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

The Effect of Endogenous Human Capital Accumulation on Optimal Taxation

William B. Peterman

2012-003

The Effect of Endogenous Human Capital Accumulation on Optimal Taxation

William B Peterman*

October 13, 2015

Abstract

This paper considers the impact of learning-by-doing on optimal tax policy in a general equilibrium heterogenous agent life-cycle model. Analytically, it identifies two main channels by which learning-by-doing alters the optimal tax policy. First, learning-by-doing creates a motive for the government to use age-dependent labor income taxes. If the government cannot condition taxes on age, then a capital tax or progressive/regressive labor income tax can be used in order to mimic age-dependent taxes. Second, a progressive/regressive labor income tax is potentially more distortionary in a model with learning-by-doing since the distortion is propagated through the additional intertemporal link between current labor and future human capital. Quantitatively, I find that both of these channels are important for the optimal tax policy. Adding learning-by-doing leads to a notably flatter optimal labor income tax due to the second channel. Moreover, including learning-by-doing causes an increase in the optimal capital tax due to the first channel. I find that when solving for the optimal tax policy in the learning-by-doing model, the welfare consequences of not accounting for endogenous human capital accumulation are equivalent to around one percent of expected lifetime consumption, a majority of which are due to adopting too progressive of a tax policy.

JEL: E24, E62, H21.

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

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

1 Introduction

Previous research documents that variation in consumption and labor, in part due to fluctuations in age-specific human capital, can have important implications for optimal taxation.[1] In particular, variation in consumption and labor may cause the optimal labor income tax to be age-dependent. Moreover, if age-dependent taxes are disallowed then it tends to be optimal to use either a non-flat labor income tax (i.e. a progressive or regressive tax) or a non-zero tax on capital to mimic these age-dependent taxes.[2] Therefore, variation in human capital over the lifetime can have implications for two fundamental questions in the optimal taxation literature.[3] First, should the income tax be progressive? Second, should capital be taxed? However, previous research that examines both of these questions in a life cycle model tends to assume that human capital is accumulated exogenously. In this paper, I examine the effect of endogenous human capital accumulation on both the optimal shape of the labor income tax policy and the optimal taxation of capital.

In particular, I determine the effect both analytically and quantitatively on optimal tax policy of including endogenous age-specific human capital accumulation through learning-by-doing (LBD) in which an agent's future human capital is affected by the hours worked today.[4] Despite being commonly used in life-cycle models, and empirical evidence supporting the relationship between current work and future productivity, previous research has not simultaneously examined the effect of LBD on both tax questions.[5] Therefore, this paper determines the effect of LBD on both parts of the optimal tax policy. Overall, this paper finds that endogenizing human capital accumulation has significant qualitative and quantitative implications for both parts of the optimal tax policy, operating through two important channels.

I begin by analytically demonstrating these two channels in a simple model. First, I demonstrate that adding LBD changes the relative incentives to work over an agent's lifetime. In the LBD setting there are two benefits to working. Working provides an agent with a wage ("wage benefit") and an increase in future human capital ("human capital benefit"). The human capital benefit only exists in the LBD model and not

[1] I define age-specific human capital as human capital that is accumulated after an agent begins working.

[2] Examples of this research includes Atkeson et al. (1999), Erosa and Gervais (2002), Garriga (2001), and Gervais (2012).

[3] See Diamond and Saez (2011) and Mirrlees et al. (2010) for a discussion of the importance of these questions and a general summary of previous findings.

[4] An alternative form of endogenous human capital accumulation that is sometimes used is learning-*or*-doing (LOD). 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. This paper ignores this form of human capital accumulation because Mulligan (1995) finds that once individuals start working they spend less than 7 percent of their time endowment in formal training (Peterman (2014) finds that there are much smaller effects on the optimal capital tax when endogenous human capital is added with LOD). Therefore, this form of human capital accumulation is more relevant to pre-work skill formulation than age-specific human capital accumulation. Moreover, Jacobs and Bovenberg (2010) finds that incorporating endogenous pre-work skill accumulation has similar effects on the optimal tax policy as LBD.

[5] Examples of studies which show that past hours worked and length of current job tenure impact current wages include Topel (1990), Cossa et al. (1999), and Altuğ and Miller (1998). Moreover, examples of life cycle studies that include LBD are Hansen and İmrohoroğlu (2009), Imai and Keane (2004), and Chang et al. (2002).

in the exogenous model leading agents to be less responsive to temporary changes in wages in the LBD model. Moreover, 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. Optimally, the social planner would tax labor income from agents when they are young and supply labor less elastically at a relatively higher rate than when they are older. If the social planner cannot use age-dependent taxes, then a tax on capital or a progressive/regressive labor income tax can be optimal in order to mimic the age-dependent taxes.[6] I refer to this first channel as the elasticity channel.

Second, I demonstrate that including LBD enhances the distortions from a non-flat labor income tax. In both the exogenous and LBD model a progressive tax distorts an agent's labor decisions because a it causes the marginal after-tax wage benefit to decline as labor income increases. However, the distortion is magnified in the LBD model because it is propagated through the additional intertemporal link between current labor and future human capital. In particular, in the LBD model a progressive labor income tax also leads the marginal after-tax human capital benefit to decline as future labor income increases. Thus a flatter labor income tax policy is optimal with LBD. I refer to this second channel as the intertemporal distortion channel.

Next, I quantitatively assess the impact of LBD on optimal tax policy in a rigorous general equilibrium overlapping generations model (OLG) that includes heterogeneity due to idiosyncratic shocks to labor productivity. To explore the effect of LBD, I solve for the optimal tax policies in two different cases – first in a model with no LBD (the exogenous model) and then again in a model with LBD (the LBD model). In the LBD model I find that the optimal tax policy is a 36 percent flat tax on capital income, a 22.3 percent tax on labor income with a fixed deduction of $10,901, and a lump-sum transfer of $365. In contrast in the exogenous model I find that the optimal tax policy is a 30 percent tax on capital, a 32.5 percent tax on labor income with a fixed deduction of $6,218, and a lump-sum transfer of $3,683. Thus, adding LBD has considerable quantitative implications on both tax questions. In particular, adding LBD reduces the optimal progressivity of the labor income tax policy and raises the optimal capital tax by 6.0 percentage points. Through a series of counterfactual experiments, I confirm that the intertemporal distortion is responsible for the flatter optimal tax policy and that the elasticity channel is responsible for the increase in the optimal capital tax.

Overall, I find that the welfare consequences of not accounting for the effects of LBD when determining the optimal tax policy are notable. In particular, I find that in the LBD model implementing the optimal tax

[6] In contrast, Garriga (2001) demonstrates that, in a specific set of models with exogenous human capital accumulation, it is not optimal to condition labor taxes on age nor to tax capital. Moreover, a host of work demonstrates a similar set of results in a two-generation model with a single cohort. Two examples of these works include Atkinson and Stiglitz (1976), and Deaton (1979).

policy from the exogenous model – which includes a more progressive labor tax and lower capital tax – as opposed to the actual optimal tax policy – which includes a flatter tax on labor and larger tax on capital – results in a welfare reduction equivalent to between 0.7 and 1.2 percent of expected lifetime consumption depending on the utility function. I find that a majority of these welfare consequences are due to the sub-optimal level of progressivity as opposed to the sub-optimal capital tax. Thus for welfare purposes, the change in the optimal progressivity from adding LBD is more significant than the change in the optimal capital tax. Furthermore, I find that the change in the optimal tax policy from adding LBD is not sensitive to either the utility function or the parameter values used to calibrate the LBD skill accumulation function. Taken as a whole, these results demonstrate that including LBD has quantitatively important effects on both the magnitude and the shapes of the optimal taxes on capital and labor.

This paper contributes to the general class of literature that explores the optimal tax policy when the set of available tax instruments are restricted. Correia (1996), Armenter and Albanesi (2009), and Jones et al. (1997), demonstrate that certain tax instruments, that otherwise would not be optimal, may become optimal when the government's set of tax instruments are restricted. This paper combines two related strands of the literature within this class of research that quantitatively determine the optimal capital tax and optimal progressivity of the income tax when the government is restricted from using age-dependent taxes.[7]

The first strand simultaneously examines both tax questions in a calibrated life-cycle model but includes human capital accumulation exogenously.[8] Atkeson et al. (1999), Erosa and Gervais (2002), Garriga (2001), and Gervais (2012) determine that generally it is optimal to condition labor income taxes on age in a life-cycle model.[9] Moreover, they demonstrate that if age-dependent taxes are not allowed, then it is possible to mimic the optimal tax policy with either a non-flat labor income tax (i.e. a progressive or regressive tax) or a non-zero tax on capital. Conesa et al. (2009), Peterman (2013), and Gervais (2012) demonstrate quantitatively in a life-cycle model that the inability to condition taxes on age can be a strong motive for a positive capital tax and a progressive/regressive tax on labor income. In particular, Conesa et al. (2009), henceforth CKK, find in a life-cycle model that is similar to my exogenous model, the optimal tax policy includes both a progressive labor income tax and a sizeable tax on capital. Although the authors find that a primary reason for the large optimal tax on capital is to mimic an age-dependent tax, they find that the primary reason for the optimal progressive labor income tax is to provide ex-ante insurance for idiosyncratic

[7]For a discussion of the optimal age-dependent tax policy in the new dynamic public finance framework see Farhi and Werning (2013), Kremer (2002), and Weinzierl (2011).

[8]There is a strand of literature that examines these questions in an infinitely lived agent model as opposed to a life-cycle model. See Diamond and Saez (2011) for a review of this literature.

[9]Atkeson et al. (1999) demonstrate a related result. They show conditions under which the optimal capital tax is zero if age-dependent taxes on labor income are allowed.

shocks to labor productivity and not to mimic age-dependent taxes.[10] In addition, Gervais (2012) finds that in some cases, even with a large tax on capital, a mild amount of progressivity in the labor income tax is optimal in order to mimic an age-dependent tax policy. Although these studies examining both the optimal tax on capital and the optimal shape of the labor income in a life-cycle model, they include human capital accumulation exogenously ignoring any affects of endogenous human capital accumulation on the optimal tax policy. In contrast, this paper both analytically and quantitatively assesses the effects of including endogenous human capital accumulation on the optimal tax policy.

This paper is related to a second strand of the literature that includes LBD but only focus on its effect on one of the tax questions.[11] For example, focusing on optimal capital taxation, Chen et al. (2011) finds that, in an infinitely lived agent model with labor search, including endogenous human capital accumulation causes the optimal capital tax to increase because a higher capital tax unravels the labor market frictions in their model.[12] Since Chen et al. (2011) only examine the effect of endogenous human capital on the optimal capital tax in an infinitely lived agent model, they are unable to assess whether LBD affects the motive to use age-dependent taxes, or whether LBD affects the efficiency of a progressive labor income tax versus a tax on capital to mimic age-dependent taxes. Focusing on the effect of LBD on the optimal amount of progressivity, both Best and Kleven (2012) and Krause (2009) demonstrate that a flatter income tax, as opposed to a progressive tax, is optimal in a two-generation model with LBD so as to not discourage human capital accumulation. However, Best and Kleven (2012) and Krause (2009) do not incorporate savings so they do not determine the effect of LBD on the optimal capital tax. Since both a tax on capital and a progressive/regressive labor income tax can be used to mimic an age-dependent tax policy, it is important to examine the effect of LBD on both questions simultaneously in a life-cycle model. This paper combines both strands of the literature and determines that although including LBD in a life cycle model changes the answer to both questions, the change in the optimal progressivity is the dominate effect for welfare purposes.

[10]For a discussion of the channel leading to the progressive tax policy see Mirrlees (1971), Stiglitz (1982), Mirrlees (1974), and Varian (1980). Peterman (2013) demonstrates that an additional motive for a positive tax on capital is that the government is unable to distinguish between accidental bequests and ordinary capital income. Further work, such as Karabarbounis (2012) and Peterman (2012), demonstrate that incorporating endogenous fluctuations in labor supply on the extensive margin can enhance this motive for the government to use a capital tax to mimic age-dependent taxes on labor income. In contrast, Cespedes and Kuklik (2013) find that when a non-linear mapping between hours and wages exists then hours tend to become more persistent and the optimal capital tax falls significantly, however is still positive.

[11]This paper primarily focuses on the tax studies that use a Ramsey approach. There is a parallel strand of the literature in dynamic public finance that also that examines optimal taxation with endogenous human capital. Examples of this strand that examines a model with endogenous human capital accumulation is Golosov et al. (2003), Stantcheva (2015a), and Stantcheva (2015b).

[12]The authors include endogenously human capital accumulation through both LBD and also training. The labor market frictions in Chen et al. (2011) cause a lower level of employment in their economy. A capital tax causes the wage discount to increase, thus causing firms to post more vacancies which in turn causes an increase in worker participation. A number of studies examine the optimal tax policy in an infinitely lived agent model with other forms of endogenous human capital accumulation. Examples include Jones et al. (1997), Judd (1999), and Reis (2007).

This paper is organized as follows: Section 2 examines an analytically tractable version of the model to demonstrate the two channels by which LBD alters the optimal tax policy. Section 3 describes the full model used in the quantitative exercise (see Appendix B for the competitive equilibrium). The calibration and functional forms are discussed in section 4. Section 5 describes the computational experiment, and section 6 presents the results. Section 7 examines the sensitivity of the results, while section 8 concludes.

2 Analytical Model

In this section, I demonstrate the two channels by which adding LBD alters the optimal tax policy. First, I show that adding LBD introduces new channels that cause the government to want to condition labor income taxes on age. If the government cannot use age-dependent taxes, then a tax on capital or a progressive/regressive labor income tax can be optimal in order to mimic these age-dependent taxes.[13] However, I show that introducing LBD enhances the distortions associated with the progressive tax and therefore makes it less likely that a progressive/regressive labor income tax would be optimal to mimic the age-dependent taxes.

I derive these analytical results in a tractable two-period version of the computational model that nests both cases when human capital is accumulated exogenously or through LBD. For tractability purposes, the features I abstract from include retirement, population growth, idiosyncratic labor productivity shocks, and conditional survivability. Additionally, I assume that the marginal products of capital and labor are constant so factor prices do not vary.[14] Since changes to the tax system do not affect the pre-tax wage or rate of return, I am able to focus on the life-cycle elements of the model. Also, because I exclude idiosyncratic labor productivity shocks, there is no within-cohort heterogeneity in the analytically tractable model. Therefore, without this within-cohort heterogeneity the social planner focuses only on efficiency and ignores the tradeoff between equity and efficiency. All of these assumptions are relaxed in the computational model.

[13]The link between age-dependent taxes and these alternative tax instruments are explored in a model with exogenous human capital in Garriga (2001), Erosa and Gervais (2002), and Gervais (2012).

[14]Since the factor prices do not vary, I suppress their time subscripts in this section.

2.1 Elasticity Channel

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

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

where $U()$ is a utility function that is increasing with respect to both arguments, β is the discount factor, $c_{j,t}$ 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 (implying that 1-$h_{j,t}$ is the percent of the time endowment consumed as leisure). Age-specific human capital is normalized to unity when the agent is young. At age two, age-specific human capital is $s_2(h_{1,t})$. In the case of LBD, $s_2(h_{1,t})$ is a function of hours worked in the previous period and I assume that $\frac{\partial s_2(h_{1,t})}{\partial h_{1,t}} > 0, \frac{\partial s_2^2(h_{1,t})}{\partial h_{1,t}^2} < 0$. This first assumption, $\frac{\partial s_2(h_{1,t})}{\partial h_{1,t}} > 0$, implies that an agent working more when they are young will increase their skills when they are old (the human capital benefit). In the case of exogenous human capital accumulation, s_2 is exogenously predetermined and thus is no longer dependent on hours worked ($\frac{\partial s_2(h_{1,t})}{\partial h_{1,t}} = 0, \frac{\partial s_2^2(h_{1,t})}{\partial h_{1,t}^2} = 0$).[15] Thus, the human capital benefit only exists with LBD and not in the exogenous model. The agent chooses consumption and hours worked in order to maximize equation 1 subject to the following standard budget constraints

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

and

$$c_{2,t+1} = (1+r(1-\tau_k))a_{1,t} + (1-\tau_{h,2})s_2(h_{1,t})h_{2,t+1}w, \tag{3}$$

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 begin by assuming that the tax rate on labor income is flat but can be conditioned on age. Moreover, I assume that the tax rate on capital income is flat and cannot be conditioned on age.[16] The agent's first-order conditions are

$$\frac{U_{h1}(t)}{U_{c1}(t)} = w(1-\tau_{h,1}) + \frac{w(1-\tau_{h,2})h_{2,t+1}s_{h1}(t+1)}{(1+r(1-\tau_k))}, \tag{4}$$

[15] Generally, I solve for results with both forms of human capital accumulation nested. Thus, I continue to represent s_2 as a function of hours worked unless specifically describing the exogenous case.

[16] 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, restricting capital tax policy to not be age-dependent is not binding.

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = ws_2(h_{1,t})(1-\tau_{h,2}), \tag{5}$$

and

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1+r(1-\tau_k)), \tag{6}$$

where $U_{c1}(t) \equiv \frac{\partial U(c_{1,t},1-h_{1,t})}{\partial c_{1,t}}$. Throughout the analytical section I highlight in red the portions of the expressions that are specific to LBD and do not exist when human capital is accumulated exogenously. In particular, equation 4 has an additional term in the case of LBD since the agent receives the human capital benefit at age 2 from working at age 1.

Using the primal approach to solve for the optimal tax policy in this model and assuming that the utility function is separable in consumption and labor ($U_{ch} = 0$), equation (see appendix A for the problem's formulation and further details) represents the relationship between the optimal age-dependent taxes,

$$\frac{(1-\tau_{h1})}{(1-\tau_{h2})} = \frac{(1-\lambda) + \frac{\lambda h_2 U_{h2h2}}{U_{h2}} - \frac{\lambda s_{h2}}{s_2 U_{h2}}\left(h_2 h_1 U_{h2h2} - h_1 U_{h2}\right)\left(1 + \frac{h_2 s_{h2}}{1+r(1-\tau_k)}\right)}{(1-\lambda) + \frac{\lambda h_1 U_{h1h1}}{U_{h1}} + \frac{h_2 \beta \lambda U_{h2}}{s_2^2 U_{h1}}\left(s_2(s_{h2} + h_1 s_{h2h2}) - h_1 s_{h2}^2\right)} - \frac{h_2 s_{h2}}{1+r(1-\tau_k)}. \tag{7}$$

Equation 7 demonstrates that the optimal age-dependent taxes will tend to be different with LBD and exogenous human capital accumulation. Thus incorporating human capital accumulation with LBD creates an additional motive for age-dependent taxes.

Previous work (see Atkeson et al. (1999), Erosa and Gervais (2002), Garriga (2001), and Peterman (2014)) demonstrates that if the social planner wants to condition labor income taxes on age but is disallowed from using these age-dependent taxes then a non-zero capital tax will be optimal in this type of model. For example, Garriga (2001) demonstrates that if the utility function is separable in consumption and labor and also homothetic in each individual argument then, in this type of simple model with exogenous human capital accumulation, the social planner does not want to condition labor income taxes on age and as such does not want to tax capital regardless of whether they can use age-dependent taxes.[17] However, even if the utility function meets these conditions incorporating LBD creates a new motive for the social planner to conditional labor income taxes on age and thus alters the Garriga (2001) result.[18]

In order to determine the direction of the effect on optimal taxes from adding LBD, I make three sufficient assumptions: (i) $h_2 U_{h2h2} < U_{h2}$, (ii) $h_2 s_{h2} = \kappa s_{2,t+1}$, and (iii) $h_2 s_{h2h2} = (\kappa - 1)s_{2,t+1}$, where κ is some

[17] In particular, if the utility function is such that $\frac{hU_{hh}}{U_h}$ is equal to a constant. A utility function that is homothetic in labor would imply that this ratio is constant.

[18] Under the Garriga (2001) utility function the black terms on the right hand side of both the denominator and numerator will simplify to the same constant. Thus without LBD the ratio equals one. However, once LBD is included the red terms are introduced to the expression for the optimal tax policy and with these additional terms the ratio does not simplify to one.

arbitrary constant.[19] Under these assumptions, Equation 7 implies that the optimal tax policy includes a higher labor tax on agents when they are young.

The intuition for why adding LBD will tend to increase the optimal relative tax on young labor income comes from examining the Frisch elasticity.[20] In particular the Frisch elasticity is,

$$\text{Frisch} = \frac{U_h U_{cc}}{h\left[U_{ch}^2 - U_{cc}U_{hh} + \frac{h'w'}{rws}\left(s'_h(U_{ch}^2 - U_{cc}U_{hh}) - U_{cc}U_h s'_{hh}\right)\right]} \tag{8}$$

where the next period is denoted with a prime for notational convenience ($s' = s_{j+1,t+1}$).[21][22] The additional expressions from LBD cause the denominator to increase, thus holding hours and consumption constant the Frisch elasticity is lower when LBD is included. Intuitively, the inclusion of the human capital benefit makes workers less responsive to a one-period change in wages since the wage benefit is now only part of their total compensation for working when LBD is included. Moreover, 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. In the stylized case where agents only live for two periods, the effect of the human capital benefit would only exist for an agent when they are young. Therefore, adding LBD causes an agent to supply labor relatively less elastically when they are old compared to when they are old. This shift in relative elasticities creates an incentive for the social planner to tax the labor income of younger agents at a relatively higher rate. I use the term "elasticity channel" to describe the effect on optimal tax policy caused by a change in the Frisch elasticity from including LBD.[23]

In such a case when the social planner would like to use age-dependent taxes but is disallowed, then a tax on capital can mimic such a tax. The intertemporal Euluer equation demonstrates why the tax on capital

[19] These are not strong assumption since the standard functional forms and calibration parameters I choose for the utility function and the LBD process in the computational model all adhere to these assumptions.

[20] These results can be derived under more general conditions than those needed to sign the effect of LBD on the optimal tax policy. In particular, the only assumption necessary is that the utility function is separable in consumption and labor.

[21] 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}.

[22] I provide the expression for the Frisch elasticity in a more general model where agents live for more than two periods in order to see how the effect of LBD varies over the life cycle. However, this expression also maps into this two period model. In particular, in a two period model the additional expression from LBD only exists for young agents and not old agents.

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

mimics the age-dependent labor income tax,

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

In particular, a positive (negative) capital tax induces a wedge on the marginal rate of substitution that is similar to a relatively larger tax on young (old) labor income.[24] Thus if age-dependent taxes are disallowed then adding LBD will cause a larger tax on capital to be optimal in order to mimic a relatively larger tax on labor income when agents are young.[25][26]

2.2 Distortions from Progressive/Regressive Tax

Next, I examine why a progressive/regressive labor income tax can mimic an age-dependent tax and how the relative efficiency of a progressive/regressive labor income tax versus a tax on capital changes when LBD is introduced.[27]

With a progressive/regressive tax on labor income the average tax rate is no longer a function of age; instead it is a function of labor income $T(h_{i,t}ws_i)$. This change in the tax function leads to a change in the agent's constraints (equations 2 and 3)

$$c_{1,t} + a_{1,t} = (1 - T(h_{1,t}w))h_{1,t}w \tag{10}$$

and

$$c_{2,t+1} = (1+r(1-\tau_k))a_{1,t} + (1 - T(s_2(h_{1,t})h_{2,t+1}w))s_2(h_{1,t})h_{2,t+1}w, \tag{11}$$

where T is the average tax rate on labor income which is assumed to be increasing and concave in labor income. The agent's first-order conditions with this new tax function are

$$\frac{U_{h1}(t)}{U_{c1}(t)} = w\left(1 - T(h_{1,t}w) - h_{1,t}T_{h1}(t)w + \beta\frac{U_{c2}(t+1)}{U_{c1}(t)}s_{h1}(t+1)h_{2,t+1}\left(1 - T(h_{2,t+1}ws_2(h_{1,t})) - s_2(h_{1,t})T_{s2}(t+1)wh_{2,t+1}\right)\right), \tag{12}$$

[24] Examining the implications of LBD on this relationship, the additional term is positive. Therefore, holding all else equal, the tax on young labor income would need to be relatively higher in order to induce the same wedge on the marginal rate of substitution in the LBD model.

[25] Although τ_k may act as a substitute for age-dependent taxees, τ_k is not a redundant instrument because it also distorts the intertemporal marginal rate of substitution between consumption.

[26] In addition, a progressive/regressive labor income tax can also be used to mimic age-dependent labor income taxes which is discussed in the next section.

[27] When allowing the government to use a progressive/regressive labor income tax the same primal approach does not yield an analytical solution for the optimal tax policy because labor choices affect the average labor tax rate. Therefore, prices cannot be removed from the intertemporal budget constraint using the first-order conditions in order to create the implementability constraint. Thus, I am unable to solve for the optimal policy.

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = ws_2(h_{1,t})\Big(1 - T(h_{2,t+1}ws_2(h_{1,t})) - h_{2,t+1}T_{h2}(t+1)\Big), \tag{13}$$

and

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1 + r(1-\tau_k)), \tag{14}$$

where $T_{h1}(t) \equiv \frac{\partial T(h_{1,t}w)}{\partial h_{1,t}w}$, or the marginal tax rate. The first-order conditions with respect to labor (equations 12 and 13) change compared to the case of flat labor income taxes (equations 4 and 5). Including a progressive/regressive tax changes the first order condition because if an agent changes the hours he works then his marginal labor tax rate also changes.

Similar to a capital tax, a progressive/regressive tax can mimic age-dependent labor taxes. Examining the Euler equation with a progressive/regressive tax (for expositional convenience I use the simpler case of exogenous human capital accumulation),

$$s_2\frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1 + r(1-\tau_k))\left(\frac{1 - T(wh_{1,t}) - h_{1,t}wT_{h1}(t)}{1 - T(wh_{2,t}s_2) - h_{2,t}wT_{h2}(t+1)}\right). \tag{15}$$

if the social planner wants to condition taxes on age but is disallowed then a progressive/regressive labor income tax can create a similar wedge in the marginal rate of substitution. In particular, if labor income increases over an agent's lifetime then a regressive labor income tax would create a similar wedge as an age-dependent flat tax with a relatively higher rate on income earned at young ages.

Although both a progressive/regressive tax on labor income or a non-zero tax on capital can create a wedge on the marginal rate of substitution, there are several reasons why a tax on capital may be more desirable, especially in a less parsimonious model. First, if the social planner wants the implicit labor income tax to monotonically decrease with age, a positive tax on capital may be ideal since it mimics a monotonically decreasing labor tax by age. In contrast, a progressive tax implicitly taxes labor income at a higher rate at ages when an agent earns more. Therefore, if labor income is not monotonically increasing or decreasing over a working agent's life then there is no way for a progressive/regressive tax policy to mimic a monotonically decreasing age-dependent tax policy.

The second reason a tax on capital may be preferable to a progressive/regressive labor income tax is a capital tax imposes a wedge on the marginal rate of substitution that is independent of the agent's labor choice. In contrast, the size of the wedge from a progressive/regressive labor income tax will depend on the amount of labor income. In a less parsimonious model that includes within-cohort heterogeneity, agents of the same age may have different labor income, making it even more difficult for the social planner to use a progressive/regressive labor income tax to mimic an age-dependent tax. In contrast, the wedge from a capital income tax will be a function of age but not labor supply so it will be the same for all agents of the same age regardless if there is within-cohort heterogeneity in labor incomes.

The general desirability of a progressive/regressive labor income tax may be weakened with LBD. Adding LBD alters an agent's tradeoffs because it introduces an additional intertemporal link between current labor and future productivity (see Equation 12). Since productivity is linked to the level of income, the distortions from the progressive tax are magnified through this channel. In the LBD model, a progressive tax policy reduces an agent's incentives to work since the progressive tax implies that the marginal human capital benefit declines as future labor income increases. Since the additional intertemporal link in the LBD model magnifies the distortions from a progressive/regressive tax, I refer to this second channel as the intertemporal distortion channel.

3 Computational Model

Next, I determine the quantitative effect of adding LBD on optimal tax policy in a rigorous version of the model that includes other channels that affect the optimal capital tax and progressivity of the labor income tax. One notable channel arises from the inclusion of within-cohort heterogeneity which causes the social planner to consider not only efficiency but to weigh the tradeoff between efficiency and equity. In particular, the social planner may use a progressive labor income tax to redistribute and provide insurance against labor income risk. I solve for the optimal tax policy in separate versions of the computational model with exogenous human capital accumulation and LBD. The exogenous model is adapted from CKK.[28]

3.1 Demographics

Time is assumed to be discrete, and the model period is equal to one year. Agents enter the economy when they start working, at age 20, and live for up to J years. The economy is populated with J overlapping generations of ages $20, 21, ..., J+20$. The size of each new cohort entering the economy grows at a constant rate n. Lifetime length is uncertain with mortality risk varying over the lifetime. Conditional on being alive at age j, Ψ_j is the probability of an agent living to age $j+1$. Since agents are not certain how long they will live, they may die while still holding assets. If an agent dies with assets, the assets are confiscated by the government and distributed equally to all the living agents as accidental bequests (beq_t).[29] All agents are

[28]Although I use their alternative utility function which is separable, I find qualitatively similar results with their benchmark utility function which is non-separable. I choose to use the separable utility function to follow the analytically tractable model.

[29]Including this formulation increases the optimal tax on capital for two reasons. First, Peterman (2013) demonstrates that with accidental bequests the optimal capital tax increases when the government cannot distinguish between gains on accidental bequests

required to retire at an exogenously set age j_r.

3.2 Individual

An individual is endowed with one unit of productive time per period that he divides between labor ($h_{i,j}$) and leisure ($1 - -h_{i,j}$). An agent earns $w\omega_{i,j}h_{i,j}$ for their labor where $\omega_{i,j}$ is the idiosyncratic productivity of agent i at age j. Agents split their income between saving with a one-period risk-free asset ($a_{i,j}$) and consumption ($c_{i,j}$). Agents choose labor, savings, and consumption in order to maximize their lifetime utility

$$u(c_{i,j},h_{i,j}) + \sum_{s=20}^{J-j-20} \beta^s \prod_{q=1}^{s} (\Psi_q) u(c_{i,s+1},h_{i,s+1}). \tag{16}$$

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 factor is $\beta\Psi_j$.

The log of an agent's idiosyncratic productivity $\omega_{i,j}$ in the exogenous model can be split into four additively separable components,

$$\log \omega_{i,j} = \varepsilon_j + \alpha_i + \nu_t + \theta_t. \tag{17}$$

and in the LBD model,

$$\log \omega_{i,j} = s_{i,j} + \alpha_i + \nu_t + \theta_t. \tag{18}$$

In this specification, based on the estimates in Kaplan (2012) from the Panel Study of Income Dynamics (PSID), ε_j or $s_{i,j}$ governs age-specific human capital. Moreover, $\alpha \sim NID(0,\sigma_\alpha^2)$ is an individual-specific fixed effect (or ability) that is realized when an agent enters the economy and stays fixed over the life-cycle, $\theta_t \sim NID(0,\sigma_\varepsilon^2)$ is a transitory shock to productivity received every period, and ν_t is a persistent shock, which follows a first-order autoregressive process:

$$\nu_t = \rho\nu_{t-1} + \psi_t \text{ with } \psi_t \sim NID(0,\sigma_\nu^2) \text{ and } \nu_2 0 = 0.^{30} \tag{19}$$

In the exogenous model an agent's age-specific human capital ε_j is exogenously determined. In the LBD model, an agent's age-specific human capital is determined by $s_{i,j} = S_{\text{LBD}}(\Omega_{j-1}, s_{i,j-1}, h_{i,j-1})$ where $\{\Omega_j\}_{j=20}^{j_r-1}$ is a sequence of calibration parameters that are set so that in the LBD model, under the baseline-fitted U.S.

and ordinary capital income. Second, if non-accidental bequests are included instead of accidental bequests then Fuster et al. (2007) demonstrates that the model is more like an infinitely lived agent model where the optimal capital tax tends to be smaller.

[30]Setting $\nu_1 = 0$ implies that this shock equals 1 when agents enter the model since $\exp^\nu$ is how the shock enters the agents idiosyncratic productivity. Moreover this setting implies that, consistent with the data, the variance in the idiosyncratic shocks grow with age.

tax policy, the agents' choices result in the same average age-specific human capital age profile in the LBD and exogenous models.

3.3 Market structure

The markets are incomplete and agents cannot fully insure against the idiosyncratic labor productivity and mortality risks by trading state-contingent assets. They can, however, partially self-insure against these risks by accumulating precautionary asset holdings, a. These savings are also used to fund consumption after retirement. The stock of assets earns a market return r_t. I assume that households enter the economy with no assets and are not allowed to borrow against future income, so that $a_{i,20} = 0$ and $a_{i,j} \geq 0$ for all i and j.

3.4 Firm

Firms are perfectly competitive with constant returns to scale production technology. Aggregate technology is represented by a Cobb-Douglas production function. Unlike the analytically tractable model, I do not assume a linear production function in the computational model, so prices are determined endogenously. The aggregate resource constraint is,

$$C_t + K_{t+1} - (1-\delta)K_t + G_t \leq K_t^{\zeta} N_t^{1-\zeta}, \tag{20}$$

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

3.5 Government Policy

The government runs a pay-as-you-go (PAYGO) social security system. In this reduced-form social security program, the government pays SS_t to all individuals that are retired, independent of the individual agent's earnings history. To finance the system, labor income is taxed at the flat rate τ_{ss} up to a maximum labor income level $\bar{y}$, as in the actual system. The payroll tax rate, $\tau_{ss,t}$, funds the balanced budget program. The social security system is not considered part of the tax policy that the government optimizes. I include this simplified social security program because excluding this program would cause an agent to overemphasize savings since all of their post-retirement consumption would need to be financed with private savings.[31]

[31]Peterman (2013) demonstrates that excluding social security can have notable effects on the optimal tax policy.

In addition to running the Social Security system, the government has two fiscal instruments to finance its consumption, G_t, which is done in an unproductive sector.[32] First, the government taxes capital income, $y_k \equiv r_t(a+beq_t)$, 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. Let $py^l_{i,j}$ be the pre-tax labor income which is equal to $py^l_{i,j} \equiv w_t s_{i,j} h_{i,j}$. Since part of these contributions are not taxable, the agent's taxable labor income is $y_{i,j} \equiv py^l_{i,j} - .5\tau_{ss}\min\{py^l_{i,j}, \bar{y}\}$, 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 and cannot be taxed directly. Second, I assume the tax rates cannot be age-dependent. Third, both of the taxes are solely functions of the individual's relevant taxable income in the current period.

4 Calibration and Functional Forms

Prior to solving the models, it is necessary to choose functional forms and calibrate the models' 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 (Simulated Method of Moments).[33] These values are in Table 2.

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 LBD and exogenous models. To assure that both 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 two models implying that the values of these parameters vary between the two models.

4.1 Demographics

Agents enter the model at age 20 and are exogenously forced to retire at 66. 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 adjust the size of each cohort's share of the population to account for

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

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

Table 1: **Calibration Parameters**

Parameter	Value	Target
Demographics		
Retire Age: j_r	66	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		
ζ	.36	Data
δ	8.33%	$\frac{I}{Y} = 25.5\%$
A	1	Normalization
Productivity Parameters:		
Persistence Shock: σ_ν^2	0.017	Kaplan (2012)
Persistence: ρ	0.958	Kaplan (2012)
Permanent Shock: σ_α^2	0.065	Kaplan (2012)
Transitory Shock: σ_ε^2	0.081	Kaplan (2012)
Government Parameters:		
Payroll Tax: τ_{ss}	0.124	CKK
Υ_0	.258	Gouveia and Strauss (1994)
Υ_1	.768	Gouveia and Strauss (1994)
LBD Parameters:		
Φ_1	.407	Chang et al. (2002)
Φ_2	.326	Chang et al. (2002)

a population growth rate of 1.1 percent.

Table 2: **Calibration Parameters**

Parameter	Exog.	LBD	Target
Calibration Parameters			
Discount Factor: β	0.994	0.992	$K/Y = 2.7$
Risk aversion: σ_1	2	2	CKK
Frisch Elasticity: σ_2	4	3.3	Frisch= $\frac{1}{2}$
Value of Leisure: χ	1.08	2.2	Avg. $h_j + n_j = \frac{1}{3}$
Government Parameters			
G	15.6	15.3	17% of Y

4.2 Preferences

Agents have time-separable preferences over consumption and leisure. I use $\frac{c^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h)^{1-\sigma_2}}{1-\sigma_2}$ as the benchmark utility function. In Section 7.1 I check the sensitivity of the results with regards to a different utility function in which labor instead of leisure directly enters the utility function.

I determine β such that the capital-to-output ratio is 2.7, in accordance with U.S. data.[34] I determine χ such that under the baseline-fitted U.S. tax policy, agents spend on average one-third of their time endowment working.[35] Following CKK, I set $\sigma_1 = 2$, which controls the relative risk aversion. Past micro-econometric studies (such as Altonji (1986), MaCurdy (1981), and Domeij and Flodén (2006)) estimate the Frisch elasticity to be between 0 and 0.5. However, more recent research has shown that these estimates may be biased downward. Reasons for this bias include utilizing weak instruments, not accounting for borrowing constraints, disregarding the life-cycle effect of endogenous-age specific human capital, omitting correlated variables such as wage uncertainty, and not accounting for labor market frictions.[36] Therefore, I set σ_2 such that if agents work one third of their time endowment then the Frisch elasticity is at the upper bound of the range (0.5).[37]

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

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

[36] Some of these studies include Imai and Keane (2004), Domeij and Flodén (2006), Pistaferri (2003), Chetty (2012), and Contreras and Sinclair (2008).

[37] This value is consistent the estimates in Kaplan (2012).

4.3 Idiosyncratic Productivity

The idiosyncratic labor productivity shocks are calibrated based on the estimates from the PSID data in Kaplan (2012).[38] These permanent, persistent, and transitory idiosyncratic shocks to individuals' productivity are distributed normal with a mean of zero and the shock parameters are set in accordance with the estimates in Kaplan (2012): $\rho = 0.958$, $\sigma_\alpha^2 = 0.065$, $\sigma_\nu^2 = 0.017$ and $\sigma_\varepsilon^2 = 0.081$. I discretize all three of the shocks in order to solve the model, using two states to represent the transitory and permanent shocks and seven states for the persistent shock.[39] For expositional convenience, I refer to the two different states of the permanent shock as high and low ability.

4.4 Age-Specific Human Capital

The age-specific human capital calibration parameters are different in the exogenous and LBD models. In the exogenous model, I set $\{\varepsilon_j\}_{j=20}^{j_r-1}$ to be consistent with the values estimated in Kaplan (2012) which are based off of the average hourly earnings by age in the Panel Survey of Income Dynamics.[40]

In the LBD model I use the same functional form for human capital accumulation as in Hansen and İmrohoroğlu (2009),

$$s_{j+1} = \Omega_j s_{i,j}^{\Phi_1} h_{i,j}^{\Phi_2}, \tag{21}$$

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. I do not set $\{s_{i,j}\}_{j=0}^{j_r-1}$ directly, rather I calibrate the sequence $\{\Omega_j\}_{j=20}^{j_r-1}$ such that the agents' equilibrium labor choices lead the average $\{s_{i,j}\}_{j=20}^{j_r-1}$ under the baseline-fitted U.S. tax code to match the age-specific human capital calibrated in the exogenous model ($\{\varepsilon_j\}_{j=20}^{j_r-1}$).[41] Similar to Hansen and İmrohoroğlu (2009), I calibrate the rest of the LBD parameters based on the estimates in Chang et al. (2002), setting $\Phi_1 = 0.407$ and $\Phi_2 = 0.326$.

[38] For details on estimation of this process, see Appendix E in Kaplan (2012).

[39] I use the Rouwenhorst method to discretize the persistent shock since Kopecky and Suen (2010) demonstrate with highly persistent processes this method is preferred.

[40] I make three adjustments to the process. The profile is smoothed by fitting a quadratic function in age, normalized such that the value equals one when an agent enters the economy, and is extended to cover ages 20 through 66.

[41] I calibrate $\{\Omega_j\}_{j=20}^{j_r-1}$ such that the sequence is smooth over the life-cycle. Although $\{\Omega_j\}_{j=20}^{j_r-1}$ has an affect on some of the other labor targets like the Frisch elasticity and average hours worked, I find that the effect is fairly minimal and it mostly affects the skills profile.

4.5 Firm

I assume the aggregate production function is Cobb–Douglas. The capital share parameter, ζ, 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.

4.6 Government Policies and Tax Functions

While calibrating parameters by matching certain targets in the models and the data it is important to incorporate a tax function that is similar to the U.S. tax code. I use the estimates of the U.S. tax code in Gouveia and Strauss (1994) for this tax policy, which I refer to as the baseline-fitted U.S. tax policy. The authors match the U.S. tax code to the data using a three parameter functional form,

$$T(Ty;\Upsilon_0,\Upsilon_1,\Upsilon_2) = \Upsilon_0(Ty - (Ty^{-\Upsilon_1} + \Upsilon_2)^{-\frac{1}{\Upsilon_1}}), \quad (22)$$

where Ty represents the sum of the taxable labor and capital income. The average tax rate is principally controlled by Υ_0, and Υ_1 governs the progressivity of the tax policy. To ensure that taxes satisfy the budget constraint, Υ_2 is left free. Gouveia and Strauss (1994) estimate that $\Upsilon_0 = .258$ and $\Upsilon_1 = .768$ when fitting the data. The authors do not fit separate tax functions for labor and capital income. Accordingly, I use one tax system on aggregate income for the baseline-fitted U.S. tax policy. However, when searching for the optimal tax policy, I allow for different tax rates on capital and labor income.

I calibrate government consumption, G, so that it equals 17 percent of output under the baseline-fitted U.S. tax policy to be consistent with CKK. When searching for the optimal tax policy, I restrict 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. To be consistent with CKK, Social security tax rates are set at 12.4% and the maximum taxable income for the social security program is set at 2.5 times the average earnings in the economy.

5 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)). The social welfare is equivalent to maximizing the ex-ante expected lifetime utility of agents before entering the model. When searching for the optimal tax policy I determine an optimal tax policy where capital and labor income are taxed at separate rates. I determine the optimal flat capital income (τ_k) tax but allow for a progressive labor income tax policy. I search over three different functional forms for the progressive labor income tax policy.[42] First, I use the three parameter functional form from Gouveia and Strauss (1994) such that labor taxes equal,

$$\tau_{h0}^{gs}(y^l - ((y^l)^{-\tau_{h1}^{gs}} + \tau_{h2}^{gs})^{-\frac{1}{\tau_{h1}^{gs}}}, \tag{23}$$

where y^l is the taxable labor income. Second, I use the two parameter functional form from Bénabou (2002) where labor taxes equal,

$$y^l - \tau_{h0}^b (y^l)^{1-\tau h1^b}, \tag{24}$$

where generally τ_0^b governs the level of taxation and τ_1^b governs the progressivity. Finally, I examine a three parameter piecemeal tax function where labor taxes equal,

$$\tau_{h0}^p \min\{0, y^l - \tau_{h1}^p\} - \tau_{h2}^p \tag{25}$$

where τ_{h0}^p is the tax rate, τ_{h1}^p is a fixed deduction, and τ_{h2}^p is a lump-sum transfer. Therefore the computational experiment is maximizing the expected utility for a newborn,

$$SWF(\tau_{h0}, \tau_{h1}, \tau_{h2}, \tau_k) = \mathbb{E}\left[u(c_1, h_1) + \sum_{s=1}^{J-j-1} \beta^s \prod_{q=1}^{s} (\Psi_q) u(c_{s+1}, h_{s+1})\right], \tag{26}$$

where τ_{h0},τ_{h1}, and τ_{h2} are the relevant labor income tax parameters for the functional form of interest, and τ_k is the flat tax rate on capital income.[43] To determine the effects of endogenous human capital accumulation, I compare the tax policies that maximize the SWF in the two models.

6 Results

In this section, I quantitatively assess the effects on the optimal tax policy of including LBD in a calibrated life-cycle model. In order to assess LBD's effect, I determine the optimal tax policies in the exogenous and LBD models, highlight the channels that cause the differences, and determine the effects on the economy of these optimal policies.

[42] All of these functional forms nest a flat labor income tax rate.

[43] The search is done separately for each tax function. In the case of the two parameter functional form τ_{h2} can be ignored.

6.1 Comparison of Model to the Data

Prior to examining the effects of LBD on the optimal tax policy, I compare both models' predictions for the life-cycle profiles to the data. Figure 1 plots the average life-cycle profiles from the models under the baseline-fitted U.S. tax policy and in the data.[44] The upper-left panel compares the average percent of the time endowment that is spent working over the lifetime and the upper-right compares the labor income. The actual labor supply and labor earnings profiles are constructed from the 1967 - 1999 waves of the Panel Survey of Income Dynamics (PSID). In the data I focus my attention on the labor supply and labor earnings for the head of the household between ages 20 and 80. The lower-left panel compares the consumption profile in the model to the per-capita expenditures on food in the PSID.[45] The lower-right panel examines the median savings in the models and median total wealth in the 2007 Survey of Consumer Finances (SCF) for individuals between the ages of 20 and 80.[46] I smooth through some of the volatility in the wealth data by using five year age bins.[47]

Focusing on the labor supply profiles, the models' predicted profiles have a different general pattern than the data. In the data, the labor supply profile is more hump shaped. In contrast, the models tend to predict the labor supply will decrease throughout the working lifetime. However, in the last few working years this decrease in labor supply is larger in the LBD model, more closely matching the data. Moreover, the models severely over predicts the amount of time young agents spend working because in the models agents cannot borrow against future earnings. Therefore, in order to accumulate precautionary savings agents work more early in their lifetime.[48] In contrast, in the data, some young households may have a means to borrow, minimizing the severity of this effect.

Despite the differences in labor supply, the profile of labor earnings in the data is more similar to the ones generated by the models (upper-right panel). All three labor earnings profiles are hump shaped with a peak just after forty years old. The LBD model does a bit better job matching the earnings profile at the end of the working lifetime because of the larger drop in labor supply for the last few years of the working lifetime. One main difference between the data and both of the model generated profiles is how labor earnings evolve after the age of 66. Agents are forced to retire at 66 in the models, but in the U.S. economy some head of

[44]Earnings, consumption, and savings from the models are converted to real 2012 dollars by equating the average earnings in each of the models and the data.

[45]Per-capita expenditures for each household are calculated as the total family's expenditures dividied by the total number of individuals in the household (including children).

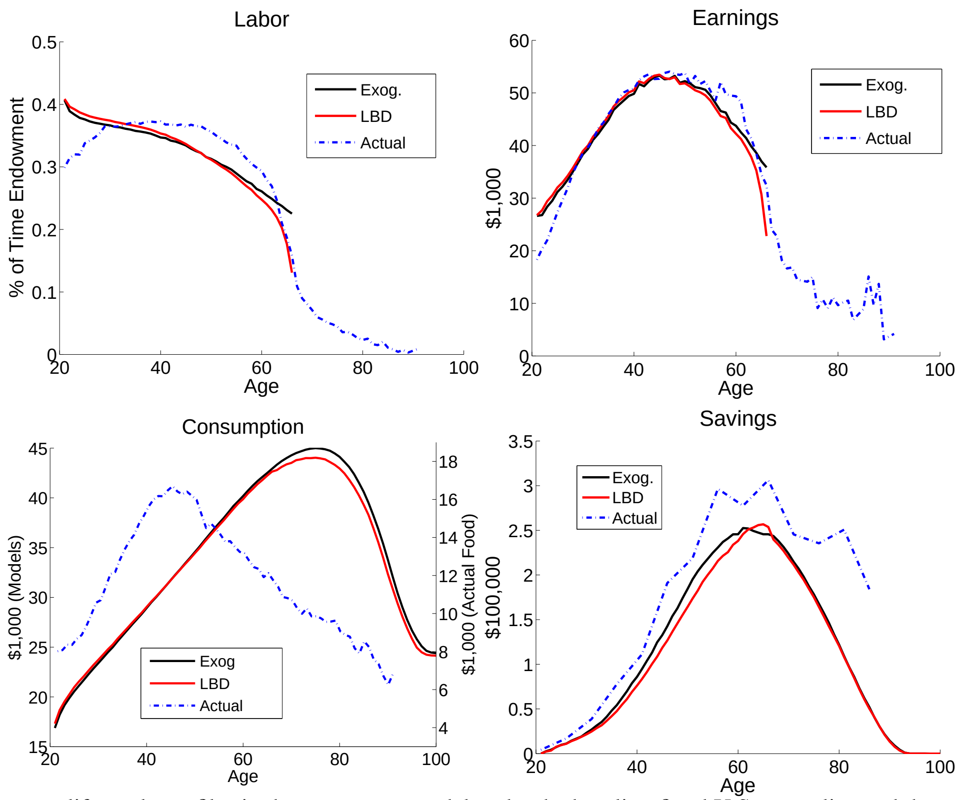
[46]When comparing the savings data to the model I choose to use the median as opposed to the mean so that the upper-tail of the distribution does not skew the comparison statistic.

[47]The data for individuals after age 80 are not included because there were few observations in the sample leading the smoothed estimates to be extremely volatile.

[48]For further discussion see Heathcote et al. (2010)

households retire after the age of 66. Thus, the earnings profile for these older households are higher in the data than in either of the models.

Figure 1: **Actual and Exogenous life-cycle Profiles**



Note: These plots are life-cycle profiles in the exogenous model under the baseline-fitted U.S. tax policy and the actual profiles in the data. The units of the consumption, earnings, and capital profiles are converted to real dollars by matching the average labor earnings in the model and in the data. The labor, earnings, and consumption profiles are the average across the cohort. The savings profiles are the median values within each cohort.

When comparing the consumption profiles, I find that all the profiles are hump-shaped. However, I find that consumption on food tends to peak earlier in the data than total consumption in either of the models. Additionally, comparing the growth in consumption from the age 20 to the peak, the model exhibits more growth in consumption over the lifetime. One possible reason for these differences is that the PSID data are limited to just expenditures on food while the model generated consumption represents all consumption. In part, this may be due to a varying percentage of total expenditures devoted to food over the lifetime.[49]

Finally, I find that the savings profiles are similar in the models and the data. All of the median savings profiles are hump-shaped, peaking between $250,000 and $300,000 at the age of 60. One smaller discrepancy is that the models predict that agents will deplete their savings more quickly than savings declines in the data. This difference could arise because the model does not include any motive for individuals leaving

[49] For example, if food is less of a luxury good than other expenditure categories then expenditures on food may peak before total expenditures.

a bequest for younger generations or holding savings in case of some unexpected end of life expenditures such as medical expenses.

Overall, both models do a fair job matching the data with respect to earnings and savings. Moreover, if anything the LBD model does a bit better job matching the earnings profile at the end of the working lifetime. However, one concern is that neither model produces a labor supply profile that matches the shape of the data. Thus, Section 7.1 examines whether the effect of LBD is consistent even when I use an alternative utility function that implies less relationship between the labor supply profile and optimal tax policy.[50]

6.2 Optimal Tax Policies in Exogenous and LBD Models

When determining the optimal tax policy, I search over all three tax functions described in Section 5. I find that in both models using the piecemeal tax function of the form of a flat tax on labor income with a fixed deduction coupled with a lump-sum transfer is optimal.[51] This tax function implies that for agents with low labor income the marginal labor income tax rate is zero and the average rate is negative.[52] Table 3 describes the optimal tax policies in the two models and Figure 2 plots the average and marginal labor tax rates by income in both models. Starting with the exogenous model, the optimal tax policy is a 30 percent tax on capital, a 32.5 percent tax on labor income with a fixed deduction of $\$6,218$, and a lump-sum transfer of $3,683. In contrast, in the model with LBD, I find that the optimal tax policy is a 36 percent tax on capital, a 22.3 percent tax on labor income with a fixed deduction of $10,901, and a lump-sum transfer of $365. There are two main differences between the optimal tax policies in the models. First, the optimal labor tax policy is much flatter in the LBD model primarily because of the much smaller lump-sum transfer. Second, the tax on capital is notably larger in the LBD model.

These changes in the optimal tax policy are due to the intertemporal distortion channel and elasticity channel. First, as Section 2.2 demonstrates, the distortions from a progressive labor tax are magnified with LBD since the current labor choice affects the level of future human capital. For example, in the LBD model working more today will increase human capital, which would imply a higher future marginal labor income tax rate with a progressive tax policy. A progressive tax provides the benefit of insurance against

[50]Although the labor supply profile can affect the optimal tax policy through numerous channels, the general shape does not seem related to the optimal tax policy when the utility function is both separable and homothetic in each consumption and labor (see Peterman (2013) for more details).

[51]It is not surprising that the optimal tax policy is of the form of a flat tax since other studies (see CKK) tend to find that the optimal policies can be closely approximated by a flat tax with a fixed deduction. When solving for the optimal with this functional form I found that it was important to use a grid search in order to ensure I found a global optimum.

[52]All working agents receive the lump-sum portion of the labor tax. In addition, any income under the fixed deduction is not subject to the marginal tax rate. Thus, if an agent's taxable labor income is less than the fixed deduction then their total labor tax bill is the negative value of the lump-sum.

Table 3: **Optimal Tax Policies**

Tax Parameters	Exog	LBD
Labor Tax Rate	32.5%	22.3%
Fixed deduction	\$6,218	\$10,901
Lump-sum	\$3,683	\$365
Capital Tax Rate	30%	36%

Figure 2: **Optimal Labor Income Tax Rates**

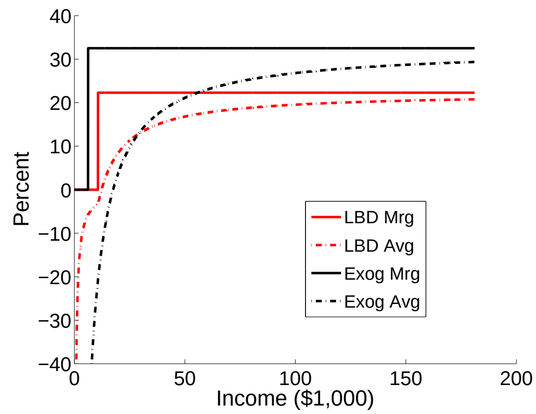
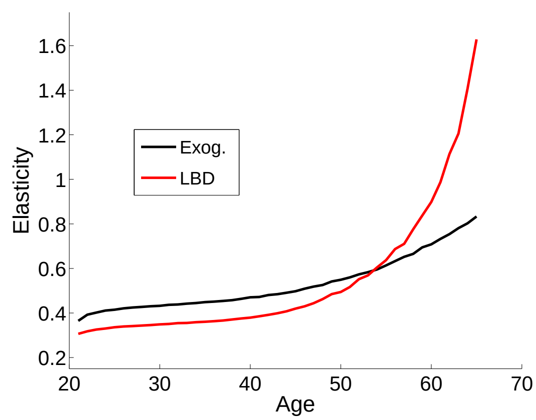


Figure 3: **Frisch Elasticity**



Note: The upper-left panel are the optimal tax policies. The upper-right panel plots the average and marginal tax rates under the optimal tax policies in each model. The lower panel plots the Frisch labor supply elasticity profiles in each model.

the idiosyncratic labor productivity but distorts agents' decisions. Thus, adding LBD changes this tradeoff between efficiency and equty leading a flatter labor income tax policy to be optimal.

Second, as Section 2.1 demonstrates, adding LBD causes agents to supply labor even more elastically as they age because the human capital benefit decreases (elasticity channel). This change in the Frisch labor supply elasticity is apparent in the Frisch labor supply elasticity profile from the models (see Figure 3). Because the Frisch elasticity tends to increase more as an agent ages in the LBD model, the social planner would like to tax labor income earned when an agent is young at a relatively higher rate than income earned when an agent is old. The social planner can mimic this type of age-dependent tax with either a tax on capital or a progressive/regressive tax on labor income. Specifically, a positive tax on capital implies that the tax on labor income is monotonically decreasing as an agent ages.[53] In contrast, the effectiveness of

[53]In particular, since the returns from saving compound, the implicit tax on labor income from a positive tax decreases at an exponential rate as an agent ages.

mimicking this type of age-dependent tax policy with a progressive/regressive labor income tax depends on the shape of the average labor earnings profile. The upper right panel of Figure 1 demonstrates that although average labor income is increasing over a slight majority of the working lifetime in the LBD model, the income profile is hump shaped. The non-monotonicity of the labor income profile implies that although a more regressive labor income tax will tend to tax labor income earned when an agent is young at a relatively higher rate, it will also tax labor income earned at the end of the working lifetime at a relatively higher rate. Therefore, it is likely that the elasticity channel is responsible for the change in the optimal capital tax and is less responsible for the change in the optimal progressivity. In Section 7.1, I confirm in a slightly different model that the elasticity channel is primarily responsible for the change in the optimal capital tax leaving the intertemporal distortion channel primarily responsible for the reduction in the progressivity of the optimal tax policy.

6.3 Welfare Effects

In order to determine the economic significance of endogenous human capital on optimal taxation, I turn to the welfare effects of not accounting for LBD when solving for various pieces of the optimal tax policy. I measure welfare in consumption equivalent variation (CEV) which is defined as the uniform increase in consumption an agent would need at each age in order to be indifferent to being born into an economy with a less optimal tax policy compared to living in an economy with the optimal tax policy.[54] First, I determine the welfare loss in the LBD model from adopting the optimal tax policy solved for in the exogenous model (which includes a progressive labor tax and a lower tax on capital) as opposed to the true optimal tax policy (which includes a flat tax on labor income and a larger tax on capital). I find ignoring LBD when solving for the optimal tax policy causes a notable welfare loss that is equivalent to .73 percent of expected lifetime consumption (see Column II of Table 4).[55]

Adopting this sub-optimal tax policy has two effects on taxes. First, it entails adopting a labor tax policy that is more progressive than optimal. Second, it entails adopting a tax policy in which the ratio of the capital-to-labor tax rates is relatively lower than optimal. In order to determine the implications of each of these effects on welfare, I determine the welfare loss when the tax policy includes the second but not the first effect. In particular, I include the progressivity parameters (lump-sum and fixed deduction) from the

[54]The increase is calculated as the ex-ante increase prior to an agent realizing their idiosyncratic wage shocks or age of death. I calculate the welfare loss as this increase in consumption for an agent living under a sub-optimal tax policy necessary to make them indifferent to living under the optimal tax policy.

[55]The labor tax rate under the tax policy solved for in the exogenous model (see the first column of Table 3) is slightly different from the labor tax rate when applying this policy to the LBD model (see column II of Table 4) in order to ensure the government budget constraint is satisfied.

Table 4: **Welfare Loss in LBD Model When Using Optimal Tax Policies from Different Models**

	Opt. I	Exog. Opt. II	Exog. Level III	Exog. Capital IV
Labor Tax Rate	22.3%	33.5%	26.4%	23.3%
Fixed deduction	$10,901	$6,272	$10,901	$11,327
Lump-sum	$365	$3,715	$365	0
Capital Tax Rate	36%	30%	24.3%	30%
CEV		0.73%	0.2%	0.05%

Note: Welfare losses are solved for in the LBD model and are relative to the optimal tax policy in Column I (see second column of Table 3). Column II calculates the welfare loss from living in an economy with the sub-optimal tax policy in which the rates are consistent with the optimal tax policy solved for in the exogenous model (leaving the labor tax rate free to clear the government budget constraint). Column III determines the welfare effects of the sub-optimal tax that includes the optimal ratio of the labor and capital tax rate solved for in the exogenous model but includes the optimal progressivity parameters from the LBD model. Column IV solves for the optimal tax policy in the LBD model restricting the capital tax rate to be the optimal rate in the exogenous model and determines the welfare loss.

true optimal LBD tax policy but a ratio of the capital-to-labor tax rates that is consistent with the optimal ratio in the exogenous model. The welfare loss from including just the sub-optimal capital-to-labor tax ratio is equivalent to .20% of expected lifetime consumption (see column III), about one quarter of the total welfare loss when both effects are included. Column IV determines how much of the welfare loss from adopting a sub-optimal capital tax rate can be mitigated when the progressivity parameters are re-optimized. In particular, column IV adopts the sub-optimal lower capital tax rate but allows the social planner to re-optimize the labor income tax parameters conditional on this lower capital tax rate. I find that coupled with this low capital tax rate the social planner prefers to include a larger fixed deduction and eliminate the lump-sum transfer leading to only a minimal welfare loss (only .05% of expected lifetime consumption) compared to the unrestricted optimal tax policy. Taken as a whole, these results demonstrates that a majority of the welfare lost in the LBD model from not accounting for LBD is due to the inclusion of a more progressive tax policy (the first effect) and almost all of the remaining welfare loss from the lower than optimal capital tax (the second effect) can be reduced by adjusting the progressivity of the labor income tax policy.

I further investigate the welfare consequences of adopting sub-optimal tax parameters in both models in Table 5 in order to determine the relative sensitivity of welfare to the different parameters. Columns I-III examine the effects in the exogenous model, and columns IV-VI examine the effects in the LBD model. Column I and IV determine the effects when the capital tax rate is increased by the equivalent of fifty percent of the optimal rate in the exogenous model. Likewise, Columns II and V examine the implications when the fixed deduction is increased by fifty percent and columns III and VI determine the welfare effects when the lump-sum transfer is increased by fifty percent. The general sizes of the welfare losses from changing each of the tax parameters are fairly similar in both models with a change in the lump-sum transfer causing the largest welfare losses and a change in the fixed deduction causing the smallest welfare losses. Moreover, the welfare from raising the capital tax 15 percentage points in the LBD model is equivalent to 0.35% of expected lifetime consumption. The size of this welfare loss demonstrates some of the reason for the relatively smaller effect from the capital tax in Table 4 column II is because of the smaller difference in the optimal capital tax rates in the two models than the differences in the progressivity parameters.

Table 5: **Welfare Effects of Misspecified Optimal Tax Policy**

	Exogenous			**LBD**		
	I	**II**	**III**	**IV**	**V**	**VI**
Labor Tax Rate	28.1%	37%	40%	16.3%	25.6%	31.1%
Fixed deduction	$6,218	$9,327	$6,218	$10,901	$14,10	$10,901
Lump-sum	$3,683	$3,683	$5,524	$365	$365	$2,207
Capital Tax Rate	45%	30%	30%	51%	36%	36%
CEV	0.26%	0.12%	0.33%	0.35%	0.09%	0.44%

Note: The welfare effects are the welfare losses from switching from the optimal tax policy in each specific model to these non-optimal tax policies.

6.4 Effects on Macroeconomy

To fully understand the effects of endogenous human capital accumulation, I analyze the aggregate economic variables and life-cycle profiles in both models under the baseline-fitted U.S. tax policy as well as the changes induced by implementing the optimal tax policies.

6.4.1 The Effects of Adding Endogenous Age-Specific Human Capital

First, I compare the effect on the aggregate economic variables and life-cycle profiles from adding LBD to the exogenous model under the baseline-fitted U.S. tax policy. Table 6 summarizes the aggregate economic variables under both the baseline-fitted U.S. tax policy and optimal tax policies. Figure 4 plots the life-cycle profiles of hours, consumption, assets, and age-specific human capital in both models.

Comparing the exogenous and LBD models under the baseline-fitted U.S. tax policy, the first and fourth columns of Table 6 demonstrate that the levels of aggregate hours, labor supply, and aggregate capital are similar in the two models. The calibrated parameters are determined so that under the baseline-fitted U.S.

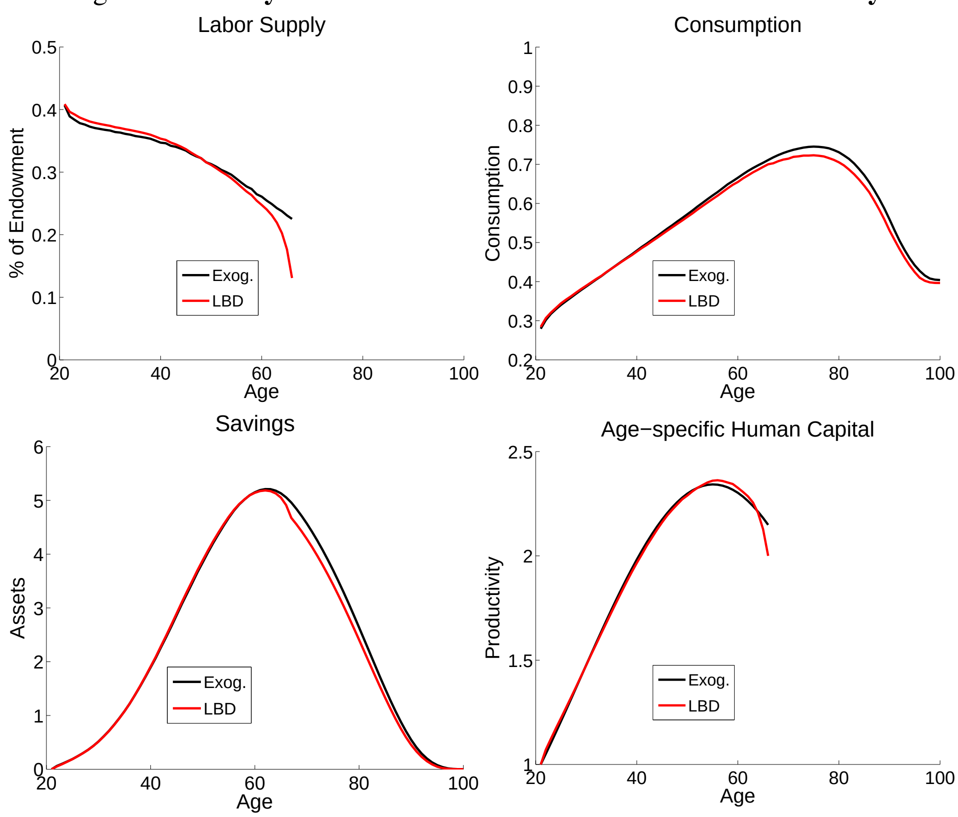
Table 6: **Aggregate Economic Variables**

	Exogenous			LBD		
Aggregate	**Baseline**	**Optimal**	***%Δ from Baseline to Optimal***	**Baseline**	**Optimal**	***%Δ from Baseline to Optimal***
Y	0.91	0.87	-4.83%	0.90	0.88	-2.56%
K	2.48	2.28	-7.93%	2.43	2.25	-7.38%
N	0.52	0.51	-3.04%	0.52	0.52	0.23%
Avg Hours	0.33	0.32	-5.49%	0.33	0.33	-0.6%
w	1.12	1.10	-1.85%	1.12	1.09	-2.79%
r	0.05	0.05	9.02%	0.05	0.06	13.84%
beq	0.03	0.03	-9.35%	0.03	0.03	-6.25%
CEV			1.12%			0.69%
Average Tax Rate	**Baseline**	**Optimal**		**Baseline**	**Optimal**	
Labor	0.19	0.18		0.19	0.15	
Capital	0.18	0.30		0.18	0.36	
Ratio	1.07	0.59		1.07	0.42	
Marginal Tax Rate	**Baseline**	**Optimal**		**Baseline**	**Optimal**	
Labor	0.2	0.32		0.2	0.22	
Capital	0.19	0.30		0.19	0.36	
Ratio	1.06	1.08		1.07	0.62	

Note: The average hours refers to the average percent of time endowment worked in the productive labor sector. Since the marginal tax rates vary with income, the reported marginal tax rates are the population and income weighted average marginal tax rates for all agents.

tax policy the models match certain targets from the data. Since many of these aggregate economic variables are targets used to calibrate each of the models, the aggregates are similar in the two models.

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



Note: These plots are life-cycle profiles of the three calibrated models under the baseline-fitted U.S. tax policy.

Although adding LBD does not have a large effect on the aggregate economic variables, it does cause some 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 solid red lines in the upper-left panel of Figure 4).[56] The upper-right panel shows that the lifetime consumption profile is a touch flatter in the LBD model compared to the exogenous model. The intertemporal Euler equation generally controls the slope of the consumption profile over an agent's lifetime. The relationship is

$$\left(\frac{c_{j+1}}{c_j}\right)^{\sigma_1} = \Psi_j \beta \widetilde{r}_t, \tag{27}$$

where $\widetilde{r}_t$ is the marginal after-tax return on capital. The right-hand-side of Equation 27 is lower in the

[56] In both the LBD and exogenous models the intratemporal marginal rate of substitution between consumption and leisure is affected by scale in such a way that high-type agents (with higher consumption) tend to work less than low-type agents. The difference between hours worked by high-type and low-type agents is of a similar magnitude in both models so it is not responsible for the differences in the two models' labor supply profiles.

LBD model, primarily because β is lower, which leads to the flatter consumption profile. Generally, agents have a similar level of savings during their working years in both models. Thus, the flatter consumption profile in the LBD model translates into less savings for the second half of the agents life because with less consumption they do not need as high of level of savings (see the lower-left panel). The lifetime age-specific human capital profiles are similar in the two models since the sequence of parameters $\{\Omega j\}_{j=20}^{j_r-1}$ is calibrated so that age-specific human capital in the LBD model matches the exogenous model (see the lower-right panel of Figure 4).

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

This subsection examines the effects on the economy of adopting the optimal tax policy in the exogenous model. In the exogenous model, the optimal capital tax is larger than the tax under the baseline-fitted U.S. tax policy so adopting the optimal tax policy causes a decrease in aggregate capital (see columns one and two of Table 6). The average marginal labor tax is higher under the optimal tax policy but the average labor tax is smaller. Thus, adopting the optimal tax policy causes a bit larger decrease in aggregate capital than in aggregate labor leading the rental rate to increase and the wage rate to decrease. Adopting the optimal tax policy in the exogenous model leads to a welfare increase that is equivalent to 1.12% of expected lifetime consumption.[57]

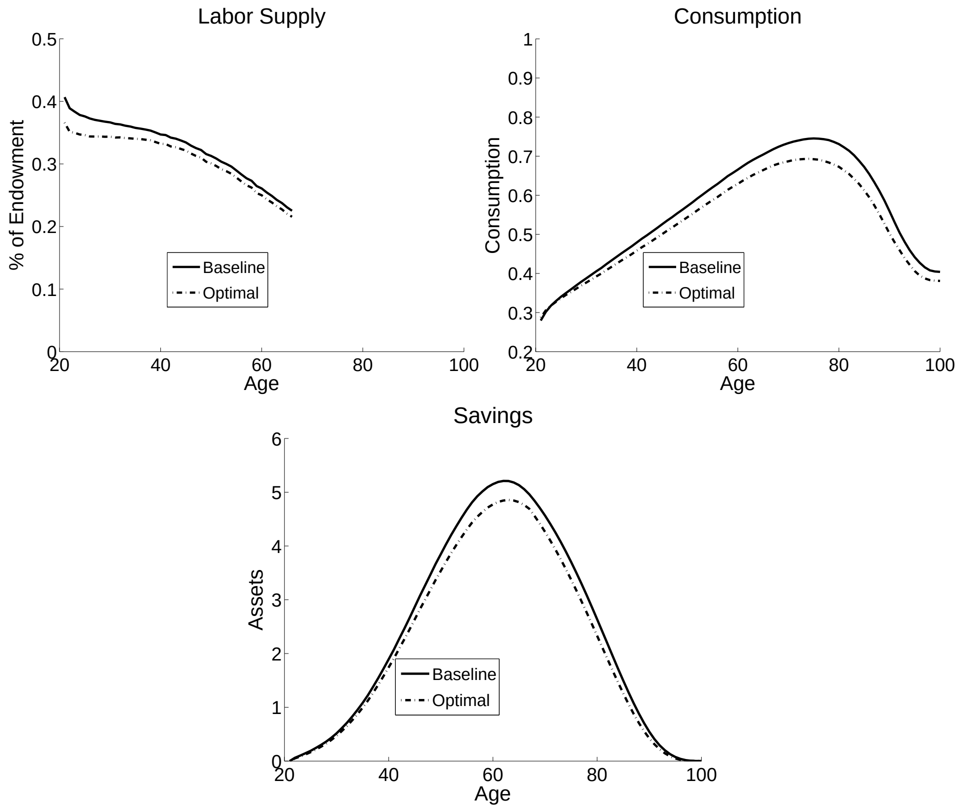
Figure 5 plots the life-cycle profiles for time worked, consumption, and assets in the exogenous model under the baseline-fitted U.S. tax policies and the optimal tax policies. Adopting the optimal tax policy in the exogenous model causes changes in all three life-cycle profiles: (i) agents work less, especially early in their life, (ii) agents save less, and (iii) the lifetime consumption profile is flatter. Overall, agents work less because of the higher marginal labor tax and the lower wage. Moreover, the higher implicit tax on young labor income due to the increase in the tax rate on capital is responsible for agents working considerably less early in their life.[58]

Implementing the optimal tax policy causes an increase in both the capital tax and the rental rate on

[57] Allowing for age-dependent taxation could potentially lead to a bit larger welfare gains. For example, Gervais (2012) finds in a similar model with exogenous human capital accumulation that the welfare loss from disallowing age-dependent taxation and restricting the social planner to a particular class of tax functions is approximately 0.4 percent of lifetime consumption. da Costa and Santos (2015) finds in a life-cycle model that the welfare gains from allowing the tax policy to be conditioned on age is between 2 and 4 percent of lifetime consumption depending on whether human capital is accumulated exogenously or endogenously. However, these welfare estimates may be an upper bound. In particular, they restrict the class of both age-independent and age-dependent tax functions to the two parameter functional form from Bénabou (2002). In my model I found that the optimal policies from this tax function were always inferior to the other two functional forms. Thus, incorporating a more flexible functional form that allows for a lump-sum would likely lead to welfare gains both with age-independent and age-dependent tax policies. However, one would suspect that these gains would be even greater for the age-independent tax policy since the optimal progressivity would be more important.

[58] In addition agents work at least a little less across all ages of their lifetime because of the higher marginal tax on labor.

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

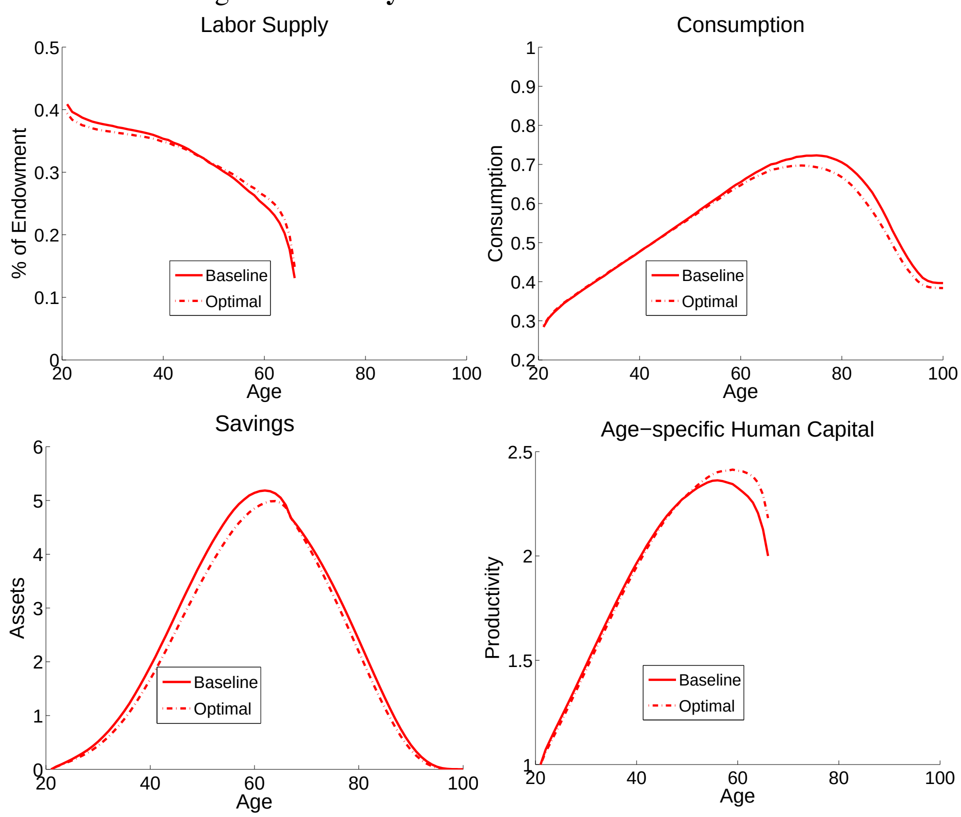


Note: Since the skills are the same in the exogenous models under the baseline-fitted U.S. tax policy and optimal tax policy, they are not plotted.

capital, leading to shifts in both the consumption and savings profiles. Overall, the increase in the capital tax dominates, so the marginal after-tax return on capital falls under the optimal tax policy, leading to a flatter consumption profile (Figure 5, upper-right panel). Moreover, because of the lower returns, agents tend to hold less savings throughout their lifetime (see the lower left panel of Figure 5).

6.4.3 The Effects of Optimal Tax Policy in the LBD Model

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



Adopting the optimal tax policy in the LBD model causes a large increase in the capital tax and a flattening of the labor tax policy. Thus, adopting the optimal tax policy in the LBD causes an even larger decrease in the average labor tax rate and a smaller increase in the average marginal labor tax rate than in the exogenous model. These tax changes lead to a similar fall in aggregate capital as in the exogenous model. However, because the optimal labor tax rate is smaller in the LBD model compared to the exogenous model, adopting the optimal tax policy causes a small increase in aggregate labor (as opposed to the small decrease in the exogenous model). Adopting the optimal tax policy leads to a welfare increase across all agents that is equivalent to 0.69% of lifetime consumption, smaller than in the exogenous model. One reason for the smaller welfare gains from adopting the optimal tax policy in the LBD model is because the optimal tax policy in the LBD model is closer to the baseline-fitted US tax policy than in the exogenous model.[59]

Implementing the optimal tax policies in the LBD model causes agents to shift time worked from earlier to later years in response to the larger capital tax, which implicitly taxes labor income from early years at a higher rate. Unlike the exogenous model where there is an overall decline in labor, in the LBD model the change is more of a reshuffling of hours from earlier in their lifetime to later in their lifetime (upper-left panel of Figure 6).[60] This shift in hours leads age-specific human capital to also be lower in the first half of agents' lifetimes and higher in the second half under the optimal tax policy. Moreover, because agents work less in the first half of their life they also tend to save less during these ages (lower-left panel of Figure 6). There is a similar flattening in the consumption profile from adopting the optimal tax policy in the LBD model.

7 Sensitivity of Results

In this section, I examine the sensitivity of the results with respect to a number of dimensions.

7.1 Alternative Utility Function

I test the sensitivity of the effect of LBD on the optimal tax policy using an alternative utility function in which the consumption and labor components are separable and each homothetic. In particular, I use the utility function $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$. I refer to this utility function as "labor utility" since labor, as opposed to leisure, enters the utility function.[61] In the labor utility function, σ_2 is the Frisch elasticity in the exogenous model.

There are three advantages to using this utility function. First, using the labor utility function implies that in the exogenous model the Frisch elasticity is not a function of the agent's labor decision. This flexibility in the Frisch elasticity allows one to decompose the role that each channel plays in the change to the optimal tax policy when including LBD.[62] Second, using this function allows me to test whether the results are sensitive to the inclusion of leisure as opposed to labor as the direct input into the utility function. Third, as Peterman

[59] In particular, in both models because of the inclusion of the lump-sum transfers in the optimal tax policies, the optimal average tax rates are more progressive than in the baseline-fitted US tax policy. However, in the case of the exogenous model the optimal average labor tax rates are even more progressive than the baseline-fitted US tax policy or the optimal policy in the LBD model.

[60] The smaller overall changes to the labor supply profile from adopting the optimal tax in the LBD model compared to the exogenous model are not due to offsetting changes across different types of agents. Instead the smaller changes in the profiles reflect that in general agents respond less to adopting the optimal tax policy in the LBD model than in the exogenous model since the increase in the marginal tax rate is small enough that it does not cause an overall reduction in the labor supply.

[61] All of the parameters are recalibrated in this model such that the model still matches the relevant targets.

[62] In the benchmark utility function leisure as opposed to labor enters the utility function. Thus in the benchmark utility function the agents labor supply decision affects the Frisch labor supply elasticity. Thus doing a similar decomposition with the benchmark utility function is not feasible.

(2013) demonstrates, the optimal tax policy tends to be less sensitive to the labor supply profile with this type of utility function. Since there are some differences between the model generated labor supply and the data, it is useful to confirm that the effects of adding LBD on the optimal tax policy are similar with the labor utility in order to confirm that the effects are not a byproduct of changes in the labor supply profiles.[63]

The first two columns of Table 7 present the optimal tax policies in the exogenous and LBD models using the labor utility function, respectively. Comparing these two columns, adding LBD causes a similar change in the optimal tax policy with the labor utility function as it does with the benchmark utility function. Specifically, adding LBD causes the optimal tax on labor income to be less progressive and the optimal tax on capital to be larger. Although the changes from adding LBD are similar, the decrease in the progressivity is a bit smaller and the increase in the capital tax is a bit larger with the labor utility function than with the benchmark utility function. I calculate the welfare loss for an agent in the LBD model when the optimal tax policy solved for in the exogenous model (which includes more progressivity and a lower tax on capital) is included, as compared to the welfare for the agent if the true optimal tax policy (which includes a flatter labor tax and a higher tax on capital) is included. I find that the welfare loss from not accounting for LBD with the labor utility function is equivalent to 1.19% of lifetime consumption, a bit larger than the same calculation under the benchmark utility (0.72%-see column I of Table 4). Thus, the overall result that incorporating LBD causes a notable change in the optimal tax policy is robust to this change in the utility function.

Table 7: **Optimal Tax Policies With Labor Utility Function**

Tax Parameters	**Exog I**	**LBD II**	**Alt. Exog III**
Labor Tax Rate	35.2%	20.1%	35.4%
Fixed deduction	$1,030	$4,241	$1,272
Lump-sum	$5,816	$909	$6,664
Capital Tax Rate	24.1%	33.5%	39.2%

Note: All models include the labor utility function. The exogenous model includes human capital accumulation exogenously. The LBD model includes human capital accumulation with LBD. The alternative exogenous model includes human capital accumulation exogenously but is altered such that the Frisch labor supply elasticity matches the LBD model.

Next, I decompose the effect of each of the channels on the optimal tax policy. The third column details the optimal tax policy in an alternative exogenous model in which human capital is accumulated exogenously, but σ_2 is set to vary by age such that the Frisch elasticity profile in the alternate exogenous model

[63]Further confirming this result, I found that when χ is calibrated such that the labor supply profiles in the exogenous models match the data (with either the benchmark or labor utility function) the changes in the optimal tax policies are quite minimal.

matches the Frisch elasticity profile in the LBD model.[64] Solving the optimal tax policy in this alternative exogenous model isolates the effect of the elasticity channel since it incorporates this channel in the exogenous model by including an age-dependent Frisch elasticity but excludes the additional intertemporal link. The optimal tax policy in this alternative exogenous model (column III) includes a similar amount of progressivity as the optimal tax policy in the exogenous model with the constant Frisch elasticity (column I). However, the optimal tax on capital is even larger in column III than when LBD is included (column II). Thus, the elasticity channel can more than account for the increase in the optimal tax on capital when LBD is included. Given that the average labor income profile is hump shaped, it is not surprising that the social planner finds a tax on capital, as opposed to a regressive labor income tax, to be the more effective way to mimic age-dependent taxes that monotonically decrease with age. Moreover, the similar progressivity of the labor income tax in the exogenous model and the alternative exogenous model demonstrates that the elasticity channel is not responsible for the change in the optimal progressivity of the labor income tax when LBD is included. Instead, the intertemporal distortion decreases the desirability of using a progressive labor tax to redistribute in the LBD model, since the additional intertemporal link in the LBD model magnifies the distortion from a progressive labor tax.

7.2 Parameterization of LBD Function

In order to test how sensitive the results are to the LBD accumulation function, I solve for the optimal tax policy with two different variants of the parameters in the LBD model. In particular, I examine how the optimal tax policy changes if I independently increase ϕ_1 or ϕ_2. I examine the effects of a one standard deviation increase in these parameter values as measured in Chang et al. (2002).[65] Table 8 describes the optimal tax policy in the benchmark LBD model and these two variants. Comparing the optimal tax policy in column I to the optimal policies in columns II and III, changing either of these parameters has only minimal implications on the optimal tax policy. Thus, the changes in the optimal tax policy due to incorporating LBD seem to be fairly robust to the choice of parameters for the LBD accumulation function.

[64] In order to keep the alternative exogenous model consistent I use the same other calibration parameters as in the benchmark exogenous model and also hold transfers and social security benefits constant between the two models. I choose to keep these constant because Peterman (2013) demonstrates that changes in both can have large impacts on the optimal tax policy.

[65] In each of these variants of the LBD model I recalibrate all of the parameters such that the model matches the specified targets in Table 2.

Table 8: **Optimal Tax Policies With Alternative Utility**

Tax Parameters	**Benchmark I**	**Higher Φ_1 II**	**Higher Φ_2 III**
Labor Tax Rate	22.3%	23.5%	22.8%
Fixed deduction	\$10,0901	\$10,872	\$10,806
lump-sum	\$365	\$661	\$546
Capital Tax Rate	36%	35.9%	36.3%

Note: All models include the benchmark utility function. Column I are the benchmark results from Table 3. Column II and III are results with the optimal tax policies when ϕ_1 and ϕ_2 are increased by one standard deviation, respectively.

7.3 Transition

This paper focuses on computing the optimal steady state tax policy from the perspective of an agent who has yet to enter the economy. This welfare criteria neglects any of the extra costs or benefits that may exist during the transition to the steady state with the new tax policy. There are three reasons why one might think that agents who are alive during this transition (transitional agents) may experience a different welfare effect compared to the steady state welfare effects. First, these living agents are already part way through their lifetime which changes the relative composition of the sources of their income. In particular, the farther along an agent is in their life the larger percentage of their future lifetime income will come from capital as opposed to labor. Since adopting the optimal steady state tax policies in both models includes a decrease in the labor tax and increase in the capital tax, adopting these policies will tend to favor transitional agents who are younger at the time of the adoption. Although this effect may cause differential welfare effects between cohorts, it will not cause a uniform increase or decrease in the welfare effects across all the cohorts.

The second reason for a discrepancy between the transitional and steady state welfare effects is that agents may directly benefit or suffer due to a different level of aggregate capital in the initial steady state (under the baseline-fitted US tax policy) and the final steady state (under the optimal tax policy). Aggregate capital tends to decrease over the transition in both models. In order to bring about this reduction, agents will need to increase their consumption over the transition. Thus, the higher level of consumption will cause this second effect to benefit transitional agents and lead to even larger welfare gains than in the steady state.

Finally, since it is likely that it will take many periods for capital to transition to the new steady state level under the optimal tax policies, both the rental rate and wage rate will not immediately jump to their new levels. Unlike the first two effects, it is somewhat harder to determine whether this will lead to more or less of a welfare benefit for transitional agents. In order to get a sense of the impact of the slow transition of capital, I calculate the welfare effects in both models in a counterfactual partial equilibrium steady state where the optimal tax policies are adopted, but aggregate capital is fixed at the level under the baseline-fitted US tax policy.[66] As opposed to the steady state welfare gain of 1.12% and .69%, I find that the welfare gains in these counterfactual steady states are .89% and .98% in the exogenous and LBD models, respectively. The steady state welfare gain is a bit smaller than the counterfactual steady state in the case of the exogenous model, and a bit larger in the case of the LBD model. Although there are some differences, agents still experience a welfare gain in this counterfactual steady states in both the LBD and exogenous models. Taken as a whole, it seems unlikely that the steady state welfare gains would be completely reversed and that on average transitional agents would not still experience a welfare gain from adoption of these optimal tax policies in both the exogenous and LBD models.

8 Conclusion

Two important questions for optimal taxation are should the income tax policy be progressive and at what rate should capital be taxed? In this paper, I examine the effect of LBD on optimal taxation and find that it affects the answers to both questions. Analytically, I demonstrate that including endogenous human capital accumulation changes the Frisch elasticity over an agent's lifetime. Thus, the elasticity channel creates a motive for the government to condition labor income taxes on age, and if disallowed, either a non-zero capital tax or progressive/regressive labor income tax can be used to mimic these age-dependent taxes. Although a progressive/regressive tax can be used to mimic an age-dependent tax, I show that the distortions from this type of tax are magnified in the LBD model due to the additional intertemporal link. Thus, the intertemporal distortion channel makes it less appealing to use a progressive tax.

Quantitatively, I find that these two channels cause the inclusion of LBD to substantially change the optimal tax policy. Specifically, including LBD causes the optimal tax policy to include a flatter labor income tax and higher capital tax rate. Moreover, not accounting for LBD causes a loss in welfare of between .7 and 1.2 percent of expected lifetime consumption depending on the utility function. A majority of this welfare loss comes from including a sub-optimal amount of progressivity. I find that these results are robust to the utility function, and the parameters used in the LBD skill accumulation process. Given the size of the effects on welfare and the robustness of the effect, these results indicate that when examining optimal taxation, accurately incorporating the skills accumulation process can be of first-order importance.

[66] I allow labor to respond immediately to the change in the tax policy. This steady state is in partial equilibrium because the equilibrium condition that aggregate capital equals the summation of all agents' level of savings is violated. However, all the other aggregate variables are consistent with the equilibrium conditions.

A Analytical Derivations

When solving for the optimal tax policy in the analytical model, I use the standard social welfare function,

$$[U(c_{2,0},1-h_{2,0})/\theta]+\sum_{t=0}^{\infty}\theta^{t}[U(c_{1,t},1-h_{1,t})+\beta U(c_{2,t+1},1-h_{2,t+1})]. \tag{28}$$

which maximizes the expected utility discounting future generations with social discount factor θ. In order to determine the optimal tax policy I use the primal approach. In this approach the social planner maximizes directly over allocations. In order to ensure that the chosen allocation can be supported in a competitive equilibrium the implementability constraint is included. In particular, the implementability constraint is formed first by combining equations 2 and 3 to form a joint intertemporal budget constraint,

$$c_{1,t}+\frac{c_{2,t+1}}{1+r(1-\tau_k)}=w(1-\tau_{h,1})h_{1,t}+\frac{w(1-\tau_{h,2})s_2(h_{1,t})h_{2,t+1}}{1+r(1-\tau_k)}. \tag{29}$$

The implementability constraint is then determined by taking the agent's intertemporal budget constraint and replacing prices and taxes with the agent's first-order conditions.[67]

$$c_{1,t}U_{c1}(t)+\beta c_{2,t+1}U_{c2}(t+1)-h_{1,t}\left(U_{h1}(t)+\frac{-\beta h_{2,t+1}s_{h1}(t+1)U_{h2}(t+1)}{s_2(h_{1,t})}\right)-\beta h_{2,t+1}U_{h2}(t+1)=0. \tag{30}$$

Suppressing the arguments of the skill accumulation formula, the Lagrangian that the Social Planner maximizes is

$$\begin{aligned}\mathscr{L}=&U(c_{1,t},1-h_{1,t})+\beta U(c_{2,t+1},1-h_{2,t+1})\\&-\rho_t(c_{1,t}+c_{2,t}+K_{t+1}-K_t+G_t-rK_t-w(h_{1,t}+h_{2,t}s_{2,t}))\\&-\rho_{t+1}\theta(c_{1,t+1}+c_{2,t+1}+K_{t+2}-K_{t+1}+G_{t+1}-rK_{t+1}-w(h_{1,t+1}+h_{2,t+1}s_{2,t+1}))\\&+\lambda_t\left(c_{1,t}U_{c1}(t)+\beta c_{2,t+1}U_{c2}(t+1)-h_{1,t}\left(U_{h1}(t)+\frac{-\beta h_{2,t+1}s_{h1}(t+1)U_{h2}(t+1)}{s_{2,t+1}}\right)-\beta h_{2,t+1}U_{h2}(t+1)\right)\end{aligned} \tag{31}$$

where ρ is the Lagrange multiplier on the resource constraint and λ is the Lagrange multiplier on the implementability constraint.[68] Assuming that the utility function is separable in consumption and labor ($U_{ch}=0$), the first-order conditions with respect to capital, labor and consumption are

$$\rho_t=\theta(1+r)\rho_{t+1} \tag{32}$$

$$\begin{aligned}w\rho_t=&(1-\lambda_t)U_{h1}(t)+\lambda_t h_{1,t}U_{h1h1}(t)\\&-\frac{h_{2,t+1}(w\theta\rho_{t+1}s_{2,t+1}^2s_{h2}(t+1)-\beta\lambda_tU_{h2}(t+1))}{s_{2,t+1}^2}[s_{2,t+1}(s_{h2}(t+1)+h_{1,t}s_{h2h2}(t+1))-h_{1,t}s_{h2}^2(t+1)]\end{aligned} \tag{33}$$

$$w\rho_{t+1}\theta s_2(t+1)=\beta\left((1-\lambda_t)U_{h2}(t+1)+\lambda_t h_{2,t+1}U_{h2h2}(t+1)-\frac{\lambda_t s_{h2}(t+1)}{s_{2,t+1}}(h_{2,t+1}h_{1,t}U_{h2h2}(t+1)-h_{1,t}U_{h2}(t+1))\right) \tag{34}$$

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

[68] In addition to the implementability constraint, the resource constraint and government budget constraint are included. However, due to Walras' Law, the government budget constraint can be excluded from the Lagrangian.

$$\rho_t = (1+\lambda_t)U_{c1}(t) + \lambda_t U_{c1c1}(t)c_{1,t} \tag{35}$$

$$\theta\rho_{t+1} = \beta[(1+\lambda_t)U_{c2}(t+1) + \lambda_t U_{c2c2}(t+1)c_{2,t+1}]. \tag{36}$$

Combining equations 35 and 36 yields the expression

$$\frac{\beta\rho_t}{\theta\rho_{t+1}} = \frac{Uc1(t) + \lambda_t(U_{c1}(t) + U_{c1c1}(t)c_{1,t})}{U_{c2}(t+1) + \lambda_t(U_{c2}(t+1) + U_{c2c2}(t+1)c_{2,t+1})} \tag{37}$$

Assuming that the utility function demonstrates constant relative risk aversion, then this expression further simplifies to,

$$\frac{\beta\rho_t}{\theta\rho_{t+1}} = \frac{Uc1(t)}{U_{c2}(t+1)}. \tag{38}$$

Combining equations 4, 5, and 6 yields,

$$\frac{(1-\tau_{h1})}{(1-\tau_{h2})} = \frac{U_{c2}(t+1)U_{h1}(t)s_{2,t+1}}{U_{c1}(t)U_{h2}(t+1)} - \frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)}. \tag{39}$$

Combining equations 38 and 39 and simplifying yields,

$$\frac{(1-\tau_{h1})}{(1-\tau_{h2})} = \frac{\theta\rho_{t+1}U_{h1}(t)s_{2,t+1}}{\beta\rho_t U_{h2}(t+1)} - \frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)}. \tag{40}$$

In order to determine the optimal labor income taxes, I combine equations 40, 32, 33, and 34 and suppress the time subscripts,

$$\frac{(1-\tau_{h1})}{(1-\tau_{h2})} = \frac{(1-\lambda) + \frac{\lambda h_2 U_{h2h2}}{U_{h2}} - \frac{\lambda s_{h2}}{s_2 U_{h2}}\left(h_2 h_1 U_{h2h2} - h_1 U_{h2}\right)\left(1 + \frac{h_2 s_{h2}}{1+r(1-\tau_k)}\right)}{(1-\lambda) + \frac{\lambda h_1 U_{h1h1}}{U_{h1}} + \frac{h_2\beta\lambda U_{h2}}{s_2^2 U_{h1}}\left(s_2(s_{h2} + h_1 s_{h2h2}) - h_1 s_{h2}^2\right)} - \frac{h_2 s_{h2}}{1+r(1-\tau_k)} \tag{41}$$

B Definition of Stationary Competitive Equilibrium

In this section I define the stationary competitive equilibrium for the model. An agent's state variables are assets a, previous periods human capital s, age j, ability α, persistent shock ν, and idiosyncratic shock θ. For a given set of exogenous demographic parameters $\{n, \Psi_j\}$, a sequence of skill accumulations parameters $\{\Omega_j\}_{j=20}^{j_r-1}$, 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 maximum amount of taxable income for social security $\bar{y}$, social security benefits SS, a production plan for the firm (N, K), an age-specific human capital accumulation function $S : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, and a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, a stationary competitive equilibrium consists of agents' decision rules $\{c, h\}$ for each state x, factor prices $\{w, r\}$, accidental bequests beq, and the distribution of individuals $\{\mu(x)\}$ such that the following holds:

1. Given prices, policies, accidental bequests, benefits, and that ω follows equation 17 or 18, the agent

maximizes equation 16 subject to

$$c+a' = w\omega h - \tau_{ss} w\omega h, +(1+r)(a+beq_t) - T^l[w\omega h - .5\tau_{ss}\min\{w\omega h,\bar{y}\})] - T^k[r(a+beq)], \quad (42)$$

for $j < j_r$, and

$$c+a' = SS + (1+r)(a+beq) - T^k[r(a+beq)], \quad (43)$$

for $j \geq j_r$.
Additionally,

$$c \geq 0, 0 \leq h \leq 1, a \geq 0, a_1 = 0. \quad (44)$$

2. Prices w and r satisfy

$$r = \zeta \left(\frac{N}{K}\right)^{1-\zeta} - \delta \quad (45)$$

and

$$w = (1-\zeta)\left(\frac{K}{N}\right)^{\zeta}. \quad (46)$$

3. The social security policies satisfy

$$\tau_{ss} = \frac{\sum\limits_{j>=j_r} ss\, \mu(x)}{\sum\limits_{j<j_r} \min\{hw\omega, \bar{y}\}\, \mu(x)}. \quad (47)$$

4. Accidental bequests are given by

$$beq = \sum (1-\Psi) a'\, \mu(x). \quad (48)$$

5. Government balances its budget

$$G = \sum T^k[r(a+beq)]\, \mu(x) + \sum_{j<j_r} T^l[w\omega h - .5\tau_{ss}\min\{w\omega h, \bar{y}\})]\, \mu(x). \quad (49)$$

6. The market clears

$$K = \sum a\, \mu(x), \quad (50)$$

$$N = \sum h\omega\, \mu(x), \quad (51)$$

and

$$\sum c\, \mu(x) + \sum a'\, \mu(x) + G = K^{\zeta} N^{1-\zeta} + (1-\zeta)K. \quad (52)$$

7. The distribution of $\mu(x)$ is stationary, that is, the law of motion for the distribution of individuals over the state space satisfies $\mu(x) = Q_\mu \mu(x)$, where Q_μ is a one-period recursive operator on the distribution.

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*, 45–85.

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

Atkeson, Andrew, Varadarajan Chari, and Patrick Kehoe, "Taxing Capital Income: A Bad Idea," *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 1999, *23* (3), 3–17.

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

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

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

Bénabou, Roland, "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 2002, *70* (2), 481–517.

Cespedes, Nikita and Michael Kuklik, "Optimal Taxation and Life Cycle Labor Supply Profile," Working Paper, Banco Central De Reserva Del Peru December 2013.

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 33209, Munich Personal RePEc Archive August 2011.

Chetty, Raj, "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 2012, *80* (3), 969–1018.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger, "Taxing Capital? Not a Bad Idea After All!," *American Economic Review*, 2009, *99* (1), 25–38.

Contreras, Juan and Sven Sinclair, "Labor Supply Response in Macroeconomic Models: Assessing the Empirical Validity of the Intertemporal Labor Supply Response from a Stochastic Overlapping Generations Model with Incomplete Markets," MPRA Paper 10533, University Library of Munich, Germany September 2008.

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

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

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.

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

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

Diamond, Peter A. and Emmanuel Saez, "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 2011, *25* (4), 165–90.

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.

Farhi, Emmanuel and Iván Werning, "Insurance and taxation over the life cycle," *The Review of Economic Studies*, 2013, *80* (2), 596–635.

Fuster, Luia, Ayşe İmrohoroğlu, and Selahattin İmrohoroğlu, "Altruism, Incomplete Markets, and Tax Reform," Mimeo 2007.

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

Gervais, Martin, "On the optimality of age-dependent taxes and the progressive U.S. tax system," *Journal of Economic Dynamics and Control*, 2012, *36* (4), 682 – 691.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski, "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 2003, *70*, 569–587.

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

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.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 2010, *118* (4), pp. 681–722.

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*, 2010, *17* (5), 451–478.

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

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

Kaplan, Greg, "Inequality and the life cycle," *Quantitative Economics*, 2012, *3* (3), 471–525.

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

Kopecky, Karen A. and Richard M.H. Suen, "Finite State Markov-chain Approximations to Highly Persistent Processes," *Review of Economic Dynamics*, 2010, *13* (3), 701 – 714.

Krause, Alan, "Optimal nonlinear income taxation with learning-by-doing," *Journal of Public Economics*, 2009, *93* (9-10), 1098–1110.

Kremer, Micahel, “Should Taxes be Independent of Age?,” Working Paper, Harvard 2002.

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.

Mirrlees, J.A., “An Exploration in the Theory of Optimal Income Taxation,” *Review of Economic Studies*, April 1971, *38* (2), 175–208.

Mirrlees, James, “Notes on welfare economics, information and uncertainty,” *Essays on Economic Behavior under Uncertainty*, 1974, pp. 243–261.

Mirrlees, James Alexander, Stuart Adam, Tim Besley, Richard Blundell, Steve Bond, Robert Chote, Malcom Gammie, Paul Johnson, Gareth Myles, and James Poterba, *Dimensions of tax design: the Mirrlees review*, Oxford University Press, April 2010.

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

_ , “Determining the motives for a positive optimal tax on capital,” *Journal of Economic Dynamics and Control*, 2013, *37* (1), 265–295.

_ , “Taxing Capital? The Importance of How Human Capital is Accumulated,” Working Paper 2013-020, Federal Reserve Board of Governors January 2014.

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.

Stantcheva, Stefanie, "Learning and (or) Doing: Human Capital Investments and Optimal Taxation," Working Paper, Harvard 2015.

— , "Optimal Taxation and Human Capital Policies over the Life Cycle," Working Paper, Harvard 2015.

Stiglitz, Joseph E., "Self-selection and Pareto efficient taxation," *Journal of Public Economics*, 1982, *17* (2), 213 – 240.

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

Varian, Hal R., "Redistributive taxation as social insurance," *Journal of Public Economics*, 1980, *14* (1), 49 – 68.

Weinzierl, Matthew, "The surprising power of age-dependent taxes," *The Review of Economic Studies*, 2011, *78* (4), 1490–1518.