A.2. Combining the first order conditions from the government's problem and suppressing the time arguments yields the following ratio for optimal labor income taxes,

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}} =$$

$$\frac{\left[1+h_{2,t+1}s_{h2}\right]\left[1+\lambda_t(1+\frac{h_{1,t}s_{h2}}{s_2})(1+\frac{1}{\sigma_2})\right]}{1+\lambda_t(1+\frac{1}{\sigma_2}) - \beta h_{2,t+1}^{1+\frac{1}{\sigma_2}}h_{1,t}^{-\frac{1}{\sigma_2}}\left[\frac{s_{h2}}{s_2}(1+\lambda_t(1+\frac{h_{1,t}s_{h2}}{s_2}))(1+\frac{1}{\sigma_2}) - h_1\left(\left(\frac{s_{h2}}{s_2}\right)^2 - \frac{s_{h2,h2}}{s_2}\right)\right]}$$

$$-\frac{h_{2,t+1}s_{h2}}{1+r(1-\tau_k)}. \qquad (23)$$

Equation 23 demonstrates that generally in the LBD model the government has an incentive to condition labor income taxes on age. This result contrasts with the exogenous model, in which the government has no incentive to condition labor income taxes on age (see equation 13).

1.2.2 LBD Enhances Motive for Postive Tax on Capital

I solve for the intertemporal Euler equation (by combining equations 19, 20 and 21) to demonstrate why including LBD causes the optimal tax policy to include age-dependent taxes and which agents the government wants to tax at a higher rate,:

$$s_2(h_{1,t})\frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1+r(1-\tau_k))\frac{1-\tau_{h,1}}{1-\tau_{h,2}} + \beta h_{2,t+1}s_{h1}(t+1). \qquad (24)$$

Comparing equation 8 and equation 24, it is clear that the LBD intertemporal Euler equation has an extra term that is positive. Therefore, holding all else equal, the tax on young labor income must be relatively

higher to induce the same wedge on the marginal rate of substitution in the LBD model.[17] If age-dependent taxes are not in the governments feasible set, then a larger tax on capital will be necessary to induce the same wedge.

By examining the Frisch elasticities in the exogenous and LBD models, it is clear why adding LBD increases the optimal relative tax on young labor income or tax on capital. Since the functional forms of these elasticities extend to a model where agents live for more than two periods, I denote an agent's age with i. In the exogenous model, the Frisch elasticity simplifies to $\Xi_{\text{exog}} = \sigma_2$. The Frisch elasticity in the LBD model is, $\Xi_{\text{LBD}} = \frac{\sigma_2}{1 - \frac{h_{i+1,t+1} w_{t+1}(h_{i,t}\sigma_2 s_{hi,hi}(t+1) - s_{hi}(t+1))}{s_{i,t}(1+r_t(1-\tau_k))w_t}}$. [18]

The Frisch elasticity in the exogenous model is constant and valued at σ_2. In the LBD model, the extra terms in Ξ_{LBD} increase the size of the denominator, thus holding hours and consumption constant between the two models, $\Xi_{\text{exog}} > \Xi_{\text{LBD}}$. Intuitively, the inclusion of the human capital benefit makes workers less responsive to a one-period change in wages since the wage benefit is only part of their total compensation for working in the LBD model. However, the human capital benefit does not have a constant effect on an agent's Frisch elasticity over his lifetime. The relative importance of the human capital benefit decreases over an agent's lifetime because he has fewer periods to use his higher human capital as he ages.[19] Therefore, adding LBD causes a young agent to supply labor relatively less elastically than an older agent. This shift in relative elasticities creates an incentive for the government to tax the labor income of younger agents at a relatively higher rate. Thus, if the government cannot condition labor income taxes on age, then the optimal tax on capital is higher in the LBD model. 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.

1.3 Learning-or-Doing

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

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

---

[17]In the case of the benchmark utility function, the government does not want to introduce a wedge in the exogenous model. In this example, holding all else equal, the government needs to include a relatively larger labor income tax on young agents in the LBD model to induce the same zero wedge on the left-hand side of the equation.

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

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

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

$$c_{1,t} + a_{1,t} = (1-\tau_{h,1})h_{1,t}w \tag{25}$$

and

$$c_{2,t+1} = (1+r(1-\tau_k))a_{1,t} + (1-\tau_{h,2})s_2(n_{1,t})h_{2,t+1}w. \tag{26}$$

The agent's first order conditions are given by

$$\frac{U_{h1}(t)}{U_{c1}(t)} = -[w(1-\tau_{h,1})], \tag{27}$$

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -ws_2(n_{1,t})(1-\tau_{h,2}), \tag{28}$$

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1+r(1-\tau_k)), \tag{29}$$

and

$$\frac{U_{n1}(t)}{U_{c2}(t+1)} = -\beta w(1-\tau_{h,2})s_{n1}(n_{1,t})h_{2,t+1}. \tag{30}$$

The first order conditions with respect to h_1, h_2, and a_1 are similar in the LOD model and the exogenous model (equations 27, 28, and 29 are similar to equations 5, 6, and 7).[20] However, since agents have the additional choice variable n_1 in the LOD model, the have a first order condition with to this variable (equation 30). This new first order condition requires an additional constraint in the government's Lagrange that equates equations 28 and 30. This constraint simplifies to $U_n(t)s_2 = \beta U_{h2}(t+1)h_{2,t+1}s_n(t+1)$. I use η_t as the Lagrange multiplier on this new constraint.

The formulation of the government's problem and resulting first order conditions are provided in appendix A.3. Combing the first order conditions yields the following relationship for optimal taxes on labor income:

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right) + \frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right) - \eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)}. \tag{31}$$

Equation 31 demonstrates that the government has an incentive to condition labor income taxes on age when LOD is introduced into the model.

---

[20]Since the first order conditions with respect to hours and savings are the same in the LOD and exogenous model, the implementability constraints are the same.

Although equation 31 shows that including LOD creates an incentive for the government to condition labor income taxes on age, it is unclear at which age the government wants to impose a relatively higher labor income tax. Comparing equations 13 and 31, there are two channels through which introducing LOD changes the optimal tax policy. The first channel results from using a utility function that is non separable in training and labor. The non separability affects the optimal tax policy through the elasticity channel since it causes LOD to alter the Frisch elasticity. This channel causes the numerator of the ratio to include the additional term $\frac{h_{1,t}}{h_{1,t}+n_{1,t}}$. As a result of this new term, the expression decreases.

The second channel results from the intertemporal link created by the additional constraints. I refer to this channel as the savings channel because this model has an additional intertemporal link since agents can save via training. 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.[21] Assuming that η_t is positive, these additional terms cause the expression to increase.[22] Thus, the two channels have opposing effects on the optimal tax policy, and the overall effect is unclear.[23]

Examining the Frisch labor supply elasticities provides intuition for how the first channel affects the optimal tax policy. In the exogenous model, the Frisch elasticity for the benchmark utility specification is constant, σ_2. Since the altered utility function is not additively separable in time spent working and training, the Frisch labor supply elasticity is not constant in the LOD model. The Frisch elasticity for the altered utility function is $\Xi_{\text{LOD}} = \frac{\sigma_2(h+n)}{h}$. This functional form implies that an agent supplies labor relatively more elastically when LOD is included in the model because the agent has a substitute for working in the form of training. Additionally, the effect on the Frisch elasticity is larger when he spends a larger proportion of his non-leisure time training (or when training is a better substitute for for generating lifetime income). Therefore, if an agent spends less time training as he ages, then he will supply labor relatively more elastically when he is young, and the government would want to tax the labor income from young agents at a relatively lower rate. One way to mimic this age-dependent tax is to decrease the tax on capital. Therefore, the elasticity channel from LOD causes a decrease in the optimal tax on capital.

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

---

[21]The 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.

[22]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.

[23]If an alternative utility function is used that is additively separable in training and hours then LOD only affects the optimal tax policy through part of the second channel and the ratio increases if η is positive. An example of such a utility function is $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi_1\frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} - \chi_2\frac{(n)^{1+\frac{1}{\sigma_3}}}{1+\frac{1}{\sigma_3}}$.

when he is young equals 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 reduced by decreasing the tax on capital or by increasing the tax on older labor income. By adopting either of these changes, the government makes it relatively more beneficial for the agent to use ordinary capital to save as opposed to training. Therefore, the government decreases the tax on capital to promote a larger ordinary capital stock.

Overall, adding LOD creates an incentive for the government to condition labor income taxes on age. If the government cannot condition labor income taxes on age, then it would want to implement a non-zero tax on capital to mimic an age-dependent tax. The tax on capital increases when LBD is added to the model. Adding LOD causes the optimal tax on capital to change, however the direction of the change is not analytically clear.

2   Computational Model

To determine the direction and magnitude of the effect of adding endogenous human capital accumulation on optimal tax policy, I solve for the optimal tax policies in the LBD and LOD models and compare them with the exogenous model. The exogenous model is adapted from CKK; however I use a different benchmark utility function so that the elasticity channel does not effect the optimal tax policy in the exogenous model. Additionally, since the authors find that neither idiosyncratic earnings risk nor heterogenous ability types are important motives for a positive tax on capital income, I exclude these sources of heterogeneity. In this section I describe the models and define the competitive equilibrium for each model.

2.1   Demographics

In the computational model, time is assumed to be discrete, and there are J overlapping generations. Conditional on being alive at age j, Ψ_j is the probability of an agent living to age $j+1$. All agents who live to an age of J die in the next period. If an agent dies with assets, the assets are confiscated by the government and distributed equally to all the living agents as transfers (Tr_t). All agents are required to retire at an exogenously set age j_r.

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

$$\mu_j = \frac{\Psi_{j-1}}{1+n}\mu_{j-1}, \text{for } i = 2,....,J, \tag{32}$$

where μ_1 is normalized such that

$$\sum_{j=1}^{J} \mu_j = 1. \tag{33}$$

2.2 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 providing labor services to the market. An agent chooses consumption as well as how to spend his time endowment in order to maximize his lifetime utility

$$\sum_{j=1}^{J} \prod_{q=1}^{j} (\Psi_{q-1}) \beta^{j-1} u(c_j, h_j + n_j), \tag{34}$$

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

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

2.3 Firm

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

$$C_t + K_{t+1} - (1 - \delta) K_t + G_t \leq K_t^{\alpha} N_t^{1-\alpha}, \tag{35}$$

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. Unlike the analytically tractable model, I do not assume a linear production function in the computational model, so prices are determined endogenously and fluctuate with regard to the aggregate capital and labor.

2.4 Government Policy

The government has two fiscal instruments to finance its consumption, G_t, which is in an unproductive sector.[24] 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 three 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.

In addition to raising resource 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 (2011) demonstrates that excluding this type of program in a model with exogenously determined retirement causes unrealistic life cycle profiles. 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. An agent's social security benefits are independent of his personal earnings history. Social security is financed by taxing labor income at a flat rate, $\tau_{ss,t}$. The payroll tax rate $\tau_{ss,t}$ is set to 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.

2.5 Definition of Stationary Competitive Equilibrium

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

---

[24]Including G_t such that it enters the agent's utility function in an additively separable manner is an equivalent formulation.

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

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

$$\sum_{j=1}^{J} \text{Max}_{c_j, h_j, a_{j+1}} \beta^{j-1} [\prod_{q=0}^{j-1} \Psi_q] u(c_j, h_j) \tag{36}$$

subject to

$$c_j + a_{j+1} = w\varepsilon_j h_j - \tau_{ss} w\varepsilon_j h_j, + (1+r)(a_j + Tr) - T^l[w\varepsilon_j h_j(1 - .5\tau_{ss})] - T^k[r(a_j + Tr)], \tag{37}$$

for $j < j_r$, and

$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k[r(a_j + Tr)], \tag{38}$$

for $j \geq j_r$. Additionally,

$$c \geq 0, 0 \leq h \leq 1 \tag{39}$$

and

$$a_j \geq 0, a_1 = 0. \tag{40}$$

2. Prices w and r satisfy

$$r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \tag{41}$$

and

$$w = (1-\alpha) \left(\frac{K}{N}\right)^{\alpha}. \tag{42}$$

3. The social security policies satisfy

$$SS = b \frac{wN}{\sum_{j=1}^{j_r-1} \mu_j} \tag{43}$$

and

$$\tau_{ss} = \frac{ss \sum_{j=j_r}^{J} \mu_j}{w \sum_{j=1}^{j_r-1} \mu_j}. \tag{44}$$

4. Transfers are given by

$$Tr = \sum_{j=1}^{J} \mu_j (1 - \Psi_j) a_{j+1}. \tag{45}$$

5. Government balances its budget

$$G = \sum_{j=1}^{J} \mu_j T^k[r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j T^l[w\varepsilon_j h_j(1 - .5\tau_{ss})]. \tag{46}$$

6. The market clears

$$K = \sum_{j=1}^{J} \mu_j a_j, \tag{47}$$

18

$$N = \sum_{j=1}^{J} \mu_j \varepsilon_j h_j, \qquad (48)$$

and

$$\sum_{j=1}^{J} \mu_j c_j + \sum_{j=1}^{J} \mu_j a_{j+1} + G = K^\alpha N^{1-\alpha} + (1-\delta)K. \qquad (49)$$

3 Calibration and Functional Forms

To determine the optimal tax policy it is necessary to choose functional forms and calibrate the model's parameters. Calibrating the models involves a two-step process. The first step is choosing parameter values for which there are direct estimates in the data. These parameter values are in table 1. Second, to calibrate the remaining parameters, values are chosen so that under the baseline-fitted U.S. tax policy certain targets in the model match the values observed in the U.S. economy.[25] These values are in table 2

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

Adding endogenous human capital accumulation to the model fundamentally changes the model. Accordingly, if the calibration parameters are the same, then the value of the targets will be different in the endogenous and exogenous models. To assure that all the models match the targets under the baseline-fitted U.S. tax policy, I calibrate the set of parameters based on targets separately in the three models. This calibration implies that these parameters are different in the exogenous and endogenous models.

---

[25]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.

3.1 Demographics

In the model, agents are born at a real world age of 20 that corresponds to a model age of 1. Agents are exogenously forced to retire at a real world age of 65. If an individual survives until the age of 100, he dies the next period. I set the conditional survival probabilities in accordance with the estimates in Bell and Miller (2002). I assume a population growth rate of 1.1 percent.

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 to 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

3.2 Preferences

Agents have time-separable preferences over consumption and labor services, and conditional on survival, they discount their future utility by β. I use the benchmark utility function for the exogenous and LBD models, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$, and an altered form of this utility function for the LOD model, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h+n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$.

I determine β such that the capital-to-output ratio matches U.S. data of 2.7.[26] 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.[27] Following CKK, I set $\sigma_1 = 2$, which controls the relative risk aversion.[28] Past micro-econometric studies estimate the Frisch elasticity to be between 0 and 0.5.[29] However, more recent research has shown that these estimates may be biased downward. Reasons for this bias include: utilizing weak instruments; not accounting for borrowing constraints; disregarding the life cycle effect of endogenous-age specific human capital; omitting correlated variables such as wage uncertainty; and not accounting for labor market frictions.[30] Furthermore, Rogerson and Wallenius (2009) show that because individuals make

---

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

[27]Using a target of one-third is standard in quantitative exercises. For examples, see CKK, Nakajima (2010), and Garriga (2001).

[28]Even though CKK use a different utility specification, their specification has a parameter that corresponds to σ_1.

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

[30]Some of these studies include Imai and Keane (2004), Domeij and Flodén (2006), Pistaferri (2003), Chetty (2009), and Contr-

decisions regarding labor on both the intensive and extensive margins "micro and macro elasticities need not be the same, and ... macro elasticities can be significantly larger."[31] Therefore, I set σ_2 such that the Frisch elasticity is at the upper bound of the range (0.5). The preference parameters are summarized in table 2.

3.3 Age-Specific Human Capital

The age-specific human capital parameters that require calibration are different in the exogenous and endogenous models. In the exogenous model, I set $\{\varepsilon_j\}_{j=0}^{j_r-1}$ so that the sequence matches a smoothed version of the relative hourly earnings estimated by age in Hansen (1993). In the LBD model, agents accumulate age-specific human capital according to the following process,

$$s_{j+1} = \Omega_j s_j^{\Phi_1} h_j^{\Phi_2}, \tag{50}$$

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,

$$s_{j+1} = \Omega_j s_{j,t}^{\kappa_1} n_j^{\kappa_2}, \tag{51}$$

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 LOD.[32] In the endogenous models I do not set $\{s_j\}_{j=0}^{j_r-1}$ directly, rather I calibrate the sequence $\{\Omega_j\}_{j=1}^{j_r-1}$ such that the agent's equilibrium labor or training choices cause $\{s_j\}_{j=0}^{j_r-1}$ under the baseline-fitted U.S. tax code to match the age-specific human capital calibrated in the exogenous model ($\{\varepsilon_j\}_{j=0}^{j_r-1}$).

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. The value of $\kappa_1 = 1$ implies that there is little depreciation of human capital when skill accumulation is the result of LOD.[33] The values of κ_2 and $\{\Omega_j\}_{j=1}^{j_r-1}$ 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.

---

eras and Sinclair (2008).

[31] Erosa et al. (2011) provides a quantitative estimate of the impact of the extensive margin on aggregate labor supply fluctuations.

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

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

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).[34]

3.4 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 the observed investment output ratio of 25.5 percent. These parameters are summarized in table 1.

3.5 Government Policies and Tax Functions

To calibrate parameters based on the targets, it is necessary to use a baseline tax function that mimics the U.S. tax code so that I can find the parameter values that imply the targets in the models that match the values in the data. I use the estimates of the U.S. tax code in Gouveia and Strauss (1994) for this tax policy, which I refer to as the baseline-fitted U.S. tax policy. The authors match the U.S. tax code to the data using a three parameter functional form,

$$T(y; \Upsilon_0, \Upsilon_1, \Upsilon_2) = \Upsilon_0(y - (y^{-\Upsilon_1} + \Upsilon_2)^{-\frac{1}{\Upsilon_1}}), \tag{52}$$

where y represents the sum of labor or capital income. The average tax rate is principally controlled by Υ_0, and Υ_1 governs the progressivity of the tax policy. 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. The authors do not fit separate tax functions for labor and capital income. Accordingly, I use a uniform tax system on both sources of income for the baseline-fitted U.S. tax policy. I calibrate government consumption, G, so that it equals 17 percent of output under the baseline-fitted U.S. tax policy, as observed in the U.S. data.[35] Therefore, Υ_2 is determined as the value that equates government spending to 17 percent of GDP. When searching for the optimal tax policy, I restrict my attention to revenue neutral changes that imply that government consumption is equal under the baseline-fitted U.S. tax policy and the optimal tax policy.

In addition to government consumption, the government also runs a balanced-budget social security program. Social security benefits are set so that the replacement rate, b, is 50 percent.[36] The payroll tax, τ_{ss},

---

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

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

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

is determined so that the social security system is balanced each period.

4 Computational Experiment

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

$$SWF(\tau_h, \tau_k) = \sum_{j=1}^{J} \beta^{j-1} [\prod_{q=0}^{j-1} \Psi_q] u(c_j, h_j), \tag{53}$$

where τ_h is the flat tax rate on labor income and τ_k is the flat tax rate on capital income.[37]

To determine the effects of endogenous human capital accumulation, I compare the tax policies that maximize the SWF in the three models. When I determine the optimal tax policy, I test different values of τ_h and determine values for τ_k so that the changes in the tax policy are revenue neutral. Therefore, the experiment is to find τ_h that satisfies

$$\max_{\tau_h} SWF(\tau_h, \tau_k) \tag{54}$$

subject to,

$$G = \sum_{j=1}^{J} \mu_j \tau_k r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j \tau_h [w s_j h_j (1 - .5\tau_{ss})]. \tag{55}$$

5 Results

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

---

[37]When searching for the optimal tax policy I search over both flat and progressive tax policies. However, I determined that the optimal tax policy was a flat taxes on both capital and labor. For notational convenience I present the computational experiment as choosing the optimal flat tax rates on capital and on labor.

5.1 Optimal Tax Policies in Exogenous, LBD, and LOD Models

Table 3 describes the optimal tax policies in the three models. The optimal tax policy in the exogenous model is an 18.2 percent flat tax on capital income ($\tau_k = 18.2\%$) and a 23.7 percent flat tax on labor income ($\tau_h = 23.7\%$).[38] While the optimal tax on capital is smaller in the exogenous model compared with CKK, it is not zero. The motives that cause a positive tax on capital in the exogenous model include: the inability of the government to borrow; agents being liquidity constrained, and the government not being able to tax transfers at a separate rate from ordinary capital income. See Peterman (2011) for a thorough discussion of the relative strengths of these motives in a model similar to the exogenous model.

The optimal tax policy in the LBD model is $\tau_k = 25.5\%$ and $\tau_h = 22.1\%$, and in the LOD model it is $\tau_k = 18.9\%$ and $\tau_h = 23.6\%$. The optimal tax on capital is approximately forty percent larger in the LBD model and four percent larger in the LOD compared to the exogenous model.

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

With respect to LBD, the alteration in the Frisch labor supply elasticity profile is the principal reason that the optimal tax on capital increases. The left panel of figure 1 plots the lifetime Frisch labor supply elasticities in the LBD model and the exogenous model under the optimal tax policy. 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 tax on capital is higher in the LBD model in order to implicitly tax younger agents, who supply labor less elastically, at a higher rate.

To quantify the effect of the elasticity channel on the optimal tax policy in the LBD model, I alter the exogenous model 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. In order to match the shapes of the profiles, I vary σ_2 in

---

[38]I checked whether a progressive tax on either capital or labor was optimal. However, I found that the optimal tax policies were always flat taxes. This result is similar to CKK who find that the optimal tax policies are flat in their model that is similar to the exogenous model. CKK find that a progressive tax on labor income is optimal only if the model includes within-cohort heterogeneity. Since all the agents within a cohort are homogenous in my models, one would expect flat taxes to be optimal. This is in contrast to Gervais (2010) who 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.

the exogenous model by age. I find that the optimal tax policy in this altered exogenous model, $\tau_k = 25.6\%$ and $\tau_h = 22.1\%$, is almost identical to the optimal tax policy in the LBD model. The optimal tax policy in the altered exogenous model demonstrates that the elasticity channel is responsible for the change in the optimal tax on capital in the LBD model.[39]

In section 1.3.1 I show that both the elasticity channel and the savings channel affect the optimal tax on capital in the LOD model. Adding LOD to the model causes young agents to supply labor relatively more elastically. The right panel of figure 1 plots the Frisch elasticity profile in the exogenous and LOD models. The elasticity channel causes a decrease in the optimal tax on capital so that younger agents who supply labor more elastically are implicitly taxed at a lower rate. Additionally, the inclusion of LOD allows individuals to use training to save, which activates the savings channel. Analytically, one can not determine the directions of the saving channel's effect of the effect on optimal tax policy.

To quantify the direction of the saving channel's effect and significance of both channels, I solve for the optimal tax policy in an alternative version of the LOD model that excludes the effect from the elasticity channel. I utilize an alternative utility function, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi_1 \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} - \chi_2 \frac{(t)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$, which is separable in training and hours worked. Since the utility function is separable, the Frisch labor supply elasticity is no longer a function of the time spent training. The Frisch elasticity with this utility function is constant, at the value σ_2, so the elasticity channel is eliminated.[40] The optimal tax policy in this model with the alternative utility function is $\tau_k = 19.9\%$ and $\tau_h = 23.3\%$. These results indicate that in this model, the savings channel results in an increase in the optimal tax on capital of 1.7 percentage points, which encourages agents to save via physical capital as opposed to human capital. I find that that the elasticity channel causes the optimal tax on capital to decrease 1 percentage point cancelling just over half of the savings channel's effect.

5.2 The Effects of Adding Endogenous Age-Specific Human Capital

This section analyzes the effect on the aggregate economic variables and life cycle profiles from adding LBD and LOD to the exogenous model under the baseline-fitted U.S. tax policy. Figure 2 plots the life cycle profiles of hours, consumption, assets, and age-specific human capital in all three models. Table 4 describes the optimal tax policies and summarizes the aggregate economic variables under both the baseline-fitted U.S. tax policy and optimal tax policies. The first, fourth, and seventh columns are the aggregate economic

---

[39]The Frisch elasticity profile in the LBD model rises rapidly in the last few period before retirement du to the function form of the human capital accumulation equation. I confirm that the optimal tax policy in an altered exogenous model without the increase in the Frisch elasticity over the last few periods still matches the optimal tax policy in the LBD model. Therefore, it is the upward slope of the profile over the whole life and not just the last few periods that is responsible for driving the results.

[40]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 1.3.1 for more details.

Figure 1: **Life Cycle Frisch Labor Supply Elasticity in Endogenous Model**
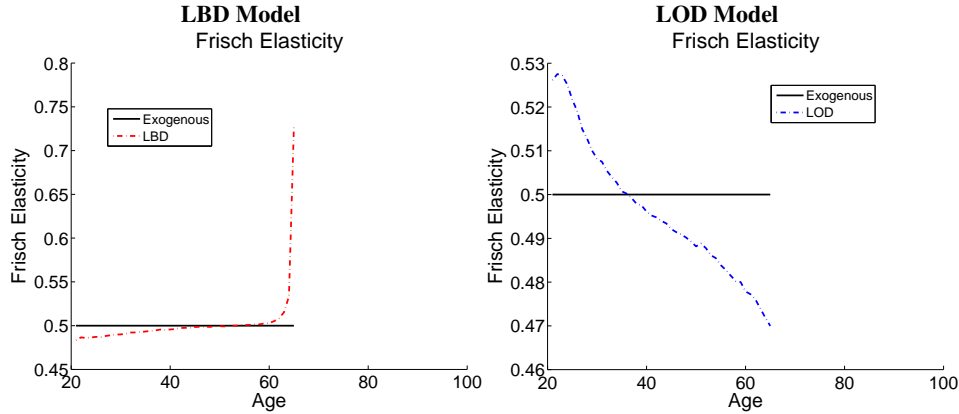
variables under the baseline-fitted U.S. tax policy in the exogenous, LBD, and LOD models, respectively. The second, fifth, and eighth columns are the aggregate economic variables under the optimal tax policies. The third, sixth, and ninth columns are the percentage changes in the aggregate economic variables induced from adopting the optimal tax policies.

Table 4: **Aggregate Economic Variables**

| | Exogenous | | | LBD | | | LOD | | |
| | | | *% Change from Baseline to* | | | *% Change from Baseline to* | | | *% Change from Baseline to* |
Aggregate	**Baseline**	**Optimal**	***Optimal***	**Baseline**	**Optimal**	***Optimal***	**Baseline**	**Optimal**	***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.

Comparing the first and fourth columns of table 4, it is clear that the levels of aggregate hours, labor supply, and aggregate capital are similar in the exogenous and LBD models. The calibrated parameters are

determined so that under the baseline-fitted U.S. tax policy the models match certain targets from the data. Since many of the aggregate economic variables are targets and these calibration parameters are determined separately in the exogenous and LBD models, the aggregates are similar in the two models.

Figure 2: **Life Cycle Profiles under Baseline-Fitted U.S. Tax Policy**
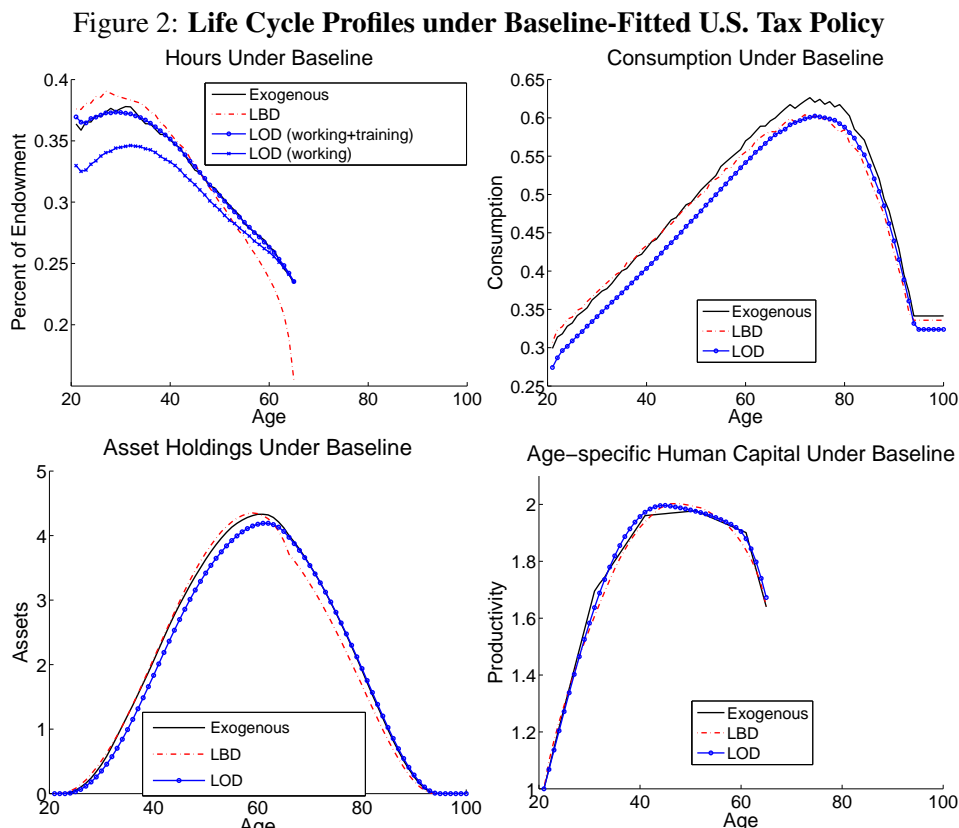
Note: These plots are life cycle profiles of the three calibrated models under the baseline-fitted U.S. tax policy. There are two labor lines for the LOD model, one solely for hours worked and the other for hours worked plus hours spent training.

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 2).

The upper-right panel shows that the lifetime consumption profile is steeper in the exogenous model. The intertemporal Euler equation controls the slope of consumption profile over an agent's lifetime. The relationship is

$$\left(\frac{c_{j+1}}{c_j} \right)^{\sigma_1} = \Psi_j \beta \widetilde{r_t}, \tag{56}$$

where $\widetilde{r_t}$ is the marginal after-tax return on capital. In order to induce the same capital to output ratio in the LBD model, β is lower which leads to the flatter consumption profile. In all three models the consumption

27

profile falls towards the end of the lifetime due to the low probability of survival. 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 the two models since the sequence of parameters $\{\Omega j\}_{j=1}^{j_r-1}$ is calibrated so that age-specific human capital matches (see the lower-right panel of figure 2).

Although the parameters values are calibrated such that the targets still match, the size of the economy is smaller in the LOD model because agents must spend time training in this model. Comparing the first and seventh columns of table 4, aggregate labor supply, and physical capital are smaller in the LOD model compared with the exogenous model because an agent spends part of his time endowment training. However, it is apparent that the relative ratios of the aggregates are similar in the two models since the factor prices are comparable.

Adding LOD also affects the life cycle profiles. Figure 2 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 that includes 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.

5.3   The Effects of the Optimal Tax Policy in the Exogenous Model

This section examines the effects on the economy of adopting the optimal tax policy in the exogenous model. In the exogenous model, the optimal tax on capital 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 figure 4). The average marginal tax on labor is also less under the optimal tax

policy than the baseline so the labor supply increases.[41] The increase in labor supply is relatively less than the increase in capital so the rental rate on capital decreases and the wage rate increases.

To compare the welfare effects of adopting the optimal tax policies in the models, I compute the consumption equivalent variation (CEV). The CEV is the uniform percentage increase in consumption, at each age, needed to make an agent indifferent between being born under the baseline-fitted U.S. tax policy and the optimal tax policy. Therefore, a positive CEV indicates a welfare increase due to tax reform. Overall, adopting the optimal tax policy in the exogenous model causes a welfare increase of 0.7 percent CEV.[42]

Figure 3: **Life Cycle Profiles in the Exogenous Model**
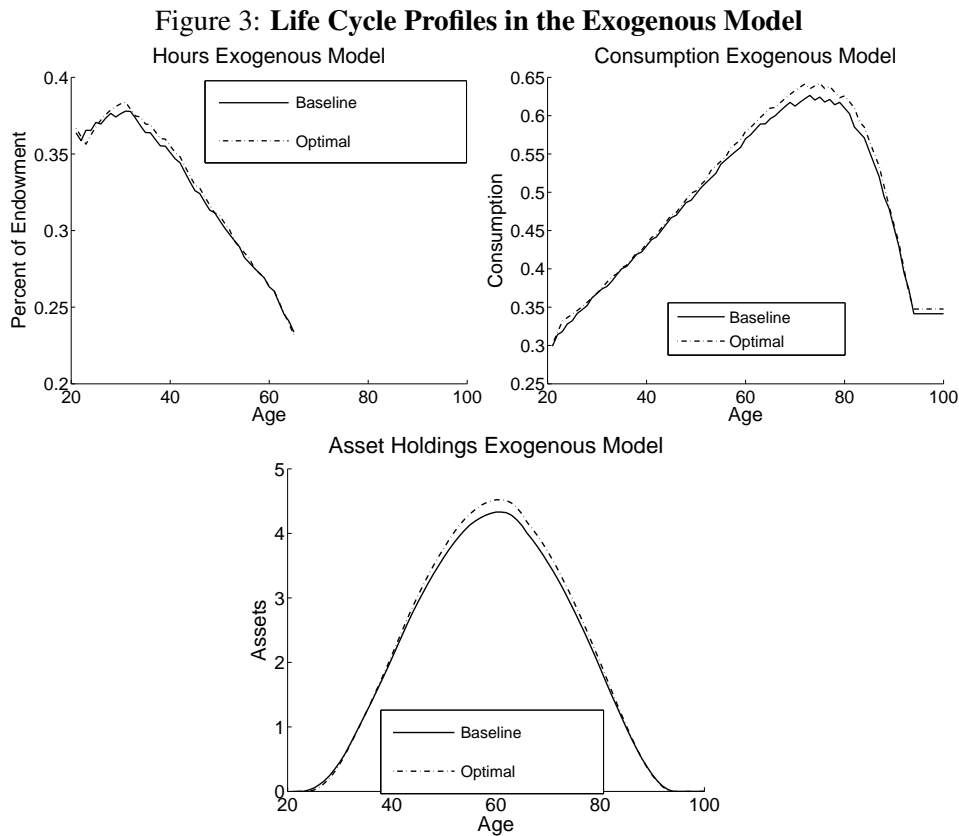
Note: Since the skills are the same in the exogenous models under the baseline-fitted U.S. tax policy and optimal tax policy, they are not plotted.

Figure 3 plots the life cycle profiles for time worked, consumption, assets and age-specific human capital in the exogenous model under the baseline-fitted U.S. tax policies and the optimal tax policies. The solid

---

[41] 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.

[42] Adopting the optimal tax policy causes two changes. First, it eliminates the progressivity of the baseline-fitted U.S. tax policy. Additionally, it adjusts the relative ratio of the tax on capital to labor. To test the relative significance of each change, I find the CEV between the baseline tax policy and a flat tax policy that raises the same amount of revenue from each income source as under the baseline tax policy. I find that eliminating progressivity represents a majority of the increase in the CEV since the model does not contain any with-in cohort heterogeneity.

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 changes in all three life cycle profiles: (i) early in their life, agents work relatively more; (ii) agents save more, especially during periods when they are wealthier; and (iii) the lifetime consumption profile steepens. Comparing the profiles in the left panel of figure 3, it is evident that agents 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 tax on capital and the rental rate on capital. These changes have competing effects on the marginal after-tax return on capital that are not consistent for all individuals. Specifically, since the baseline fitted US tax on capital is progressive and the optimal tax is flat, the change in the tax rate has an uneven effect on an agent's net return over his lifetime. The decrease is larger for agents who hold more savings since their marginal tax rate was higher under the progressive baseline fitted US tax policy. Overall, the after tax return increases for middle-aged agents and decreases for younger and older agents when the optimal tax policy is adopted. In response to these changes, middle-aged individuals increase there savings under the optimal tax policy. In contrast, younger and older agents decrease their savings (see the lower left panel of figure 3).

Similar to savings, the slope of an agents consumption profile is controlled by the marginal after-tax return to capital (see equation 56). Therefore, adopting the optimal tax policy causes a steeper consumption profile for middle-aged agents who have a higher after-tax return to capital (figure 3, upper-right panel).

5.4 The Effects of Optimal Tax Policy in the LBD Model

Adopting the optimal tax policy in the LBD model causes an increase in the tax on capital and a decrease in the tax on labor (see column four, five, and six of table 4). Since adopting the optimal tax policies causes the capital tax to change in different directions in the exogenous and LBD models, the aggregate economic variables react differently. The changes in the tax policy cause a small increase in the capital stock and a large increase in aggregate labor supply in the LBD model. The larger rise in labor compared to capital translates into an decrease in the wages and a increase in the rental rate on capital. Adopting the optimal tax policy causes a welfare gain of 0.9 percent CEV.

Implementing the optimal tax policies also causes the life cycle profiles to change (see figures 4). The labor supply profile (upper-right panel of figure 4) is affected by the increase in the tax on capital which implicitly taxes labor income from early years at a higher rate. This change results in the shift of time worked from earlier to later years. Because agents work more in their middle years, age-specific human capital is also higher for middle aged agents (see the lower-right panel).

30

Figure 4: **Life Cycle Profiles in the LBD Model**
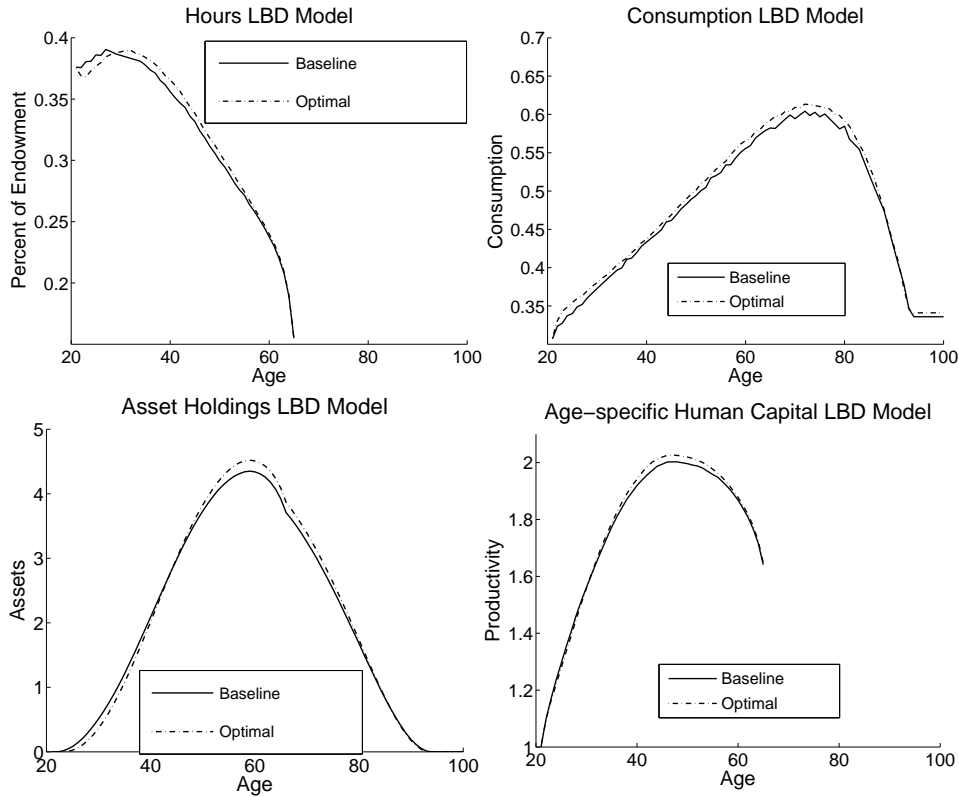
Applying the optimal tax policy introduces two opposing effects on the agent's lifetime asset profile. First, agents increase their savings under the optimal tax policy because the economy is larger. Second, the larger tax on capital under the optimal tax policy decreases the average marginal after-tax return on capital, causing agents to hold fewer assets. The first effect is constant for all agents. The second effect is not constant for all agents, but it is negatively proportional to an agent's capital income because the baseline-fitted U.S. tax policy is progressive and the optimal tax policy is flat. This means that the second effect is relatively stronger when agents save less and relatively weaker when they save more. As is apparent in the lower-left panel of figure 4, adopting the optimal tax policy causes agents to save less at ages when they had lower savings under the baseline-fitted U.S. tax policy (early and late in life), and to save more at ages when they held larger savings under the baseline-fitted U.S. tax policy (in the middle of their life).

Finally, implementing the optimal tax policy causes the consumption profile to uniformly shift upward (see the upper-right panel). The profile shifts upward due to an increase in the overall size of the economy.[43]

Figure 5: **Life Cycle Profiles in the LOD Model**
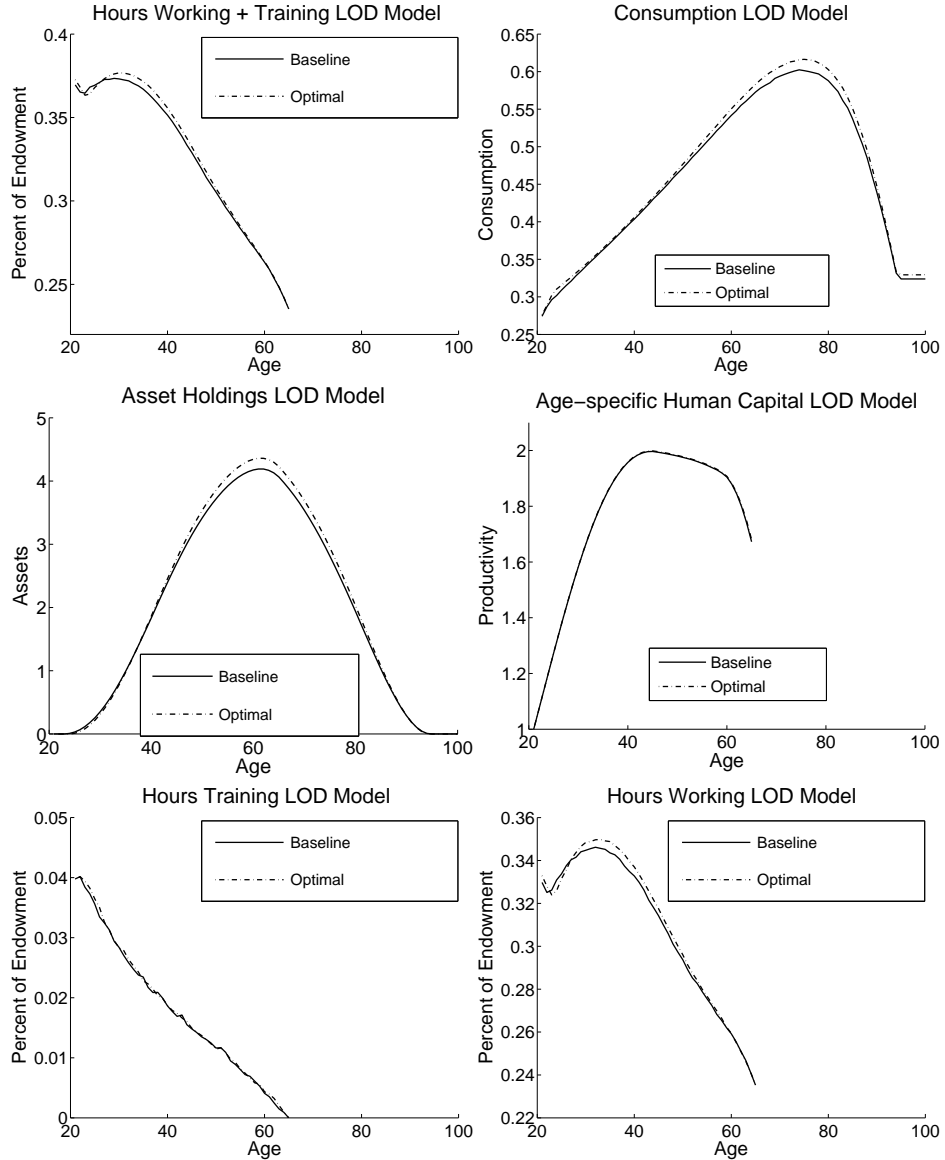
Note: The upper-left panel is a plot of labor and the sum of labor and training.

5.5 The Effects of Optimal Tax Policy in the LOD Model

Although the optimal tax on capital is larger in the LOD model than in the exogenous model, 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 a decrease in the tax on 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. The CEV from adopting the adopting the optimal tax policy is 0.8 percent in the LOD model, similar to the exogenous model.

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 3 and 5): (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. The average marginal tax on capital is smaller in the optimal tax policy compared with the tax rate in baseline-fitted U.S. tax policy, meaning that the implicit tax on young labor income is smaller than old labor income. Therefore, agents react by shifting hours worked to earlier in their lifetime (see lower-right panel of figure 5). Agents train a similar amount under the optimal tax policy so human capital is also similar (see the middle-left and middle-right profile).

As with the exogenous model, adopting the optimal tax policy causes a decrease in both the tax on capital and the rental rate on capital. These have counteracting effects on the agent's savings decisions. During the early and later years of an agent's life, the tax on capital falls less since the baseline-fitted U.S. tax policy is progressive; therefore the decrease in the rental rate dominates and the agent holds less savings. In the middle of an agent's life the tax on capital is larger under the baseline-fitted U.S. tax policy, so the drop in the tax from adopting the optimal policy dominates, and agents hold more savings.

Adopting the optimal tax policy causes an agent's consumption profile to be steeper. The slope of the profile is controlled by the after tax return on capital. Therefore, the change in the slope is more pronounced for ages when agents hold more assets.

6 Sensitivity Analysis

Next I check the sensitivity of the results with respect to the utility specification by determining the effects of endogenous human capital accumulation on optimal tax policy in a model with an alternative utility function, $U(c_{1,t}, 1-h_{1,t}) = \frac{(c_{1,t}^{\gamma}(1-h_{1,t})^{1-\gamma})^{1-\varsigma}}{1-\varsigma}$. This utility function is the benchmark specification in CKK. I

---

[43] 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.

refer to this utility function as the nonseparable utility function. This function includes an additional motives for a positive tax on capital. Atkeson et al. (1999), Erosa and Gervais (2002), Garriga (2001), and CKK demonstrate that a utility function that is neither homothetic or separable in consumption and labor creates a motive for a positive tax on capital. Under such a utility specification, the labor supply elasticity is a negative function of hours worked. An agent's labor supply elasticity profile tends to slope upwards in simulations using this utility function since their labor supply profile generally slopes downward. The optimal tax on capital is therefore larger to implicitly tax younger labor income that is supplied less elastically at a higher rate. I begin by presenting the new calibration parameters followed by the optimal tax policies in all three models.

6.1 Changes in Calibration

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

Table 5 lists the calibration parameters for the nonseparable utility parameters. The Frisch elasticity in the exogenous model for this utility function is $\frac{(1-h)}{h}\frac{1-\gamma(\varsigma-1)}{\varsigma}$. The average Frisch elasticity implied by the calibration in the exogenous model is 1.13, which is more than twice as large as with the benchmark utility specification in the exogenous model. However, section 3.2 expresses reasons why a larger Frisch elasticity may be in line with unbiased empirical estimates. Note, the Frisch elasticity is a decreasing function in hours, meaning it is no longer constant in the exogenous model (as long as hours worked vary over the lifetime).

Table 5: **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

Adding LBD and LOD changes workers' incentives and choice variables, and the relevant preference parameters change accordingly. The main difference is that in the LBD model agents enjoy the human

capital benefit. Because working is more valuable in the LBD model, the value for γ is smaller to generate the same level of aggregate hours. A lower γ decreases the relative importance of consumption compared to leisure, which implies a lower Frisch elasticity in the LBD model than in the exogenous model. The rest of the parameters are similar in the exogenous and LBD models.

Adding LOD provides agents with an alternative method of saving. Since agents use human capital as part of their savings, to induce the capital to output ratio to be the same in the LOD model, β must be higher. A higher value for β implies agents are more patient, placing a higher value on future consumption. To finance future consumption, agents increase their savings, and the capital to output ratio increases. A higher value for β also implies that agents will spend more time training. Since γ is set to target the sum of time spent in non-leisure activities, the value for γ must drop to keep agents working and training for one-third of their endowment.

6.2 Optimal Tax Policies in Nonseparable Models

There is a larger motive for a positive tax on capital in all the models with nonseparable 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 the life, the Frisch elasticity profile is upward sloping in all the models. I find that adding LOD or LBD only changes the degree by which the Frisch elasticity is upward sloping. The upward sloping Frisch elasticity profile motivates a large positive tax on capital.[44] Second, there are less degrees of freedom when calibrating the model so the Frisch elasticity is larger in the nonseparable model. Therefore, the government would prefer to rely on a capital tax, as opposed to a labor income tax. Both of these motives cause the optimal tax on capital to be larger in a model which uses the nonseparable utility.

Table 6 lists the optimal tax policies for the nonseparable models. Even with the nonseparable utility — in which the exogenous model contains a large motive for a tax on capital — adding either form of endogenous human capital causes the optimal tax on capital to increase by a larger percent. In the case of LBD, the optimal tax on capital increases by approximately 45 percent because including LBD causes an increase in the slope of the Frisch elasticity profile. Adding LOD causes approximately a 15 percent increase in the optimal tax on capital. Once again, there are two channels by which LOD changes the optimal tax on capital. The savings channel which causes the optimal tax on capital to increase and the elasticity channel which causes the optimal tax on capital to decrease. Although adding LOD causes a decrease in the slope of the Frisch elasticity profile, the Frisch elasticity profile is still upward sloping. Therefore, the impact of

[44] Adding LBD causes the Frisch elasticity to be even steeper and further enhances this motive for a positive tax on capital

the elasticity channel in the LOD is smaller with the non-separable utility function. However, the relative impact of the elasticity channel is not affected by the non-separable utility function. Therefore, including LOD causes a larger increase in the optimal capital tax with the non-separable utility function than with the benchmark utility function.

Table 6: **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

7 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 including endogenous human capital accumulation creates a motive for the government to condition labor income taxes on age and in their absence, it is optimal to use a non-zero tax on capital to mimic these age-dependent taxes. Quantitatively, I finds that including either form of endogenous human capital accumulation causes the optimal tax on capital to rise. Furthermore, the form by which one includes endogenous human capital accumulation matters. I find that the optimal tax on capital is approximately 35 percent larger with LBD compared to LOD. Many of the previous computational life cycle studies model age-specific human capital exogenously. However, these findings demonstrate that this assumption has large impacts on optimal tax policy.

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

to increase in numerical simulations.

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

References

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

Altuğ, Sumru and Robert Miller, "The Effect of Work Experience on Female Wages and Labour Supply," *Review of Economic Studies*, 1998, *65*.

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.

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

Chamley, 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.

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

Cossa, Ricardo, James Heckman, and Lance Lochner, "Wage Subsidies and Skill Formation: A Study of the Earned Income Tax Credit," Working Paper, University of Rochester 1999.

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.

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

Frazis, Harley and Mark Loewenstein, "On-the-Job-Training," *Foundations and Trends in Microeconomics*, 2006, *2* (5), 363–440.

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

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

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

Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan, "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," 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.

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.

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.

Mincer, Jacob, "Job Training: Costs, Returns, and Wage Profiles," Working Paper 3208, NBER 1989.

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

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

Peterman, William, "Determining the Motives for a Positive Optimal Tax on Capital," Finance and Economics Discussion Series 2011-55, Federal Reserve Boark of Governors 2011.

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.

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

Shaw, Kathryn L., "Life-Cycle Labor Supply with Human Capital Accumulation," *International Economic Review*, 1989, *30* (2), pp. 431–456.

Topel, Robert, "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," Technical Report 3294, NBER 1990.

A   Analytical Derivations

For Online Publication

A.1   Exogenous

The Lagrangian for this specification is

$$
\begin{aligned}
\mathscr{L} = {} & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\
& - \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\varepsilon_2)) \\
& - \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\varepsilon_2)) \\
& + \lambda_t(c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi h_{1,t}^{1+\frac{1}{\sigma_2}} - \beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}})
\end{aligned}
\tag{57}
$$

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

$$
\rho_t = \chi h_{1,t}^{\frac{1}{\sigma_2}}(1 + \lambda_t(1 + \frac{1}{\sigma_2}))
\tag{58}
$$

$$
\rho_{t+1}\theta\varepsilon_2 = \beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}(1 + \lambda_t(1 + \frac{1}{\sigma_2}))
\tag{59}
$$

$$
\rho_t = \theta(1+r)\rho_{t+1}
\tag{60}
$$

$$
\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1}
\tag{61}
$$

and

$$
\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1}.
\tag{62}
$$

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

$$
\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta}
\tag{63}
$$

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

$$
\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1}{\varepsilon_2}\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\sigma_2}}.
\tag{64}
$$

Combining equation 63 and 64 yields

$$
\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1}{\varepsilon_2}\left(\frac{\beta\rho_t}{\rho_{t+1}\theta}\right)\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\sigma_2}}.
\tag{65}
$$

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

$$
\frac{\rho_t\beta}{\varepsilon_2\rho_{t+1}\theta}\left(\frac{h_{2,t+1}}{h_{1,t}}\right)^{\frac{1}{\sigma_2}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})}.
\tag{66}
$$

Combining equation 66 and 65 generates the following expression for labor taxes

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t(1+\frac{1}{\sigma_2})}{1+\lambda_t(1+\frac{1}{\sigma_2})} = 1. \tag{67}$$

A.2 LBD

The Lagrangian for this LBD specification is

$$\begin{aligned}
\mathscr{L} = &\frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\
&- \rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}s_2)) \\
&- \rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}s_2)) \\
&+ \lambda_t(c_{1,t}^{1-\sigma_1}+\beta c_{2,t+1}^{1-\sigma_1}-\chi h_{1,t}^{1+\frac{1}{\sigma_2}} + \frac{\chi\beta h_{2,t+1}^{1+\frac{1}{\sigma_2}}h_{1,t}s_{h1}(t+1)}{s_2} - \beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}})
\end{aligned} \tag{68}$$

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

$$\begin{aligned}
\rho_t(1+h_{2,t+1}s_{h1}(t+1)) = &\chi h_{1,t}^{\frac{1}{\sigma_2}}(1+\lambda_t(1+\frac{1}{\sigma_2})) - \theta\rho_{t+1}h_{2,t+1}s_{h1}(t+1) \\
&+ \lambda_t\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}\beta h_{1,t}\left[\frac{s_{h1}(t+1)^2}{s_2^2} - \frac{s_{h2,h2}(t+1)}{s_2}\right]
\end{aligned} \tag{69}$$

$$\rho_{t+1}\theta s_2 = \beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}\left[1+\lambda_t(1+\frac{1}{\sigma_2})+(1+\frac{1}{\sigma_2})\frac{h_{1,t}s_{h1}(t+1)\lambda_t}{s_2}\right] \tag{70}$$

$$\rho_t = \theta(1+r)\rho_{t+1} \tag{71}$$

$$\rho_t = c_{1,t}^{-\sigma_1}+\lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \tag{72}$$

and

$$\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1}+\beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1}. \tag{73}$$

The first order conditions with respect to capital and consumption are the same in the exogenous (60, 61, and 62) and LBD models (71, 72, and 73). Therefore equation 14 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

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta} \tag{74}$$

Taking the ratio of the agent's first order conditions, equations 19 and 20 and combining with equation 74 yields

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}} = \left(\frac{h_{1,t}}{h_{2,t+1}}\right)^{\frac{1}{\sigma_2}}\left(\frac{\rho_{t+1}\theta s_2}{\beta\rho_t}\right) - \frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)}. \tag{75}$$

Combining equations 75, 69 and 70 the ratio of the optimal taxes on labor is,

$$
\frac{1-\tau_{h,1}}{1-\tau_{h,2}} =
$$

$$
\frac{\left[1+h_{2,t+1}s_{h2}(t+1)\right]\left[1+\lambda_t(1+\frac{h_{1,t}s_{h2}(t+1)}{s_2})(1+\frac{1}{\sigma_2})\right]}{1+\lambda_t(1+\frac{1}{\sigma_2})-\beta h_{2,t+1}^{1+\frac{1}{\sigma_2}}h_{1,t}^{-\frac{1}{\sigma_2}}\left[s_{h2}(t+1)(1+\lambda_t(1+\frac{h_{1,t}s_{h2}(t+1)}{s_2}))(1+\frac{1}{\sigma_2})-h_1\left((\frac{s_{h2}(t+1)}{s_2})2-\frac{s_{h2,h2}(t+1)}{s_2}\right)\right]} \tag{76}
$$

$$
-\frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)}.
$$

A.3 LOD

The Lagrangian for the LOD model is

$$
\begin{aligned}
\mathscr{L} =& \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{(h_{1,t}+n_{1,t})^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta\frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi\frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\
&- \rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}s_2)) \\
&- \rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}s_2)) \\
&+ \lambda_t(c_{1,t}^{1-\sigma_1}+\beta c_{2,t+1}^{1-\sigma_1}-\chi h_{1,t}^{1+\frac{1}{\sigma_2}}-\beta\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}) \\
&+ \eta_t(\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}s_{n1}(t+1)-\chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}})
\end{aligned} \tag{77}
$$

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

$$
\rho_t = \chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}}\left[1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}\right] \tag{78}
$$

$$
\rho_{t+1}\theta s_2 = \beta\chi h_{2,t+1}^{\frac{1}{\sigma_2}}\left[1+\lambda_2\left(1+\frac{1}{\sigma_2}\right)-\eta_t\left(1+\frac{1}{\sigma_2}\right)s_{n1}(t+1)\right] \tag{79}
$$

$$
\rho_t = \theta(1+r)\rho_{t+1} \tag{80}
$$

$$
\rho_t = c_{1,t}^{-\sigma_1}+\lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \tag{81}
$$

$$
\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1}+\beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1} \tag{82}
$$

and

$$
\theta\rho_{t+1}h_{2,t+1}s_{n2}(t+1) = \tag{83}
$$

$$
\frac{\chi(h_{1,t}+n_{1,t})^{\frac{1}{\sigma_2}}\left(\lambda_t h_{1,t}+\eta_t s_2+\sigma_2(h_{1,t}+n_{1,t})(1+\eta_t s_{n2}(t+1))\right)-\beta\chi\eta_t\sigma_2 h_{2,t+1}^{1+\frac{1}{\sigma_2}}(h_{1,t}+n_{1,t})s_{n2,n2}(t+1)}{\sigma_2(h_{1,t}+n_{1,t})} \tag{84}
$$

The first order conditions with respect to capital and consumption are the same in the exogenous (60, 61, and 62) and LOD models (80, 81, and 82). Therefore equation 14 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

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta} \tag{85}$$

Taking the ratio of the agent's first order conditions, equations 27 and 28 and combining with equation 85 yields

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \left(\frac{h_{2,t+1}}{h_{1,t}+n_{1,t}}\right)^{\frac{1}{\sigma_2}}\left(\frac{\beta\rho_t}{\rho_{t+1}\theta s_2}\right). \tag{86}$$

Taking the ratio of equations 78 and 79 yields,

$$\left(\frac{h_{2,t+1}}{h_{1,t}+n_{1,t}}\right)^{\frac{1}{\sigma_2}}\left(\frac{\beta\rho_t}{\rho_{t+1}\theta s_2}\right) = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)}. \tag{87}$$

Combining equations 86 and 87 generates the following expression for the ratio of the optimal labor taxes,

$$\frac{1-\tau_{h,2}}{1-\tau_{h,1}} = \frac{1+\lambda_t\left(1+\frac{h_{1,t}}{\sigma_2(h_{1,t}+n_{1,t})}\right)+\frac{\eta_t s_2}{\sigma_2(h_{1,t}+n_{1,t})}}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\eta_t s_{n1}(t+1)\left(1+\frac{1}{\sigma_2}\right)}. \tag{88}$$

B Competitive Equilibrium

For Online Publication

B.1 LBD Model

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

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

$$\sum_{j=1}^{J} \text{Max}_{c_j, h_j, a_{j+1}} \beta^{j-1} [\prod_{q=0}^{j-1} \Psi_q] u(c_j, h_j) \tag{89}$$

subject to

$$c_j + a_{j+1} = w s_j h_j - \tau_{ss} w s_j h_j, + (1+r)(a_j + Tr) - T^l[w s_j h_j (1 - .5\tau_{ss})] - T^k[r(a_j + Tr)], \tag{90}$$

$$s_{j+1} = S_{\text{LBD}}(\Omega_j, s_j, h_j), \tag{91}$$

for $j < j_r$, and

$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k[r(a_j + Tr)], \tag{92}$$

for $j \geq j_r$. Additionally,

$$c \geq 0, 0 \leq h \leq 1, \tag{93}$$

$$a_j \geq 0, a_1 = 0. \tag{94}$$

2. Prices w and r satisfy

$$r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \tag{95}$$

$$w = (1-\alpha)\left(\frac{K}{N}\right)^{\alpha} \tag{96}$$

3. The social security policies satisfy

$$SS = b \frac{wN}{\sum_{j=1}^{j_r-1} \mu_j} \tag{97}$$

$$\tau_{ss} = \frac{ss \sum_{j=j_r}^{J} \mu_j}{w \sum_{j=1}^{j_r-1} \mu_j} \tag{98}$$

4. Transfers are given by

$$Tr = \sum_{j=1}^{J} \mu_j (1 - \Psi_j) a_{j+1} \tag{99}$$

5. Government budget balance:

$$G = \sum_{j=1}^{J} \mu_j T^k[r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j T^l[ws_j h_j(1 - .5\tau_{ss})] \tag{100}$$

6. Market clearing:

$$K = \sum_{j=1}^{J} \mu_j a_j \tag{101}$$

$$N = \sum_{j=1}^{J} \mu_j s_j h_j \tag{102}$$

$$\sum_{j=1}^{J} \mu_j c_j + \sum_{j=1}^{J} \mu_j a_{j+1} + G = K^\alpha N^{1-\alpha} + (1-\delta)K \tag{103}$$

B.1.1 LOD Model

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

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

$$\sum_{j=1}^{J} \text{Max}_{c_j, h_j, n_j, a_{j+1}} \beta^{j-1} [\prod_{q=0}^{j-1} \Psi_q] u(c_j, h_j, n_j) \tag{104}$$

subject to

$$c_j + a_{j+1} = ws_j h_j - \tau_{ss} ws_j h_j, + (1+r)(a_j + Tr) - T^l[ws_j h_j(1 - .5\tau_{ss})] - T^k[r(a_j + Tr)], \tag{105}$$

$$s_{j+1} = S_{\text{LOD}}(\Omega_j, n_j, h_j), \tag{106}$$

for $j < j_r$, and

$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k[r(a_j + Tr)], \tag{107}$$

for $j \geq j_r$. Additionally,

$$c \geq 0, 0 \leq h \leq 1, \tag{108}$$

$$a_j \geq 0, a_1 = 0. \tag{109}$$

2. Prices w and r satisfy

$$r = \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta \tag{110}$$

$$w = (1-\alpha)\left(\frac{K}{N}\right)^\alpha \tag{111}$$

3. The social security policies satisfy

$$SS = b \frac{wN}{\sum_{j=1}^{j_r-1} \mu_j} \tag{112}$$

$$\tau_{ss} = \frac{ss \sum_{j=j_r}^{J} \mu_j}{w \sum_{j=1}^{j_r-1} \mu_j} \tag{113}$$

4. Transfers are given by

$$Tr = \sum_{j=1}^{J} \mu_j (1 - \Psi_j) a_{j+1} \tag{114}$$

5. Government budget balance:

$$G = \sum_{j=1}^{J} \mu_j T^k [r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j T^l [ws_j h_j (1 - .5\tau_{ss})] \tag{115}$$

6. Market clearing:

$$K = \sum_{j=1}^{J} \mu_j a_j \tag{116}$$

$$N = \sum_{j=1}^{J} \mu_j s_j h_j \tag{117}$$

$$\sum_{j=1}^{J} \mu_j c_j + \sum_{j=1}^{J} \mu_j a_{j+1} + G = K^\alpha N^{1-\alpha} + (1 - \delta)K \tag{118}$$