Interpretation of the Results

We prepared this attachment to provide complete intuition and more detailed analysis of our results. We start with the simplest model with inelastic labor supply and add the model elements one by one. The first class of models with exogenous labor supply can be seen as an extreme case of zero elasticity of labor supply is zero, and thus it presents a useful comparison benchmark for our results in the paper. We show that while it is possible to generate a rationale for positive capital income taxation even in this case, this rationale is not robust to the introduction of progressive labor income taxes

We then study various versions of life cycle models with endogenous labor supply and demonstrate, both theoretically and quantitatively, that there is a robust and quantitatively significant reason to tax capital. More specifically, as long as households endogenously choose labor supply over a nontrivial and realistically modeled life cycle it is optimal to tax capital at a significantly positive rate. We also briefly discuss the relationship of our results to those obtained by Aiyagari (1995) who argues that uninsurable idiosyncratic income risk (possibly in conjunction with tight borrowing constraints) may provide a rationale for taxing capital in the long run even in models with infinitely lived households.

Finally, we add a short section on the transition dynamics. Although our quantitative analysis in the paper focuses on steady state results, we argue that our main result would likely be strengthened if we could explicitly compute the optimal policy along the transition path..

In what follows we will index different versions of our model by Mx; the table at the end of the attachment provides a summary of the key quantitative results of all versions of the model that we study in this attachment.

1 Models with Exogenous Labor Supply

In this section we first demonstrate that, contrary to what intuition may suggest, even with exogenous labor supply there may be a role for capital income taxation, despite the fact that labor is supplied inelastically. But we also argue that in this class of models the case for positive capital income taxation is not robust if one allows for a progressive labor income tax, a crucial element of our quantitative model.

1.1 The Simplest Model with Exogenous Labor Supply (M1)

In this subsection we analyze the simplest OLG model with exogenous labor supply for which some analytical results can be derived. Part of the intuition here (but only a part) extends to our full quantitative model with endogenous labor supply.[1]

1.1.1 Environment

The population grows at constant rate n. Households live for J periods and inelastically supply one unit of labor in every period of their life. There is an exogenous stream of government expenditures that also grows at rate n. Households value consumption stream according to

$$\sum_{j=1}^{J} \beta^{j-1} u(c_{jt}),$$

where u has the usual properties.

Let by N_t denote the size of the cohort born in period t and normalize the size of the first newborn cohort to 1. Then $N_t = (1+n)^t$. We can then write the resource constraint (in terms of per capita of newborn agents) as

$$\sum_{j=1}^{J} \frac{c_{jt}}{(1+n)^{j-1}} + (1+n)K_{t+1} - (1-\delta)K_t + G = f(K_t)$$

where K_t is per-capita (of newborns) capital and G is per capita (of newborns) government consumption.

1.1.2 Social Planner Problem

Suppose the social planner wants to maximize steady state welfare of a newborn agent. Simply taking first order conditions and then imposing steady state time constancy of allocation leads to the first best allocation characterized by

$$f'(K^*) = \delta + n, \tag{1}$$

that is, the golden rule which defines optimal capital per capita K^*. Aggregate consumption is given by

$$C^* = f(K^*) - G - (n+\delta)K^*$$

and the allocation of consumption across ages is determined by the equations

$$[(1+n)\,\beta]^{j-1}\, u'(c_j) = u'(c_1)$$

for all j. Thus the consumption profile in the optimal allocation is determined exclusively by the relationship between n and β. In particular, if $\beta = \frac{1}{1+n}$ then

$$c_1^* = c_2^* = \ldots = c_J^* = \frac{C^*}{\sum_{j=1}^{J} \frac{1}{(1+n)^{j-1}}},$$

whereas if $\beta > \frac{1}{1+n}$ consumption is increasing over the lifetime and decreasing if $\beta < \frac{1}{1+n}$.

[1]We thank Jonathan Heathcote for useful discussions leading to this analysis.

1.1.3 Competitive Equilibrium

Suppose the government can levy exogenously set proportional consumption taxes at rate τ_c, and can decide upon proportional labor and capital income taxes (τ_l, τ_k). Note that because of the exogeneity of labor supply the labor income tax is effectively a lump-sum tax. Thus intuition suggests that one should be able to implement the first best allocation as a competitive equilibrium.

With Government Debt: Denote the per capita amount of assets owned by the government by B (if $B < 0$ the government is in debt). The government budget constraint in per capita terms reads as

$$G = \tau_c C + \tau_l w \sum_{j=1}^{J} \frac{1}{(1+n)^{j-1}} + \tau_k r K + (r-n)B.$$

The household intertemporal budget constraint of a newborn agent reads as

$$\sum_{j=1}^{J} \frac{(1+\tau_c)c_j}{(1+r(1-\tau_k))^{j-1}} = \sum_{j=1}^{J} \frac{(1-\tau_l)w}{(1+r(1-\tau_k))^{j-1}}.$$

The goods market clearing condition reads as

$$f(K) = \sum_{j=1}^{J} \frac{c_j}{(1+n)^{j-1}} + G + (n+\delta)K$$

and the asset market clearing condition is given as follows. Define private asset positions at the beginning of the period as $s_1 = 0$ and then

$$s_j = (1-\tau_l)w + (1+r(1-\tau_k))s_{j-1} - (1+\tau_c)c_j$$

The steady state asset market clearing condition reads as

$$\sum_{j=1}^{J} \frac{s_j}{(1+n)^{j-1}} + B = K$$

Claim 1 *The social optimum can be decentralized as a competitive equilibrium. The optimal capital income tax equals* $\tau_k = 0$.

Proof. The equilibrium interest rate is given by

$$r = f'(K) - \delta$$

Since we want to decentralize the optimal allocation, $K = K^*$ and thus, using (1)

$$r^* = f'(K^*) - \delta = n.$$

From the first order condition of the household we find

$$[(1+r(1-\tau_k))\,\beta]^{j-1}\,u'(c_j) = u'(c_1)$$

For the equilibrium consumption profile to coincide with the socially optimal consumption profile, we require $\tau_k = 0$. The government budget constraint now reads as

$$G = \tau_c C + \tau_l w \sum_{j=1}^{J} \frac{1}{(1+n)^{j-1}},$$

which determines τ_l. Since the socially optimal allocation is feasible, the goods market clearing condition is satisfied. The asset market clearing condition is satisfied by the appropriate choice of government debt B^*. It remains to check whether the socially optimal allocation satisfies the household intertemporal budget constraint. But using the resource constraint and the government budget constraint

$$\begin{aligned}
& \sum_{j=1}^{J} \frac{(1+\tau_c)c_j^*}{(1+r(1-\tau_k))^{j-1}} \\
= & \sum_{j=1}^{J} \frac{(1+\tau_c)c_j^*}{(1+n)^{j-1}} = C^* + \tau_c C^* \\
= & f(K^*) - G - (n+\delta)K^* + G - \tau_l w \sum_{j=1}^{J} \frac{1}{(1+n)^{j-1}} \\
= & f(K^*) - (r+\delta)K^* - \sum_{j=1}^{J} \frac{\tau_l w}{(1+n)^{j-1}} \\
= & \sum_{j=1}^{J} \frac{(1-\tau_l)w}{(1+n)^{j-1}} = \sum_{j=1}^{J} \frac{(1-\tau_l)w}{(1+r(1-\tau_k))^{j-1}},
\end{aligned}$$

we find indeed that the household budget constraint is satisfied at the optimal allocation. ■

Thus with government debt the golden rule optimal allocation can be decentralized as a competitive equilibrium. In this (unique) decentralization, capital income tax equals zero. The government potentially has to hold substantial amounts of negative government debt, the extent of which is determined by how much households save privately (given β and r^*) and thus increases with the impatience of households.

Without Government Debt: The result that the optimal allocation can be decentralized as a competitive equilibrium relies crucially on there being government debt. If one requires $B = 0$ (as we and Hubbard and Judd (1986) in their benchmark model as well as Imrohoroglu (1998) do), then the golden rule social optimum can in general not be decentralized as a competitive equilibrium. The following example makes this clear. Suppose $\beta = \frac{1}{1+n}$. Then in the social optimum consumption is constant at $c_j^* = c^*$. From the intertemporal budget constraint

$$(1+\tau_c)c^* \sum_{j=1}^{J} \frac{1}{(1+r(1-\tau_k))^{j-1}} = (1-\tau_l)w \sum_{j=1}^{J} \frac{1}{(1+r(1-\tau_k))^{j-1}}$$

and thus

$$c_j^* = \frac{(1-\tau_l)w}{(1+\tau_c)},$$

which in turn implies $s_j = 0$ for all j. But with $B = 0$ the asset market clearing condition is not satisfied at the social optimum since $K^* > 0$. Intuitively, nothing guarantees that the steady state aggregate private asset demand equals to the aggregate capital stock. With government debt the government owns the capital stock (i.e. issues negative government debt). This example also shows that potentially large amounts of negative government debt is required to decentralize the golden rule allocation. Ceteris paribus, the higher is β, the more households are willing to accumulate privately and the lower is the amount of assets the government has to accumulate.

Thus without government debt (that is, imposing $B = 0$) the golden rule optimal allocation can almost surely not be decentralized. In this case the optimal capital income tax need not be zero (but it as well can be negative as it can be positive). While the optimum cannot be fully characterized analytically, simple simulations show the following trade-off, as first discussed in Imrohoroglu (1998).

1. A higher τ_k leads to lower private saving, thus a lower aggregate capital stock and thus to lower *aggregate* consumption[2].

2. A higher τ_k tilts the consumption profile. The shape of the consumption profile is determined by $\beta(1+r(1-\tau_k))$. Since a change in τ_k also changes r in general equilibrium, we cannot say anything general about the direction of the consumption profile effect. Furthermore, depending on $\beta < 1$ vs. $\beta > 1$, an upward or downward sloping life cycle consumption profile is preferred by the household.

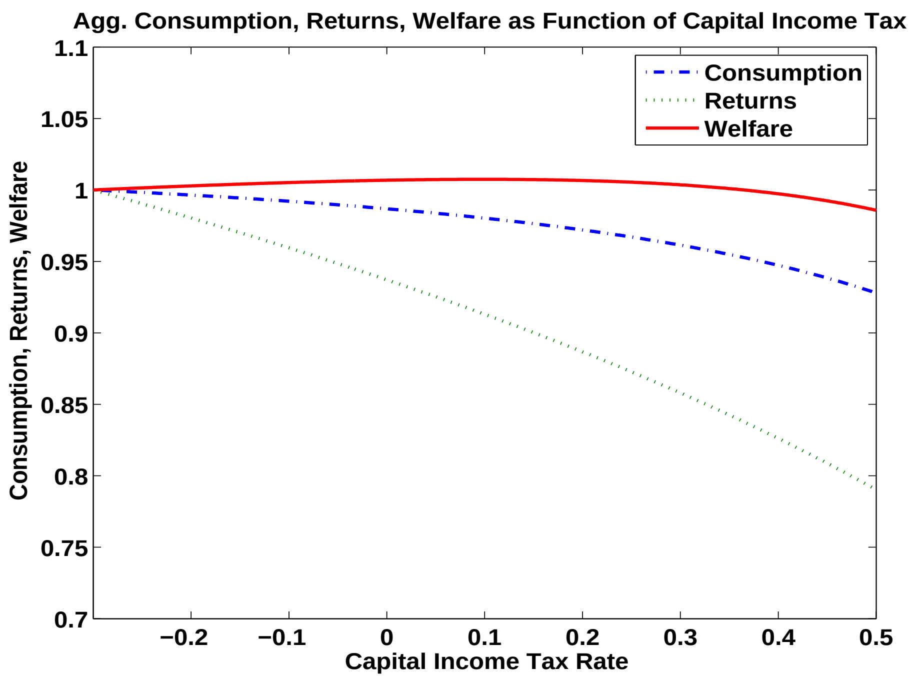
Typical results from simulations of this simple model reveal the following. First, while the optimal capital income tax is in general not equal to zero, it is quantitatively close to zero. For the parameterization that reproduces the same empirical targets as out quantitative model,[3] the optimal capita income tax is 9.9%. The implied labor income tax equals $\tau_l = 19.5\%$. Second, the optimal capital income tax is significantly larger than the (very negative) tax required to maximize per capita consumption. For many parameterizations (including ones used in the quantitative model) there does not exist a τ_k negative enough such that the resulting competitive equilibrium reaches the golden rule capital stock which maximizes per capita consumption. Figure 1 shows the interest rate, aggregate consumption and steady state welfare of a newborn as a function of the capital income tax. Figure 2 shows the age-consumption profile for a capital income tax of -30% and 30% and demonstrates that even though aggregate (and thus per capita) consumption is significantly higher with negative capital income taxes (by more than 5%), because of the consumption profile effect (with high capital income taxes, more is consumed when it is valued more) welfare is higher under the high capital income tax regime, relative to the low capital income tax regime.

[2]As long as the capital stock is smaller than the golden rule, which is the case for all parameterizations even remotely close to the ones we use in the quantitative model and the ones required for the model to hit our empirical targets.

[3]The parameters are given by $\sigma = 2, \beta = 0.9825, \alpha = 0.36, \delta = 0.0833, n = 1.1\%$ and $\tau_c = 5\%$.

Third, comparative statics with respect to β reveal the following. A higher β reduces the optimal capital income tax,[4] and reduces the pre-tax return r and after tax return (at the optimum) $(1-\tau_k)r$. With a higher β households value consumption in the future more and thus prefer a steeper upward sloping consumption profile, and thus, ceteris paribus, a lower capital income tax.

Figure 1:



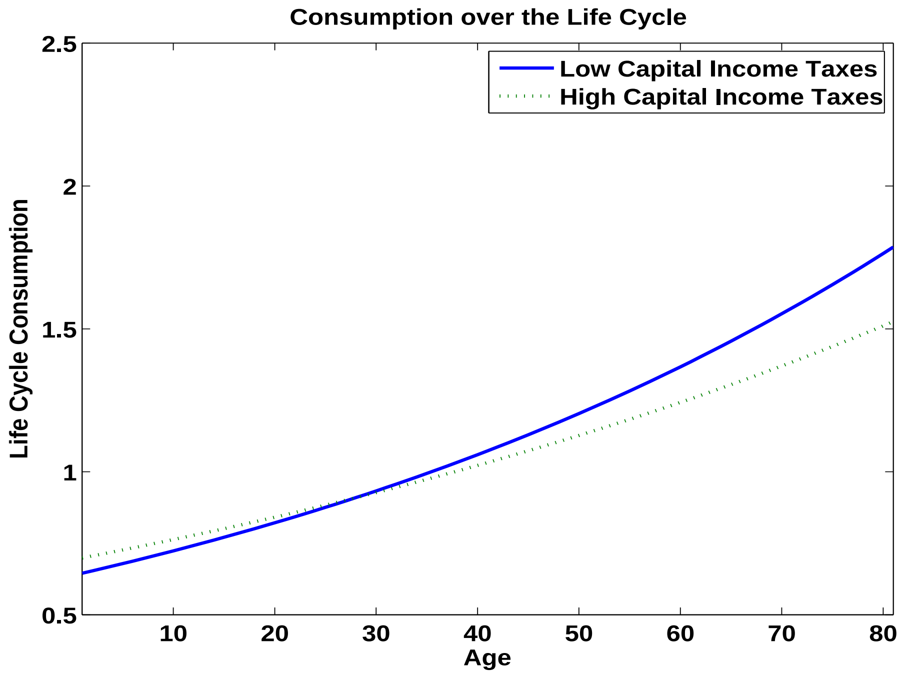
Our results for this model, qualitatively and even quantitatively, are remarkably close to those documented by Imrohoroglu (1998).[5] He uses the same objective function and also restricts himself to steady states without government debt. In his model agents face idiosyncratic earnings shocks and tight borrowing constraints, the combination of which may provide an independent role for positive capital taxation. What our results demonstrate is that even in the absence of such idiosyncratic risk and potentially binding borrowing constraints the optimal capital income tax may be positive, if one abstracts from government debt. Quantitatively, however, the optimal capital income tax is close to zero in this model, for parameter combinations in a large neighborhood of the ones consistent with our empirical targets. Furthermore, large negative government debt is required in this model to implement the socially optimal golden rule allocation, which is consistent with our results with government debt from the quantitative model.

Note that in this simple model the use of progressive labor income taxes does not change the result. Since there is no type heterogeneity or idiosyncratic risk, and since labor income does not depend on age (that is to say, the labor efficiency profile is flat) there is no need

[4]At $\beta = 1.01$, for example, the optimal capital income tax equals 3%.

[5]See, in particular, his sensitivity analysis with respect to the time discount factor β.

Figure 2:



or ability of the government to condition labor income taxes implicitly on age by making the tax code progressive. This result changes drastically once we introduce a more realistic hump-shaped life cycle earnings profile.

1.2 The Life Cycle Model with Exogenous Labor Supply (M2)

We now modify the model above to introduce stochastic mortality, a more realistic labor income process and a PAYGO social security system. These features enable the model to generate a realistic, hump-shaped life cycle consumption profile. Essentially, the fact that $\beta(1 + r(1 - \tau_k)) > 1$ induces consumption to initially rise early on in the life cycle where mortality risk is close to zero.[6] Later in life mortality risk drives down the effective time discount factor and consumption declines.[7]

The results from this extension can be summarized as follows. First, holding the time discount factor fixed at $\beta = 0.9825$, and keeping all other parameters constant as well we obtain an optimal capital income tax that is roughly the same as in the simple model above. The corresponding labor income tax, when the tax system is restricted to be linear, is also

[6]Of course this statement depends on the calibrated β and the capital income tax rate τ_k. It holds for all parameterizations that come anywhere close to our empirical target and for a wide range of capital income tax rates, including the optimal one.

[7]Hansen and Imrohoroglu (2006) have stressed that these model features alone are sufficient to generate a hump in consumption. For our benchmark calibration uninsured mortality risk and a nontrivial life cycle structure of earnings alone lead to a peak in consumption that is, however, too late in the life cycle, relative to the data.

approximately equal to that of the simple model.[8] However, the capital-output ratio generated by the model is significantly lower than our calibration target. Relative to the benchmark, the growing labor income profile reduces incentives to save early in life, as does the positive probability of dying and the PAYGO social security system.

Second, recalibrating all parameters to match the same targets as above (and as in the paper) requires a substantially higher $\beta = 1.001$, which, incidentally is exactly the time discount rate required for our benchmark model to match the target. With a higher β households value future consumption more, and thus the optimal capital income tax (a tax on future consumption) falls. Quantitatively, the optimal capital income tax falls to close to zero and the labor income tax rises to offset the loss in revenues.

Third, this model also shows the potential of progressive income taxation to reallocate resources across households of different ages. So far all results were derived under the restriction of a proportional labor income tax. With progressive labor income taxes the optimal tax code is to tax labor at a 100% rate above a tax-exempt labor income threshold, effectively generating a flat after-tax age-earnings profile. The optimal capital income tax is negative, −23%.

With a labor efficiency (earnings) profile that is upward-sloping progressive labor income taxes redistribute labor income towards young and away from old households. In that way it changes the after-tax labor earnings profile and thus, holding the consumption profile fixed, the life cycle asset accumulation profile as well as aggregate asset demand. For a given consumption profile, shifting after tax labor income towards younger ages increases aggregate saving and thus the capital stock and the *level* of aggregate consumption. But a higher capital stock may also affect the life cycle profile of consumption by lowering the rate of return to capital. The life cycle profile of consumption in the simple model without idiosyncratic income risk, endogenous labor supply and binding borrowing constraints is exclusively determined by the size of $\psi_j \beta(1 + r(1 - \tau_k))$, where ψ_j is the conditional survival probability at age j. A rise in r that is *offset* by a fall in τ_k (in the progressive tax scenario, relative to the proportional tax case) may leave $\beta(1 + r(1 - \tau_k))$ and thus the life cycle consumption pattern unchanged though. In fact, comparing the optimal linear to the optimal nonlinear labor income tax code we find that even though the capital income tax is substantially different in both scenarios, $\beta(1 + r(1 - \tau_k))$ is even slightly lower in the progressive income tax case (1.0408 vs 1.0422).[9]

Obviously this result is an artefact of the exogenous labor supply assumption that ignores the incentive effects of high marginal tax rates on labor supply. However, it demonstrates an important insight that is relevant for our full quantitative model. A positive capital income tax may substitute for age-dependent labor income taxation. If the labor income tax code is allowed to be progressive, however, it can also be used to serve this purpose. With exogenous labor supply there is no reason not to exploit this instrument very heavily. With endogenous labor supply the negative incentive effects on labor supply make it optimal to also rely on capital income taxes to play this role.

[8] The exact numbers are $\tau_k = 13\%$ and $\tau_l = 22\%$.

[9] Adding type heterogeneity to this model and recalibrating it leaves the results virtually unchanged.

1.3 The Role of Borrowing Constraints and Idiosyncratic Risk in Models with Exogenous Labor Supply (M3)

Before turning to models with endogenous labor supply, we show how the previous results relate to the existing literature on the role of borrowing constraints for optimal capital income taxation in life cycle models. Hubbard and Judd (1987) and Imrohoroglu (1998) argue that in life cycle models where households have upward-sloping labor earnings and face tight borrowing constraints the government should not rely on labor income taxes alone, since high labor income taxes translate directly into low consumption especially for young households for whom the constraint tends to be binding. This effect can be especially severe, as Imrohoroglu (1998) argues, if households in addition face idiosyncratic income shocks that forces households with low shocks to consume little when young.

However, as already pointed out by Hubbard and Judd (1986), if the labor income tax code is allowed to be progressive it is possible to tax young households with lower labor income at lower average rates than older households that are not borrowing constrained anymore. Thus the case for positive capital income taxes due to tight borrowing constraints may be weaker if the labor income tax is not restricted to be proportional. In order to evaluate this argument in the context of our model we now add tight borrowing constraints and idiosyncratic risk to the model. Apart from relatively small details, this is the model analyzed by Imrohoroglu (1998), and it is the *same* as our benchmark model with endogenous labor supply being shut down.

Our results confirm the intuition developed above. Restricting the government to proportional labor income taxes, the optimal capital income tax is significantly positive at 24%, despite the fact that the labor income tax remains a lump-sum tax. This capital income tax is (as expected) *higher* than in the model without borrowing constraints. However, as soon as we allow the labor income tax to be progressive, in effect redistributing after-tax labor earnings across different generations (and within generations, across types and households with different income realizations), the optimal capital income tax becomes a negative −16%, financed by a high marginal labor income tax with a substantial deduction. Thus while borrowing constraints in life cycle models, especially in conjunction with idiosyncratic risk, can provide a rationale for substantially positive capital income taxes even if labor is supplied inelastically, that rationale is not robust to the introduction of a more flexible labor income tax. Once the labor income tax is allowed to spare young and poor households (this is effectively what the deduction does) there is no need for the government to use the capital income tax to shift resources towards younger generations.

2 Models with Endogenous Labor Supply

The previous section showed that while theoretically there is room for positive capital income taxes even in models with exogenous labor supply, due to the absence of government debt or due to the presence of borrowing constraints and idiosyncratic risk, if the government can use progressive labor income taxes the implied capital income tax is close to zero or even negative. We will now argue that in realistically calibrated life cycle models with *endogenous labor supply* a significantly positive capital income tax is a robust part of the optima tax system even if the government can implement a progressive labor income tax schedule. In models with exogenous labor supply the distortions of the labor-leisure choice induced by progressive taxes make it suboptimal for the government to exclusively rely on the labor income tax to mimic age-dependent taxes. As we will demonstrate theoretically and quantitatively, a positive capital income tax is useful in this dimension and therefore optimally used by the government. Furthermore, in the *absence* of a flexible labor income tax code (as assumed in most of the existing literature – certainly all the theoretical work on optimal taxation in the Ramsey tradition) the capital income tax plays a potential role in redistributing within age cohorts. One of our contributions is to show that with a flexible labor income tax schedule the capital income tax *need* not play that role. Redistribution and insurance across households with different income levels is effectively accomplished via a progressive labor income tax system.

We will first discuss the optimal capital tax result theoretically in the simplest possible model and then show exactly what model elements are responsible for our quantitative results.

2.1 Theory

We now present the simplest model with endogenous labor supply that allows us to convey the intuition for our quantitative results. Our discussion draws on work by Atkeson, Chari and Kehoe (1999), Erosa and Gervais (2002) and Garriga (2003). For simplicity we abstract from population growth.[10] Households live for 2 periods. Preferences over consumption and leisure are given by

$$U(c_{1t}, l_{1t}) + \beta U(c_{2t+1}, l_{2t+1}).$$

A household of age 1 possesses labor productivity 1 and a household of age 2 possesses labor productivity ε. They maximize lifetime utility subject to the following budget constraints.

$$c_{1,t} + s_t = (1 - \tau_{l1,t}) l_{1,t} \tag{2}$$

$$c_{2,t+1} = (1 - \tau_{l2,t+1}) \varepsilon l_{2,t+1} + (1 + r(1 - \tau_{k,t+1})) s_t \tag{3}$$

Note that the labor income tax rate is allowed to differ by the age of the household. We will investigate below under what conditions it is optimal for the government to use age-dependent taxation. Also note that below we will make sufficient assumptions to guarantee that the interest rate is time-constant and the wage per efficiency unit of labor is constant at 1.

The household optimality conditions are given by

$$\frac{U_{l_{1,t}}}{U_{c_{1,t}}} = -(1 - \tau_{l1,t}) \tag{4}$$

$$\frac{U_{l_{2,t+1}}}{U_{c_{2,t+1}}} = -(1 - \tau_{l2,t+1}) \varepsilon \tag{5}$$

$$\frac{U_{c_{1,t}}}{U_{c_{2,t+1}}} = \beta(1 + r(1 - \tau_{k,t+1})) = \frac{U_{l_{1,t}}}{U_{l_{2,t+1}}} * \frac{(1 - \tau_{l2,t+1})\varepsilon}{(1 - \tau_{l1,t})} \tag{6}$$

[10] None of the substantive results derived in this section depends on this assumption.

which together with the intertemporal budget constraint determine the optimal allocation $(c_{1t}, l_{1t}, c_{2t+1}, l_{2t+1})$, given prices and taxes.

The resource constraint of the economy reads as

$$c_{1,t} + c_{2,t} + K_{t+1} - (1-\delta)K_t + G = rK_t + L_t, \tag{7}$$

where K_t is capital per capita and effective labor supply is given by

$$L_t = l_{1,t} + \varepsilon l_{2,t}.$$

Thus we specify the technology in such a way that marginal products of capital and labor are constant. This assumption allows us to clearly focus on the life cycle elements of the model in that the tax system does not affect wages or rates of return.

Finally, the budget constraint of the government is given by

$$\tau_{l1,t} l_{1,t} + \tau_{l2,t} \varepsilon l_{2,t} + \tau_{k,t} r s_{t-1} = G.$$

Atkeson et al. (1999), Erosa and Gervais (2002) and Garriga (2003) use the primal approach to characterize optimal taxation for this model, under the assumption that the benevolent government discounts future generations with social discount factor θ; that is, the objective of the government is given by[11]

$$\left[U(c_{2,0}, l_{2,0})/\theta\right] + \sum_{t=0}^{\infty} \theta^t \left[U(c_{1,t}, l_{1,t}) + \beta U(c_{2,t+1}, l_{2,t+1})\right]. \tag{8}$$

2.1.1 The Primal Approach

It is common in this literature to use the primal approach, that is, to use the first order conditions of households to replace prices and taxes in the intertemporal household budget constraint. The government maximizes over allocations directly, and faces as restriction the resource constraint and the so-called implementability constraint

$$U_{c_{1,t}} c_{1,t} + \beta U_{c_{2,t+1}} c_{2,t+1} + U_{l_{1,t}} l_{1,t} + \beta U_{l_{2,t+1}} l_{2,t+1} = 0, \tag{9}$$

which is nothing else but the household intertemporal budget constraint with prices and taxes substituted from the first order conditions (4)-(6). This equation summarizes the restrictions on the government that household optimal behavior imposes. The primal approach to optimal taxation then proceeds by letting the government maximize over allocations directly and then using (4)-(6) to back out the appropriate taxes.

Without further restrictions on the government tax instruments the government chooses $\{c_{1t}, l_{1t}, c_{2t+1}, l_{2t+1}, K_{t+1}\}$, given K_0, to maximize (8) subject to (7) and (9).[12] We will first analyze the solution to the unrestricted Ramsey problem and then discuss the implications of imposing additional restrictions. We will do this for both the separable and nonseparable preference specifications used in the quantitative model of our paper.

[11] The case of $\theta = 1$, which corresponds to our steady state assumption, leads to the technical complication that social welfare is not finite. This can be handled by defining optimality by an overtaking criterion, as discussed e.g. in Blanchard and Fischer ((1989), p. 102 and footnote 12 or in Green and Zhou (2002).

[12] The restrictions already implied by the formulation are proportional taxation and the absence of consumption taxes. If, in addition, the government is restricted to use only age independent taxation, the additional

2.1.2 The Separable Case

Attach Lagrange multiplier μ_t to the resource constraint (7) and λ_t to the implementability constraint (9). If utility is of the form

$$\frac{c^{1-\sigma_1}}{1-\sigma_1}+\chi\frac{(1-l)^{1-\sigma_2}}{1-\sigma_2},$$

then the first order conditions with respect to consumption and capital read as

$$c_{1,t}^{-\sigma_1} = \mu_t-\lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \quad (10)$$
$$\beta c_{2,t+1}^{-\sigma_1} = \theta\mu_{t+1}-\lambda_t\beta(1-\sigma_1)c_{2,t+1}^{-\sigma_1} \quad (11)$$
$$\mu_t = \theta\mu_{t+1}(1+r) \quad (12)$$

Combining these equations yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1}=\beta(1+r). \quad (13)$$

Comparing this result to (6) from the household problem immediately implies that without further restrictions on the age-dependence of labor income taxes the optimal capital income tax in this case is zero.

Taking first order conditions with respect to labor supply and using (6) again gives as optimal labor income taxes[13]

$$\frac{1-\tau_{l2,t+1}}{1-\tau_{l1,t}}=\frac{1+\lambda_t\left(1+\sigma_2\frac{l_{1,t}}{1-l_{1,t}}\right)}{1+\lambda_t\left(1+\sigma_2\frac{l_{2,t+1}}{1-l_{2,t+1}}\right)}. \quad (14)$$

In the steady state this implies that $\tau_{l1}>\tau_{l2}$ if and only if $l_1>l_2$, that is, the government optimally conditions labor income taxes on age, and taxes labor more heavily when it is supplied to a larger extent. Note that for this utility specification the Frisch labor supply elasticity is given by

$$\varepsilon_l=\frac{1-l}{\sigma_2 l}$$

constraints imposed on the Ramsey problem are

$$\varepsilon\frac{U_{l_{1,t}}}{U_{c_{1,t}}}=\frac{U_{l_{2,t}}}{U_{c_{2,t}}}.$$

If in addition the government is required to balance the budget in each period, then using Walras' law this constraint translates into

$$K_{t+1}=s_t$$

where s_t in turn can be defined purely in terms of consumption and labor allocations using (2) and substituting out taxes from the first order conditions, that is

$$s_t=-\frac{U_{l_{1,t}}}{U_{c_{1,t}}}l_{1,t}-c_{1,t}.$$

[13] Appendix A contains the algebra for the derivation of (14).

and thus we can re-interpret this result as stating that $\tau_{l1} > \tau_{l2}$ if and only if $\varepsilon_{l1} < \varepsilon_{l2}$, that is, the government finds it optimal to tax labor income more heavily when it is supplied more inelastically, which given our preferences is the case when the the household chooses to work more. In our quantitative model households find it optimal to work more when young for reasons explained below, so the result here implies that with age-dependent taxes the government should tax labor less heavily at older ages.[14]

Now suppose that labor income taxes cannot depend on age, that is, suppose the government is restricted to set $\tau_{l1} = \tau_{l2}$. This imposes the additional restriction

$$\varepsilon \frac{U_{l_{1,t}}}{U_{c_{1,t}}} = \frac{U_{l_{2,t}}}{U_{c_{2,t}}}$$

on the Ramsey problem. From the analysis above we know that this restriction is not binding in the steady state if and only households have a completely flat labor supply profile. In appendix B we show that if this restriction binds then the optimal tax on capital income is nonzero. From the intertemporal Euler equation for labor (leisure) of the household

$$\frac{\varepsilon U_{l_1}}{U_{l_2}} = \beta(1 + r(1 - \tau_k)) \frac{(1 - \tau_{l1})}{(1 - \tau_{l2})}$$

we see that the government can induce a given intertemporal labor (leisure) profile by either age-dependent labor income taxes or capital income taxes. The government can create the same intertemporal tax wedge $(1+r(1-\tau_k))\frac{(1-\tau_{l1})}{(1-\tau_{l2})}$ for the intertemporal labor supply decision by either $\tau_{l1} > \tau_{l2}$ and $\tau_k = 0$ or, if age-dependent labor income taxation is not permitted, by $\tau_{l1} = \tau_{l2}$ and $\tau_k > 0$. That is, a positive capital income tax mimics a labor income tax that is falling with age. Intuitively, if the goal of the government is to tax labor at a lower rate for old people (that is, implicitly tax leisure at a higher rate when old), one way to achieve this is to tax the future at a higher rate by taxing capital income.

Relating our discussion here to the literature, Atkeson Chari and Kehoe (1999), Erosa and Gervais (2002) and Garriga (2003) all contain the result that with age-dependent labor income taxes and a utility function that is separable between consumption and leisure the optimal capital income tax is zero. No theoretical results are derived for the case in which labor income tax rates cannot be conditioned on age, but Erosa and Gervais (2002) show quantitative results for the separable case with $\sigma_1 = \sigma_2$ in which labor income taxes are restricted to be age-independent (and proportional). They demonstrate that in this case the optimal long run capital income tax is positive (see their Table 2). Both Erosa and Gervais (2002) and Garriga (2003) provide the intuition we build upon how capital income taxes can be used as a partial substitute for age-dependent labor income taxes.

2.1.3 The Nonseparable Case

If preferences are nonseparable in consumption and leisure and of the Cobb-Douglas form

$$u(c, l) = \frac{\left(c^{\gamma}(1 - l)^{1-\gamma}\right)^{1-\sigma}}{1 - \sigma}$$

[14] Equation (6) in turn shows that the optimal labor supply profile is crucially determined by the age-efficiency profile (as determined by ε in the simple model) as well as the relationship between the time discount factor and after-tax real interest rate.

the analysis becomes more messy in terms of algebra, but the case for nonzero capital income taxes is strengthened. Now, even if labor income taxes can be conditioned on age, the capital income tax in general is nonzero.[15] Furthermore, the labor supply profile determines both the optimal profile of age-dependent labor income taxes as well as the sign of the capital income tax. In particular, if labor supply falls over the life cycle, then the optimal capital income tax is positive (as long as $\sigma > 1$ which we assume in the quantitative version of our paper).

Finally, the argument above that a positive capital income tax can be used to generate the same intertemporal tax wedge as a labor income tax that is declining with age was independent of the assumed utility function and thus applies here as well.

The key conclusion from the theoretical analysis is that endogenous labor supply, coupled with life cycle model elements that generate a labor supply profile that is not constant across ages (even in the steady state) generates a robust role for positive capital income taxation, as long as the government cannot condition the labor income tax code on the age of the household (and in the nonseparable case, even with age-dependent labor income taxes). The capital income tax indirectly allows the government to tax leisure (labor) at different ages at different rates. Since labor at different ages is supplied with different elasticities, the government in general makes use of the capital income tax for this reason.

2.2 Quantitative Work

The theoretical discussion has argued that in a model in which household labor supply and consumption change over the life cycle, if the government cannot condition the tax function on age it optimally uses the capital income tax to mimic age-dependent labor income taxes. The theoretical discussion was carried out restricting the government to proportional taxes, as is common in the literature on optimal taxation in the Ramsey tradition. However, our analysis of the model with exogenous labor supply has demonstrated that progressive labor income taxes can also be used as a tool for implicitly taxing households of different ages at different rates. Of course, if labor supply is endogenous, progressive taxes have adverse incentive effects. Therefore it is an open question, requiring the following quantitative analysis, to what extent the capital income tax and to what extent the progressive labor income tax is being used to tax labor income at different ages at different rates.

2.2.1 The Model without Life Cycle Elements (M4)

We start by documenting results for the model described in the previous section (that is, abstract from any idiosyncratic risk, household heterogeneity and life cycle elements, apart from finite life). We recalibrate all parameters to attain the same empirical targets as before. The calibrated $\beta = 0.979$ is very similar to model M1, the same model with exogenous labor supply.

The calibrated parameters imply that for all τ_k the interest rate in equilibrium is such that $\beta(1+r(1-\tau_k)) > 1$. Therefore the life cycle consumption and leisure profiles are upward sloping, and labor supply declines over the life cycle, even in the absence of explicit life cycle model elements. Therefore, as our discussion of the theory above suggests, labor at older ages

[15] The derivation of this result is contained in Appendix C

should be taxed less heavily than labor at younger ages. In the absence of age-dependent labor income taxes a positive capital income tax plays the role. The capital income tax is quantitatively significant, at about 20%. It is, however, only about half the size as in the benchmark model, for reasons to be explained below. The labor income tax is essentially proportional, at a rate of 17%. This is not surprising given that the model does not contain any insurance or redistributive role of taxation.[16]

2.2.2 The Model with Life Cycle Elements (M5)

Now we introduce all the life cycle elements into the model, most notably social security, mortality risk, and age-dependent labor productivity. These model elements strongly affect the optimal life cycle profile of consumption, saving and leisure.[17] Since now labor efficiency and hence earnings are increasing over the working age and the government provides social security, for given time discount factor households save less.[18] Thus, in order to achieve the same empirical target the newly calibrated time discount factor β increases to $\beta = 1.009$. This in turn generates consumption and leisure profiles that are more strongly upward sloping, relative to model M4. For leisure (labor supply) this effect is somewhat mitigated by the fact that labor productivity is growing until about age 50, which, ceteris paribus, shifts labor supply towards later ages. Overall however, labor supply differs more by age than in model M4. As a consequence the optimal capital income tax, inducing indirect age-dependent leisure taxation, is used more heavily. The optimal capital income tax amounts to 34%, with a labor income tax that is essentially proportional with a marginal rate of 14%.[19]

2.2.3 Taxes and Redistribution (M6)

We model ex-ante heterogeneity among households (as opposed to ex-post heterogeneity due to different stochastic labor productivity realizations) as permanent labor productivity differences. Given our social welfare function the government has a motive to redistribute between households of different ability types. If low productivity households were simply scaled-down versions of high-productivity households (that is, if the model were completely homothetic), then a proportional capital income tax is completely ineffective tool to redistribute. In our model the only modest sources of nonhomotheticity are the lump-sum redistribution of bequests and the social security system with its flat benefits (but also remember that payroll

[16] Note that with a flat labor efficiency profile and declining labor over the life cycle labor income is moderately declining over the life cycle. Thus a progressive labor income tax does allow the government to tax older households that supply less labor at lower rates. The optimal tax code is very slightly progressive as a result, but not substantially so because of the labor supply distortions that high marginal rates create.

[17] Mortality risk also implies that the time discount factor has to be adjusted by the conditional probability of survival. For households of working age this adjustment is close to irrelevant, but it becomes quantitatively important for elderly households. This model feature also helps to generate the falling portion of the hump-shaped consumption profile over the life cycle as mortality-adjusted time discount rates decline with age.

[18] Mortality risk induces some extra precautionary saving, but the effect is quantitatively small given that the social security system provides effective insurance against longevity risk.

[19] With an upward sloping labor efficiency profile it is even harder for the government to use progressive labor income taxation to tax labor income of older households at lower rates than of younger households, since these households have higher labor earnings despite the fact that they work less.

taxes are capped at an upper income level $\bar{y}$). Therefore it is to be expected that the capital income tax is not used (in fact, cannot be used effectively) for redistributive purposes between ex-ante different households. In addition, a progressive labor income tax that taxes high-earnings households at higher rates is exactly the appropriate tool for redistribution of this type.[20] Our quantitative analysis confirms this suspicion. Adding type heterogeneity and recalibrating all parameters[21] does not change the results for the optimal capital income tax in a significant way, relative to the model without type heterogeneity. The optimal capital income tax falls slightly, to 32%. The main change from before is that now substantial progressivity in the labor income tax is optimal, in order to redistribute from high earnings to low earnings households. One way to see this is to notice that a_0, the marginal tax rate for highest earnings households in the Gouveia-Strauss tax function, increases from 14% to 18%, mainly to finance lower tax rates of low-earnings households.

2.2.4 Idiosyncratic Income Risk (M7)

Introducing uninsurable idiosyncratic labor productivity risk has two main effects. First, it increases precautionary saving (already present because of mortality risk), and thus lowers the time discount factor needed to match the target wealth-to-output ratio. The required β falls moderately, from $\beta = 1.009$ to $\beta = 1.005$. This, other things being equal, flattens desired (labor) leisure profiles and thus calls for more moderate capital income taxation. Second, a progressive labor income tax can act as partial substitute for missing insurance markets. Thus we would expect the labor income tax to become more progressive as a result of the introduction of idiosyncratic labor productivity risk, with higher marginal rates at the high end of the earnings distribution and higher capital income taxation to pay for lower rates for households with bad earnings realizations.

The resulting optimal capital income tax is 35%, with a labor income tax schedule that has a highest marginal rate of $a_0 = 23\%$ and is essentially a flat tax with a sizeable deduction of close to \$10,000. Thus while the optimal size of the capital income tax is not significantly affected by the presence of idiosyncratic risk, the optimal shape of the labor income tax responds strongly, since it is the labor income tax that can provide effective partial insurance against labor productivity and thus labor earnings risk.

2.2.5 The Role of Borrowing Constraints with Endogenous Labor Supply

As Hubbard and Judd (1987) and Imrohoroglu (1998) explain (and we reviewed in the section on exogenous labor supply), in the presence of borrowing constraints it may be optimal to tax capital somewhat instead of taxing labor at high rates, since young households at the borrowing constraint have only labor income, and labor income is low when young, thus high labor income taxes automatically mean low consumption for the young. Not stressing the potential importance of tight borrowing constraints for the optimal capital income tax result was an important omission of the previous version of the paper, although fortunately (for

[20] In fact, if the government were allowed and able to condition taxes on type, it should do so. Designing a progressive income tax is a partial substitute for type-dependent taxation.

[21] The required parameter values are essentially identical to those from the previous model (absent redistributive social security and accidental bequests this is a theorem).

us) not a quantitatively serious one, as the results below indicate. With exogenous labor supply and proportional labor income taxes the use of capital income taxes was optimal in order to prevent taxing young households with low labor income, which tend to be borrowing constrained, at too high rate. We showed that once progressive labor income taxes are allowed, the labor income tax can effectively take care of this problem, with no need to resort to capital income taxation. However, with endogenous labor supply the highly progressive labor income tax found optimal with exogenous labor supply is clearly suboptimal. It is therefore conceivable that, despite having access to progressive labor income taxes the presence of a tight borrowing constraint provides a partial rationale for capital income taxes.

To gauge the importance of tight borrowing constraints for our results we compare the results of our benchmark model with tight borrowing constraints to the results of model M7, which is identical to the benchmark model but has natural borrowing limits.[22] We find very similar results for both versions of the model.[23] The optimal capital income tax is $\tau_k = 35\%$ with the loose natural borrowing constraint and $\tau_k = 36\%$ with the tight borrowing constraint. In both cases the labor income tax is essentially a flat income tax with a sizeable deduction. So while a tight borrowing constraint increases the optimal capital income tax, the effect is of minor quantitative importance. To some degree this result does not come as a surprise, given the reported life cycle asset profiles (see Figure 1 in this and the previous version of the paper) for the benchmark model with tight constraints. That figure indicated that optimal asset holdings are positive for almost all households; only few young households (those with bad productivity shocks) actually hit the borrowing constraint.[24] Consequently, when we relax the constraints, our result, while not identical to the case with tight borrowing constraints, are not significantly affected.

[22]Strictly speaking, whenever there is a positive probability of dying at each age the loosest borrowing limit that is consistent with no default in equilibrium is our strict borrowing limit $a' \geq 0$ at all ages. In our exercises we ignore default due to premature death and define the age-dependent natural borrowing constraint as follows. First, define the minimal after-tax resources that an agent working full time has available

$$\begin{aligned} y_{\min}(i,j) &= SS + (1+r)Tr - T \text{ if } j \geq j_r \\ y_{\min}(i,j) &= (1-\tau_{ss})\alpha_i \varepsilon_j \eta_{\min} w + (1+r)Tr - T \text{ if } j < j_r \end{aligned}$$

In the last period of life the household evidently cannot borrow at all, thus

$$a'_J \geq \bar{A}(i,J) = 0$$

The natural borrowing limits for all other ages and then given recursively by

$$\bar{A}(i,j) = \frac{-y_{\min}(i,j) + \bar{A}(i,j+1)}{1+r}.$$

[23]We again recalibrate the β. In the model with tight borrowing constraints a $\beta = 1.001$ is needed to match the target. With tight constraints households cannot borrow and thus, for a given β, aggregate saving increases, relative to the case with natural borrowing limits. Consequently β falls, relative to the case with loose borrowing limits.

[24]This result in turn hinges on the calibrated time discount factor β. Had the time discount factor been significantly lower, more people would be borrowing constrained, and the quantitative role of the borrowing constraint for the capital income tax would have been larger. However, with such lower β the aggregate wealth level in the economy is counterfactually low, relative to our empirical target.

3 A Side Remark on Transitional Dynamics

We characterize the optimal tax system in the steady state. In many optimal policy analyses focusing on steady state welfare is problematic since it by construction ignores the transitional costs associated with the economy converging to the steady state. For example, in many papers on social security reform towards funded systems the new steady state is characterized by a substantially higher capital stock, and thus higher wages and consumption of households. Ignoring the transition in which this higher capital stock is built up through higher investment (and thus lower consumption) may lead to a bias towards the policy favoring a higher steady-state capital stock.

While we are fully aware of this issue (in fact, is was the main point of our paper Conesa and Krueger, 1999), we like to point out that our optimal policy is not subject to this critique. The substantially positive capital income tax that we found optimal for the reasons explained above leads to a capital stock (and capital-output ratio) that is substantially lower than the one associated with the status quo tax system. In other words, along the transition path the capital stock falls and is partly consumed by transitional generations.

While it is computationally infeasible to compute the entire optimal tax transition (which would involve maximizing over entire sequences of policy parameters, for each sequence computing an entire transition of the economy), we conducted some experiments intended to mimic such an exercise. In particular, restricting the tax code to a proportional capital income tax and a proportional labor income tax plus deduction, we asked what is the optimal once-and for all tax reform. That is, starting at the status quo the marginal labor and capital income tax rate (τ_l, τ_k) are chosen once and for all, and the deduction in every period along the transition and in the steady state adjusts to guarantee budget balance.

In order to define optimality we again have to define a social welfare function that now incorporates the wellbeing of transitional generations. In order to highlight the role of the transition we used a utilitarian social welfare function among all generations *currently alive*. That is, we define social welfare as

$$\int v(s) d\Phi_0(s)$$

where Φ_0 is the cross-sectional distribution in the initial steady state with the status quo (that is, Gouveia-Strauss estimated) tax code and v is the remaining lifetime utility of an agent with current state s, under the policy (and associated factor prices) along the transition path.

The results we obtained confirm our expectation that our steady state results provide a conservative lower bound for the optimal capital income tax. The optimal capital income tax is $\tau_k = 65\%$, with the labor income tax at $\tau_l = 10\%$, together with a sizeable deduction. The implied steady state capital-output ratio is even lower than under our optimal steady state tax result. Transitional generations do not bear the full burden of the lower capital stock and hence lower wages (since the drop in both takes time) but benefit from a larger share of output being available for consumption (since net investment is negative along the transition). As a result the tax system is even more strongly tilted towards capital income taxation.

We do not want to over-emphasize this result, since this exercise stops short of providing a fully optimal tax reform (tax rates are set only once, at the beginning of the reform). Also, the government only values households currently alive and ignores welfare of generations born in the future. Furthermore, the literature on optimal capital taxation focuses on the optimal capital income tax in the long run, which we want to compare ourselves to. With this exercise we simply want to confirm that, for the application at hand, the restriction on maximizing steady state welfare is not subject to the well-known bias found in other applications.

4 A Remark on Infinitely Lived Models with Idiosyncratic Risk

Aiyagari (1995) argues that uninsurable idiosyncratic risk alone (possibly in conjunction with tight borrowing limits) provides a rationale for positive capital income taxation even in the steady state of a model with infinitely lived households. In contrast to most of the public finance literature in his benchmark model the government also chooses optimally the level of government consumption which enters the households' utility function additively. In contrast to our work the government also does not face any restriction on how much government debt to issue.

When we repeat our analysis in an infinite horizon model with idiosyncratic risk and tight borrowing constraints, as Aiyagari (1994, 1995) we find consistently high capital income *subsidies* for a variety of model specifications and parameterizations investigated. The optimal capital subsidy was in the order of 45% in the Aiyagari model with endogenous labor supply, calibrated to the same targets as our benchmark model. Results in the model with exogenous labor supply yields similar results. In all of the many specifications we considered the optimal capital income tax was significantly negative. These results are consistent with the theoretical analysis in Davila et al. (2007), and confirm our earlier findings from the life cycle model that the presence of uninsured idiosyncratic risk alone does not provide a rationale for positive capital income taxes in our environment.

Appendix A: Derivation of Labor Income Tax Rates in the Separable Case

The first order conditions of the Ramsey problem with respect to labor are

$$\begin{aligned} \mu_t &= \chi\left(1-l_{1,t}\right)^{-\sigma_2}\left[1+\lambda_t\left(1+\frac{l_{1,t}\sigma_2}{1-l_{1,t}}\right)\right] \\ \theta\mu_{t+1} &= \frac{\beta}{\varepsilon}\chi\left(1-l_{2,t+1}\right)^{-\sigma_2}\left[1+\lambda_t\left(1+\frac{l_{2,t+1}\sigma_2}{1-l_{2,t+1}}\right)\right] \end{aligned}$$

and thus

$$\frac{1+\lambda_t\left(1+\sigma_2\frac{l_{1,t}}{1-l_{1,t}}\right)}{1+\lambda_t\left(1+\sigma_2\frac{l_{2,t+1}}{1-l_{2,t+1}}\right)} = \frac{\beta\mu_t}{\varepsilon\theta\mu_{t+1}}\left(\frac{(1-l_{2,t+1})}{(1-l_{1,t})}\right)^{-\sigma_2} \tag{15}$$

From the households first order conditions for labor we obtain

$$\frac{1-\tau_{l2,t+1}}{1-\tau_{l1,t}} = \frac{1}{\varepsilon}\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} * \left(\frac{(1-l_{2,t+1})}{(1-l_{1,t})}\right)^{-\sigma_2} \tag{16}$$

Plugging in for $\left(\frac{(1-l_{2,t+1})}{(1-l_{1,t})}\right)^{-\sigma_2}$ from (15) in (16) yields

$$\frac{1-\tau_{l2,t+1}}{1-\tau_{l1,t}} = \frac{1}{\varepsilon}\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} * \frac{\varepsilon\theta\mu_{t+1}}{\beta\mu_t}\frac{1+\lambda_t\left(1+\sigma_2\frac{l_{1,t}}{1-l_{1,t}}\right)}{1+\lambda_t\left(1+\sigma_2\frac{l_{2,t+1}}{1-l_{2,t+1}}\right)} \tag{17}$$

Finally using the fact that (from combining (12) and (13))

$$\frac{\theta\mu_{t+1}}{\beta\mu_t} = \left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{\sigma_1}$$

expression (17) simplifies to

$$\frac{1-\tau_{l2,t+1}}{1-\tau_{l1,t}} = \frac{1+\lambda_t\left(1+\sigma_2\frac{l_{1,t}}{1-l_{1,t}}\right)}{1+\lambda_t\left(1+\sigma_2\frac{l_{2,t+1}}{1-l_{2,t+1}}\right)}$$

which is equation (14) in the main text.

Appendix B: Nonzero Capital Income Tax in the Separable Case

Adding Lagrange multiplier η_t to the constraint

$$\frac{U_{l_{1,t}}}{U_{c_{1,t}}} = \frac{U_{l_{2,t}}}{\varepsilon U_{c_{2,t}}} \tag{18}$$

and re-deriving the first order conditions with respect to $c_{1,t}$ and $c_{2,t+1}$ yields

$$\mu_t = c_{1,t}^{-\sigma 1}\left(1+\lambda_t(1-\sigma_1)+\eta_t\chi\frac{\sigma_1(1-l_{2,t})^{-\sigma_2}}{c_{1,t}}\right) \tag{19}$$

$$\theta\mu_{t+1} = \beta c_{2,t+1}^{-\sigma_1}\left(1+\lambda_t(1-\sigma_1)-\eta_{t+1}\chi\frac{\sigma_1(1-l_{1,t+1})^{-\sigma_2}}{c_{2,t+1}}\right) \tag{20}$$

$$\mu_t = \theta\mu_{t+1}(1+r) \tag{21}$$

and thus

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{-\sigma_1} = \beta(1+r) * \frac{\left(1+\lambda_t(1-\sigma_1)+\eta_t\chi\frac{\sigma_1(1-l_{2,t})^{-\sigma_2}}{c_{1,t}}\right)}{\left(1+\lambda_t(1-\sigma_1)-\eta_{t+1}\chi\frac{\sigma_1(1-l_{1,t+1})^{-\sigma_2}}{c_{2,t+1}}\right)}$$

The Euler equation of the household problem reads as before

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

Combining the two yields, in the steady state

$$\frac{1+r(1-\tau_k)}{1+r} = \frac{\left(1+\lambda(1-\sigma_1)+\eta\chi\frac{\sigma_1(1-l_2)^{-\sigma_2}}{c_1}\right)}{\left(1+\lambda(1-\sigma_1)-\eta\chi\frac{\sigma_1(1-l_1)^{-\sigma_2}}{c_2}\right)} \tag{22}$$

and thus $\tau_k \neq 0$ as long as $\eta \neq 0$, that is, as long as the constraint (18) is binding in the steady state. The sign of τ_k depends on the sign of η, that is, it depends on whether an unconstrained social planner would like to set $\tau_1 > \tau_2$ or $\tau_1 < \tau_2$. As argued in the main text, if an unconstrained planner would set $\tau_1 > \tau_2$, she can mimic this by choosing $\tau_1 = \tau_2$ and $\tau_k > 0$.

Appendix C: Analytical Derivations for the Nonseparable Case

The first order conditions of the Ramsey problem with respect to consumption at both ages and capital read as

$$\mu_t = \frac{\gamma\left[c_{1,t}^{\gamma}(1-l_{1,t})^{1-\gamma}\right]^{1-\sigma}}{c_{1,t}}\left(1+\lambda_t(1-\sigma)\left[\gamma-\frac{(1-\gamma)l_{1,t}}{1-l_{1,t}}\right]\right) \tag{23}$$

$$\theta\mu_{t+1} = \frac{\beta\gamma\left[c_{2,t+1}^{\gamma}(1-l_{2,t+1})^{1-\gamma}\right]^{1-\sigma}}{c_{2,t+1}}\left(1+\lambda_t(1-\sigma)\left[\gamma-\frac{(1-\gamma)l_{2,t+1}}{1-l_{2,t+1}}\right]\right) \tag{24}$$

$$\mu_t = \theta\mu_{t+1}(1+r)$$

Combining these equations yields.

$$\frac{\frac{\left[c_{1,t}^{\gamma}(1-l_{1,t})^{1-\gamma}\right]^{1-\sigma}}{c_{1,t}}}{\frac{\left[c_{2,t+1}^{\gamma}(1-l_{2,t+1})^{1-\gamma}\right]^{1-\sigma}}{c_{2,t+1}}} * \frac{\left(1+\lambda_t(1-\sigma)\left[\gamma-\frac{(1-\gamma)l_{1,t}}{1-l_{1,t}}\right]\right)}{\left(1+\lambda_t(1-\sigma)\left[\gamma-\frac{(1-\gamma)l_{2,t+1}}{1-l_{2,t+1}}\right]\right)} = \beta(1+r)$$

Using this result together with the first order condition of the household problem (6) yields, similar to equation (22),

$$\frac{1+r(1-\tau_{k,t+1})}{1+r} = \frac{\left(1+\lambda_t(1-\sigma)\left[\gamma-\frac{(1-\gamma)l_{2,t+1}}{1-l_{2,t+1}}\right]\right)}{\left(1+\lambda_t(1-\sigma)\left[\gamma-\frac{(1-\gamma)l_{1,t}}{1-l_{1,t}}\right]\right)} \tag{25}$$

and in general $\tau_{k,t+1} \neq 0$ even if labor income taxes can be conditioned on age. In particular, suppose $\sigma > 1$, then $\tau_{k,t+1} > 0$ if and only if $l_{1,t} > l_{2,t+1}$, that is, if labor supply declines over the life cycle.[25]

The first order conditions of the Ramsey problem with respect to labor supply at young age are given as follows.

[25] Note that from equations (23) and (24) both the numerator and denominator on the right hand side of (25) have to be positive.

$$\mu_t = -\left[1 + \lambda_t(1-\sigma)(\gamma - \frac{(1-\gamma)l_{1,t}}{1-l_{1,t}}) + \frac{\lambda_t}{1-l_{1,t}}\right] U_{l_{1,t}}$$

Likewise the first order condition with respect to $l_{2,t+1}$ becomes

$$\theta\mu_{t+1} = -\frac{\beta}{\epsilon} * \left[1 + \lambda_t(1-\sigma)(\gamma - \frac{(1-\gamma)l_{2,t+1}}{1-l_{2,t+1}}) + \frac{\lambda_t}{1-l_{2,t+1}}\right] U_{l_{2,t+1}}$$

Combining these with the household optimality conditions with respect to the labor supply, after some algebra we obtain

$$\frac{1-\tau_{l1,t}}{1-\tau_{l2,t+1}} = \frac{1 + \frac{\lambda_t}{1+\lambda_t(1-\sigma)\gamma - (1+\lambda_t(1-\sigma))l_{2,t+1}}}{1 + \frac{\lambda_t}{1+\lambda_t(1-\sigma)\gamma - (1+\lambda_t(1-\sigma))l_{1,t}}}$$

We observe that the optimal age-dependent labor income tax schedule is determined by the optimal labor supply over the life cycle. As long as $1 + \lambda(1-\sigma) > 0$, in the steady state we have that $\tau_{l_1} > \tau_{l_2}$ if and only if $l_1 > l_2$, that is, a declining labor supply profile implies labor income taxes that decline with age.[26]

Combining two results we conclude the following: with nonseparable preferences not only the age-dependent labor tax but also the capital income tax depends on the optimal labor allocation over the life cycle. That is, even *with* age-dependent labor income taxes the optimal capital income tax is $\tau_k \neq 0$ in general, and is strictly positive if $\sigma > 1$ and the labor supply profile is declining with age, $l_1 > l_2$.

Finally, the fact that if the government is restricted to age-independent labor income taxes the capital income tax can be used to mimic it is still valid. The intertemporal tax wedge with respect to labor supply is given by

$$\frac{1-\tau_{l1,t}}{1-\tau_{l2,t+1}}(1 + r(1-\tau_{k,t+1})) = (1+r)\frac{1 + \lambda_t(1-\sigma)(\gamma - \frac{(1-\gamma)l_{2,t+1}}{1-l_{2,t+1}}) + \frac{\lambda_t}{1-l_{2,t+1}}}{1 + \lambda_t(1-\sigma)(\gamma - \frac{(1-\gamma)l_{1,t}}{1-l_{1,t}}) + \frac{\lambda_t}{1-l_{1,t}}}$$

and thus a given wedge can be implemented with age-dependent labor income taxes or capital income taxes.

[26]The condition $1 + \lambda(1-\sigma) > 0$ is surely satisfied for $\sigma < 1$, but may (or not) be satisfied for $\sigma > 1$.

Model	Endog. Labor Supply	Tight Borrow. Const.	Type Hetero.	Idio. Income	Life Cycle	β	r	τ_k	τ_l	Prog. Labor Tax
M1	No	No	No	No	No	0.983	4.5%	10%	19%	No
M2	No	No	No	No	Yes	1.001	3,2%	-24%	100%	Yes
M3	No	Yes	No	Yes	Yes	1.001	4.3%	-34%	100%	Yes
M4	Yes	No	No	No	No	0.979	4.7%	20%	17%	No
M5	Yes	No	No	No	Yes	1.009	5.6%	34%	14%	No
M6	Yes	No	Yes	No	Yes	1.009	5.2%	32%	18%	Yes
M7	Yes	No	Yes	Yes	Yes	1.005	5.6%	35%	23%	Yes
Benchmark	Yes	Yes	Yes	Yes	Yes	1.001	5.6%	36%	23%	Yes

Table 1: Summary of Quantitative Results

References

[1] Aiyagari, R.,1994, Uninsured Idiosyncratic Risk and Aggregate Saving, *Quarterly Journal of Economics* 109, 659-684.

[2] Aiyagari, R.,1995, Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting, *Journal of Political Economy* 103, 1158-1175.

[3] Atkeson, A., V.V. Chari and P. Kehoe, 1999, Taxing Capital Income: A Bad Idea, *Federal Reserve Bank of Minneapolis Quarterly Review*, 23, 3-17.

[4] Blanchard, O. and S. Fischer, 1989, *Lectures on Macroeconomics*, MIT Press, Cambridge, MA.

[5] Conesa, J.C. and D. Krueger, 1999, Social Security Reform with Heterogeneous Agents, *Review of Economic Dynamics*, 2, 757-795.

[6] Davila, J., J. Hong, P. Krusell and V. Rios-Rull, 2007, "Constrained Efficiency in the Neoclassical Growth Model with Uninsurable Idiosyncratic Shocks," *Working paper*, University of Pennsylvania.

[7] Erosa, A. and M. Gervais, 2002, Optimal Taxation in Life Cycle Economies, *Journal of Economic Theory* 105, 338-369.

[8] Garriga, C., 2003, Optimal Fiscal Policy in Overlapping Generations Models, *Working Paper.*

[9] Green, E. and R. Zhou, 2002, "Dynamic Monetary Equilibrium in a Random Matching Economy, Econometrica, 70, 929-970.

[10] Hansen, G and S. Imrohoroglu, 2006, Consumption and Hours Over the Life Cycle: The Role of Annuities, *Working paper*, University of Southern California.

[11] Hubbard, G. and K. Judd, 1986, Liquidity Constraints, Fiscal Policy, and Consumption, *Brookings Papers on Economic Activity*, 1986, 1-50

[12] Imrohoroglu, S., 1998, A Quantitative Analysis of Capital Income Taxation, *International Economic Review*, 39, 307-328.