Taxing Capital? Not a Bad Idea After All!*

Juan Carlos Conesa
Universitat Autonoma de Barcelona

Sagiri Kitao
New York University

Dirk Krueger
Goethe University Frankfurt, CEPR and NBER

June 20, 2006

Abstract

In this paper we quantitatively characterize the optimal capital and labor income tax in a lifecycle model with idiosyncratic, uninsurable income shocks, where households also differ permanently with respect to their labor productivity. The welfare criterion we employ is ex-ante (before ability is realized) expected (with respect to uninsurable productivity shocks) utility of a newborn in a stationary equilibrium. Embedded in this welfare criterion is a concern of the policy maker for insurance against idiosyncratic shocks and redistribution between agents of different ability. Such insurance and redistribution can be achieved by progressive labor income taxes or taxation of capital income, or both. The policy maker has then to trade off these concern against the standard distortions of these taxes for the labor supply and capital accumulation decision.

We find that the optimal capital income tax rate is not only positive, but is significantly positive. The optimal (marginal and average) tax rate on capital of is $\tau_k = 36\%$, in conjunction with a progressive labor income tax code that is, to a first approximation, a flat tax of 23% with a deduction that corresponds to about $6,000 (relative to an average income of households in the model of $35,000). In our model how much the government should tax capital and labor depends crucially on how elastic capital accumulation and labor supply of middle-aged individuals are with respect to the respective tax rates. These households are both highly productive in their jobs and in the middle of accumulating retirement wealth. They supply labor quite elastically, whereas their savings are fairly inelastic with respect to the marginal capital income tax rate. As a corollary, the capital income tax is high; in fact, substantially higher at the margin than the labor income tax.

Keywords: Progressive Taxation, Capital Taxation, Optimal Taxation

J.E.L. classification codes: E62, H21, H24

*We thank participants of the MEA conference on OLG models, UAB and UPF for many useful suggestions. The authors can be reached at juancarlos.conesa@uab.es, sagiri.kitao@nyu.edu and dirk.krueger@wiwi.uni-frankfurt.de.

1 Introduction

Should the government tax capital income? The seminal contributions of Chamley (1986) and Judd (1985) argue that standard economic theory provides a negative answer to this question. The government should not tax capital, at least not in the long run. The survey articles by Chari and Kehoe (1999) and Atkeson, Chari and Kehoe (1999) argue that this result is robust to a relaxation of a number of stringent assumptions made by Chamley and Judd.

Chamley and Judd derive their result under the assumptions that households are infinitely lived and face no risk (either aggregate or idiosyncratic), or equivalently, can fully insure against idiosyncratic risk and trade a full set of Arrow securities against aggregate uncertainty. If, on the other hand, idiosyncratic risk is not insurable, Aiyagari (1995) suggests that positive capital taxation may be optimal, in order to cure the overaccumulation of capital as a result of precautionary savings behavior by households.[1] His quantitative results suggests, however, that the optimal capital income tax is small.[2] Even if insurance markets are complete, or equivalently households face no idiosyncratic risk, Hubbard and Judd (1997) demonstrate that financial market frictions in the form of borrowing constraints may make the taxation of capital income desirable.

Both the original Chamley-Judd result as well as its response by Aiyagari relied on models with infinitely lived agents. Characterizing the structure of optimal taxes in a model that explicitly models the life cycle of households in an overlapping generations economy, Erosa and Ventura (2002) and Garriga (2003) demonstrate that the optimal capital income tax can be, and in general is different from zero, at least if the tax code is anonymous in that the tax schedule a household faces is not allowed to depend on the age of the household. It is an open question, however, how large the optimal capital income tax, relative to the optimal labor income tax is in a realistically calibrated life cycle model in which households face borrowing constraints and idiosyncratic income risk in the same order of magnitude as in the data.

The goal of this paper is therefore to quantitatively characterize the optimal capital and labor income tax in a model that nests both model elements previously identified in the literature as having potential for generating positive capital income taxes: imperfect insurance against idiosyncratic income shocks due to missing insurance markets and borrowing constraints, as well as an explicit life cycle structure. In our model households differ according to their age and their history of income realizations. In addition, we allow agents to be heterogenous with respect to their initial ability to generate income, modelled as a fixed effect in their labor productivity. To the extent that society values

[1] Central to Aiyagari's (1995) result is that the Ramsey government does not have to respect the fact that insurance markets against idiosyncratic income risk are incomplete. We will discuss this point in detail below.

[2] Golosov et al. (2003) also argue, in a dynamic private information economy with idiosyncratic income shocks, for an optimal capital income tax rate that is ex-post different from zero, but still equal to zero in expectation for each household.

an equitable distribution of welfare this model element induces a positive role for taxes that redistribute from more to the less able households.

In order to determine the optimal tax system in our model with rich cross-sectional heterogeneity we need to take a stand on the social welfare function employed in evaluating policies. The welfare criterion we employ is ex-ante (before ability is realized) expected (with respect to uninsurable productivity shocks) utility of a newborn in a stationary equilibrium. Embedded in this welfare criterion is a concern of the policy maker for insurance against idiosyncratic shocks and redistribution between agents of different ability, since taking an extra dollar from the highly able and giving it to the less able, ceteris paribus, increases social welfare since the value function characterizing lifetime utility is strictly concave in ability to generate income.[3] Such insurance and redistribution can be achieved by progressive labor income taxes or taxation of capital income (which mainly accrues to the wealthy), or both. The policy maker has then to trade off these concern against the standard distortions of these taxes for the labor supply and capital accumulation decision.

We find that the optimal capital income tax rate is not only positive, but is significantly positive. The optimal (marginal and average) tax rate on capital of is $\tau_k = 36\%$, in conjunction with a progressive labor income tax code that is, to a first approximation, a flat tax of 23% with a deduction that corresponds to about $\$6,000$ (relative to an average income of households in the model of $\$35,000$). The intuition for this result is as follows: how much the government should tax capital and labor depends crucially on how elastic capital accumulation and labor supply are with respect to their corresponding marginal tax rate. In our life cycle economy those contributing most to tax revenue are middle-aged individuals which are both highly productive in their jobs (and hence have high labor income) and in the middle of accumulating savings for retirement (and therefore pay the bulk of the capital income tax bill). But these agents supply labor quite elastically, whereas their saving choices (which at their age is mainly life cycle saving rather than precautionary saving due to idiosyncratic income shocks) is fairly inelastic with respect to the marginal capital income tax rate.[4] As a corollary, the capital income tax is substantial; in fact, substantially higher at the margin than the labor income tax.

Since one would expect our findings to depend crucially on the exact specification of household preferences with respect to leisure (and thus the labor supply elasticity) with respect to leisure, we investigate how sensitive our results are with respect to this specification. Replacing the Cobb-Douglas utility specification between consumption and leisure which is often used in macroeconomics (and which we therefore employ as a benchmark, but which implies a

[3] Of course redistribution and insurance are two sides of the same coin: what is redistribution between households of different abilities ex post (after ability is realized) is insurance against low ability ex ante (before birth).

[4] Saez (2003) carries out an empirical investigation into the link between marginal taxes and income elasticity of the rich. His estimated elasticities are in line with the elasticities we compute in our model.

Note that in models where hosueholds live forever the life cycle savings motive, crucial in our model, is absent by construction.

rather high labor supply elasticity) with a preference specification which implies labor supply elasticities consistent with the micro evidence (for males) delivers optimal tax rates on capital which are somewhat lower, but still significantly different from zero. In particular, the optimal capital income tax falls to $\tau_k = 21\%$, and the optimal labor income tax schedule is roughly a flat tax of 34% with deduction of now $\$9,000$. Thus our main finding of a significant capital income tax and a flat labor income tax with sizeable deduction is robust, but not surprisingly the exact mix between taxing capital and labor income shifts towards higher labor income taxes with lower labor supply elasticities.

We then revisit Aiyagari's (1995) claim that one justification of a positive capital income tax is the presence of uninsured idiosyncratic labor productivity risk. In order to isolate the role of this model element for our result of significantly positive capital income taxes we characterize the optimal tax code in a version of the model that abstracts from idiosyncratic risk, and also eliminates ex-ante differences in ability. The resulting model is nothing else but the classic Auerbach and Kotlikoff (1987) large scale OLG model in which households face borrowing constraints. For this specification we find that the optimal labor income tax schedule is not progressive anymore, and that capital income taxes are even *higher* than in our benchmark model. While the optimal marginal (and average) labor income tax rate is 11.4% the corresponding capital income tax rate is 41.8%. Hence the presence of uninsurable idiosyncratic risk *reduces* the optimal capital income tax, in contrast to the findings by Aiyagari (1995). The key difference between his thought experiment and ours is that we maintain throughout the restriction that insurance markets against idiosyncratic risk are absent, whereas Aiyagari (1995) compares the market equilibrium to what a social planner would choose that does not have to obey the restriction on allocations that missing insurance markets against idiosyncratic uncertainty impose. A paper that makes a very related point is Davila et al. (2005) which suggests that a social planner that has to respect the condition that any allocation she may choose has to be implementable with the given incomplete market structure would likely opt for a *higher* capital stock than arising in the market equilibrium. Their analysis suggests what we verify in our work: that capital income taxes are *lower* in the presence of uninsurable idiosyncratic income risk, compared to a model where this risk is absent.

Finally we demonstrate that even in our model it is possible to generate optimal capital income taxes close to zero. However this result emerges only in the rather uninteresting (and arguably unrealistic) case in which the government accumulates so much negative debt (that is, it owns assets) in the steady state that it can finance almost all government outlays by interest earned on these assets. In such a circumstance there is little need to generate any tax revenue, and thus little need to raise revenue from capital income taxes.[5]

Besides contributing to the large literature on the optimal size of the capital income tax, our study is related to the literature on optimal taxation more

[5] This is still a nontrivial result since it is conceivable that positive labor income taxes would be used to finance subsidies for capital accumulation.

broadly, and to the optimal progressivity of the income tax code in particular. Mirrlees (1971) characterizes the optimal tax code when the policy maker faces a trade-off between providing efficient incentives for household labor supply and achieving an equitable after-tax income distribution. The studies by Mirrlees (1974) and Varian (1980), recently extended to an environment in which households can save by Reiter (2004), replace the policy maker's concern for equity by an insurance motive; by making after-tax incomes less volatile, a progressive tax system may provide partial income insurance among ex-ante identical households and thus may be called for even in the absence of ex-ante heterogeneity of households and a public desire for equity.

We follow the tradition of this literature that explicitly models the policy makers concerns for equity and insurance, and its trade-off with providing the right incentives for savings and labor supply decisions, but take a quantitative approach. Previously, this strategy was adopted by Altig et al. (2001), Ventura (1999), Castañeda et al. (1999), Domeij and Heathcote (2001) and Nishiyama and Smetters (2005) in their positive analysis of fundamental tax reforms. On the normative side, the contributions by Bohacek and Kejak (2004) and Conesa and Krueger (2006) characterize the optimal progressivity of the income tax code, without allowing this tax code to differentiate between labor and capital income. As such these papers cannot directly contribute to the discussion about the optimal size of the capital income tax when capital taxes are an alternative tool to provide redistribution/insurance.[6] In work that is complementary to ours Smyth (2005) allows differential tax treatments of labor and capital income and characterizes the (potentially nonlinear) tax system that maximizes a weighted sum of lifetime utility of all agents alive in the steady state. Since in his world households are identical at birth, by construction his analysis also does not capture a potentially positive, purely redistributive motive (in the sense used in this paper) for capital and progressive labor taxation, but rather only its insurance aspect.

The paper is organized as follows. In the next section we lay out the economic environment and define equilibrium. Section 3 discusses the calibration of the model and section 4 explains the optimal tax experiments we are implementing in the calibrated model. Results from our benchmark model are presented in section 5, and section 6 contains a sensitivity analysis of our results with respect to the importance of uninsurable idiosyncratic income risk and our utility specification with respect to leisure. Finally, section 7 concludes the paper.

[6] Conesa and Krueger (2006) find an optimal tax code that is roughly a flat tax with generous deduction and thus comes close to the proposal of Hall and Rabushka (1995). Saez (2002) studies the optimal size of the deduction (and thus the optimal progressivity of the tax code) within the restricted set of flat tax systems with deduction.

2 The Economic Environment

The model we use is an extended version of the one used in Conesa and Krueger (2006), augmented to allow for a meaningful distinction between capital and labor income taxation.

2.1 Demographics

Time is discrete and the economy is populated by J overlapping generations. In each period a continuum of new agents is born, whose mass grows at a constant rate n. Each agent faces a positive probability of death in every period. Let $\psi_j = prob(\text{alive at } j+1|\text{alive at } j)$ denote the conditional survival probability from age j to age $j+1$. At age J agents die with probability one, i.e. $\psi_J = 0$. Therefore, even in the absence of altruistic bequest motives, in our economy a fraction of the population leaves (unintended) bequests. These are denoted by Tr_t and redistributed in a lump-sum fashion across individuals currently alive. At a certain exogenous age j_r agents retire and receive social security payments SS_t, which are financed by proportional labor income taxes $\tau_{ss,t}$, up to an income threshold $\bar{y}$ above which no further payroll taxes are paid.

2.2 Endowments and Preferences

Individuals are endowed with one unit of productive time in each period of their lives and enter the economy with no assets, besides transfers emanating from accidental bequests. They spend their time supplying labor to a competitive labor market or consuming leisure.

Individuals are heterogeneous along three dimensions that affect their labor productivity and hence their wage. First, agents of different ages differ in their average, age-specific labor productivity ε_j, which will govern the average wage of an age cohort. Retired agents (those with age $j \geq j_r$) by assumption are not productive at all, $\varepsilon_j = 0$. Second, since we do not model differences in ability, education choices or other factors that affect earnings potentials explicitly, we introduce, as a second source of heterogeneity, group-specific differences in productivity. We assume that agents are born as one of M possible ability types $i \in \mathbf{I}$, and that this ability does not change over an agents' lifetime, so that agents, after the realization of their ability, differ in their current and future earnings potential. The probability of being born with ability α_i is denoted by $p_i > 0$. This feature of the model, together with a social welfare function that values equity, gives a welfare-enhancing role to redistributive fiscal policies.

Finally, workers of same age and ability face idiosyncratic uncertainty with respect to their individual labor productivity. Let by $\eta_t \in \mathbf{E}$ denote a generic realization of this idiosyncratic labor productivity uncertainty at period t. The stochastic process for labor productivity status is identical and independent across agents and follows a finite-state Markov chain with stationary transitions over time, i.e.

$$Q_t(\eta, E) = Prob(\eta_{t+1} \in E | \eta_t = \eta) = Q(\eta, E). \tag{1}$$

We assume that Q consists of only strictly positive entries which assures that there exists a unique, strictly positive, invariant distribution associated with Q which we denote by Π. All individuals start their life with average stochastic productivity $\bar{\eta} = \sum_\eta \eta \Pi(\eta)$, where $\bar{\eta} \in \mathbf{E}$. Different realizations of the stochastic process then give rise to cross-sectional productivity, income and wealth distributions that become more dispersed as a cohort ages. In the absence of explicit insurance markets for labor productivity risk a progressive tax system may be an effective, publicly administered tool to share this idiosyncratic risk across agents.

At any given time individuals are characterized by (a_t, η_t, i, j), where a_t are asset holdings (of one period, risk-free bonds), η_t is stochastic labor productivity status at date t, i is ability type and j is age. An agent of type (a_t, η_t, i, j) deciding to work ℓ_j hours commands pre-tax labor income $\varepsilon_j \alpha_i \eta_t \ell_j w_t$, where w_t is the wage per efficiency unit of labor. Let by $\Phi_t(a_t, \eta_t, i, j)$ denote the measure of agents of type (a_t, η_t, i, j) at date t.

Preferences over consumption and leisure $\{c_j, (1 - \ell_j)\}_{j=1}^J$ are assumed to be representable by a standard time-separable utility function of the form:

$$E \left\{ \sum_{j=1}^{J} \beta^{j-1} u(c_j, 1 - \ell_j) \right\}, \tag{2}$$

where β is the time discount factor. We discuss the exact form of the period utility function u below. Expectations are taken with respect to the stochastic processes governing idiosyncratic labor productivity and the time of death.

2.3 Technology

We assume that the aggregate technology can be represented by a standard Cobb-Douglas production function. The aggregate resource constraint is given by:

$$C_t + K_{t+1} - (1 - \delta) K_t + G_t \leq A K_t^\alpha N_t^{1-\alpha} \tag{3}$$

where K_t, C_t and N_t represent the aggregate capital stock, aggregate consumption and aggregate labor input (measured in efficiency units) in period t, and α denotes the capital share. The calibration constant A normalizes units in our economy[7], and the depreciation rate for physical capital is denoted by δ. As standard with a constant returns to scale technology and perfect competition, without loss of generality, we assume the existence of a representative firm operating this technology.

[7] We decided to abstract from technological progress, since we will be considering preference specifications that are not consistent with the existence of a balanced growth path, but allow us to endow households with a labor supply elasticity consistent in magnitude with microeconometric evidence.

2.4 Government Policy

The government engages in three activities in our economy: it absorbs resources as government spending, it levies taxes and it runs a balanced budget social security system. The social security system is defined by benefits SS_t for each retired household, independent of that household's earnings history. Social security taxes are levied up to a maximum labor income level $\bar{y}$. The payroll tax rate $\tau_{ss,t}$ is set to assure period-by-period budget balance of the system. We take the social security system as exogenously given and not as subject of optimization of the policy maker.

Furthermore the government faces a sequence of exogenously given government consumption $\{G_t\}_{t=1}^{\infty}$ and has three fiscal instruments to finance this expenditure. First it levies a proportional tax $\tau_{c,t}$ on consumption expenditures, which we also take as exogenously given in our analysis. Second, the government taxes capital income of households, $r_t(a + Tr_t)$, at a constant marginal capital tax rate $\tau_{K,t}$. Here r_t denotes the risk free interest rate, a denotes asset held by the household, and Tr_t, denote transfers from accidental bequests. Finally, the government can tax each individual's taxable labor income according to some, potentially progressive, labor income tax schedule T. Define as

$$yp_t = w_t \alpha_i \varepsilon_j \eta \ell_t$$

a household's pre-tax labor income, where w_t denotes the wage per efficiency unit of labor. A part of this pre-tax labor income is accounted for by the part of social security contributions paid by the employer

$$ess_t = 0.5\tau_{ss,t} \min\{yp_t, \bar{y}\}$$

which is not part of taxable income under the current U.S. tax code. Thus we define as taxable labor income

$$y_t = \begin{cases} yp_t - ess_t & \text{if } j < j_r \\ 0 & \text{if } j \geq j_r \end{cases}$$

We impose the following restrictions on labor and capital income taxes. First, tax rates cannot be personalized as we are assuming anonymity of the tax code. Second, the capital income tax is a proportional tax, as described above. Labor income taxes can be made an arbitrary function of individual taxable labor income in a given period. We denote the tax code by $T(\cdot)$, where $T(y)$ is the labor total income tax liability if taxable labor income equals y. Our investigation of the optimal tax code then involves finding the labor income tax *function* T and the capital tax rate τ_K that maximizes social welfare, defined by a particular social welfare function specified below.

Finally, notice that we do not allow for government debt. We will maintain this assumption both in the benchmark economy and in our baseline scenario for finding the optimal tax schedules. We postpone the introduction of government debt to the sensitivity analysis and the discussion of our results in section 6.3.

2.5 Market Structure

We assume that workers cannot insure against idiosyncratic labor income uncertainty by trading explicit insurance contracts. Also annuity markets insuring idiosyncratic mortality risk are assumed to be missing. However, agents trade one-period risk free bonds to self-insure against the risk of low labor productivity in the future. The possibility of self-insurance is limited, however, by the assumed inability of agents to sell the bond short; that is, we impose a stringent borrowing constraint upon all agents. In the presence of survival uncertainty, this feature of the model prevents agents from dying in debt with positive probability.[8]

2.6 Definition of Competitive Equilibrium

In this section we will define a competitive equilibrium and a stationary equilibrium. Individual state variables are individual asset holdings a, individual labor productivity status η, individual ability type i and age j. The aggregate state of the economy at time t is completely described by the joint measure Φ_t over asset positions, labor productivity status, ability and age.

Therefore let $a \in \mathbf{R}_+$, $\eta \in \mathbf{E} = \{\eta_1, \eta_2, ..., \eta_n\}$, $i \in \mathbf{I} = \{1, \ldots, M\}$, $j \in \mathbf{J} = \{1, 2, ...J\}$, and let $\mathbf{S} = \mathbf{R}_+ \times \mathbf{E} \times \mathbf{J}$. Let $\mathbf{B}(\mathbf{R}_+)$ be the Borel σ -algebra of $\mathbf{R}_+$ and $\mathbf{P}(\mathbf{E})$, $\mathbf{P}(\mathbf{I})$, $\mathbf{P}(\mathbf{J})$ the power sets of $\mathbf{E}, \mathbf{I}$ and $\mathbf{J}$, respectively. Let $\mathbf{M}$ be the set of all finite measures over the measurable space $(\mathbf{S}, \mathbf{B}(\mathbf{R}_+) \times \mathbf{P}(\mathbf{E}) \times \mathbf{P}(\mathbf{I}) \times \mathbf{P}(\mathbf{J}))$.

Definition 1 *Given a sequence of social security replacement rates* $\{b_t\}_{t=1}^{\infty}$, *consumption tax rates* $\{\tau_{c,t}\}_{t=1}^{\infty}$ *and government expenditures* $\{G_t\}_{t=1}^{\infty}$ *and initial conditions* K_1 *and* Φ_1, *a competitive equilibrium is a sequence of functions for the household,* $\{v_t, c_t, a'_t, \ell_t : \mathbf{S} \to \mathbf{R}_+\}_{t=1}^{\infty}$, *of production plans for the firm,* $\{N_t, K_t\}_{t=1}^{\infty}$, *government labor income tax functions* $\{T_t : \mathbf{R}_+ \to \mathbf{R}_+\}_{t=1}^{\infty}$, *capital taxes* $\{\tau_{K,t}\}_{t=1}^{\infty}$, *social security taxes* $\{\tau_{ss,t}\}_{t=1}^{\infty}$ *and benefits* $\{SS_t\}_{t=1}^{\infty}$, *prices* $\{w_t, r_t\}_{t=1}^{\infty}$, *transfers* $\{Tr_t\}_{t=1}^{\infty}$, *and measures* $\{\Phi_t\}_{t=1}^{\infty}$, *with* $\Phi_t \in \mathbf{M}$ *such that:*

1. *given prices, policies, transfers and initial conditions, for each* t, v_t *solves the functional equation (with* c_t, a'_t *and* ℓ_t *as associated policy functions):*

$$v_t(a, \eta, i, j) = \max_{c, a', \ell}\{u(c, \ell) + \beta\psi_j \int v_{t+1}(a', \eta', i, j+1) Q(\eta, d\eta')\} \qquad (4)$$

[8] If agents were allowed to borrow up to a limit, it may be optimal for an agent with a low survival probability to borrow up to the limit, since with high probability she would not have to pay back this debt. Clearly, such strategic behavior would be avoided if lenders could provide loans at different interest rates, depending on survival probabilities (i.e. age). In order to keep the asset market structure simple and tractable we therefore decided to prevent agents from borrowing altogether, very much in line with much of the incomplete markets literature in macroeconomics; see Aiyagari (1994) or Krusell and Smith (1998) for representative examples.

subject to:[9]

$$c+a' = w_t\varepsilon_j\alpha_i\eta\ell - \tau_{ss,t}\min\{w_t\varepsilon_j\alpha_i\eta\ell,\bar{y}\} + (1+r_t(1-\tau_{K,t}))(a+Tr_t), \quad for \;\; j < j_r, \tag{5}$$

$$c+a' = SS_t + (1+r_t(1-\tau_{K,t}))(a+Tr_t) - T_t[y_t], for \;\; j \geq j_r, \tag{6}$$

$$a' \geq 0, c \geq 0, 0 \leq \ell \leq 1. \tag{7}$$

2. *Prices w_t and r_t satisfy:*

$$r_t = \alpha A\left(\frac{N_t}{K_t}\right)^{1-\alpha} - \delta, \tag{8}$$

$$w_t = (1-\alpha)A\left(\frac{K_t}{N_t}\right)^{\alpha}. \tag{9}$$

3. *The social security policies satisfy*

$$\tau_{ss,t}\int \min\{w_t\alpha_i\varepsilon_j\eta\ell_t,\bar{y}\}\Phi_t(da\times d\eta\times di\times dj) = SS_t\int \Phi_t(da\times d\eta\times di\times\{j_r,...,J\}). \tag{10}$$

4. *Transfers are given by:*

$$Tr_{t+1} = \int (1-\psi_j)a'_t(a,\eta,i,j)\Phi_t(da\times d\eta\times di\times dj) \tag{11}$$

5. *Government budget balance:*

$$\begin{aligned} G_t \;\; = \;\; & \int \tau_{K,t}r_t(a+Tr_t)\Phi_t(da\times d\eta\times di\times dj) + \\ & \int T_t[y_t]\Phi_t(da\times d\eta\times di\times dj) + \\ & \tau_{c,t}\int c_t(a,\eta,i,j)\Phi_t(da\times d\eta\times di\times dj) \end{aligned} \tag{12}$$

6. *Market clearing:*

$$K_t = \int a\Phi_t(da\times d\eta\times di\times dj) \tag{13}$$

$$N_t = \int \varepsilon_j\alpha_i\eta\ell_t(a,\eta,i,j)\Phi_t(da\times d\eta\times di\times dj) \tag{14}$$

$$\int c_t(a,\eta,i,j)\Phi_t(da\times d\eta\times di\times dj) + \int a'_t(a,\eta,i,j)\Phi_t(da\times d\eta\times di\times dj) + G_t = AK_t^{\alpha}N_t^{1-\alpha} + (1-\delta)K_t \tag{15}$$

[9] Taxable labor income y_t was defined above.

7. *Law of Motion:*

$$\Phi_{t+1} = H_t(\Phi_t) \tag{16}$$

where the function $H_t : \mathbf{M} \rightarrow \mathbf{M}$ *can be written explicitly as:*

(a) *for all* $\mathcal{J}$ *such that* $1 \notin \mathcal{J}$*:*

$$\Phi_{t+1}(A\times E\times\mathcal{I}\times\mathcal{J}) = \int P_t((a,\eta,i,j);A\times E\times\mathcal{I}\times\mathcal{J})\Phi_t(da\times d\eta\times di\times dj) \tag{17}$$

where

$$P_t((a,\eta,i,j);A\times E\times\mathcal{I}\times\mathcal{J}) = \begin{cases} Q(e,E)\psi_j & \text{if } a'_t(a,\eta,i,j) \in A, i \in \mathcal{I}, j+1 \in \mathcal{J} \\ 0 & \text{else} \end{cases} \tag{18}$$

(b)

$$\Phi_{t+1}((A\times E\times\mathcal{I}\times\{1\}) = (1+n)^t \begin{cases} \sum_{i\in\mathcal{I}} p_i & \text{if } 0 \in A, \bar{\eta} \in E \\ 0 & \text{else} \end{cases} \tag{19}$$

Definition 2 *A stationary equilibrium is a competitive equilibrium in which per capita variables and functions as well as prices and policies are constant, and aggregate variables grow at constant gross growth rate* $(1 + n)$*.*

3 Functional Forms and Calibration of the Benchmark Economy

In order to carry out the numerical determination of the optimal tax code in our model we first have to choose a model parameterization. We now describe our choices to that effect.

3.1 Demographics

In our model households are born at age twenty, corresponding to model age 1. They become unproductive and hence retire at model age 46 (age 65 in real time) and die with probability 1 at model age 81 (age 100 in the real world). The population grows at an annual rate of $n = 1.1\%$, the long-run average in the U.S. Finally our model requires conditional survival probabilities from age j to age $j+1$, ψ_j, which we take from the study by Bell and Miller (2002). Table 1 summarizes our choices of demographic parameters.

Table I: Demographics Parameters

Parameter	Value	Target
Retir. Age: j_r	46 (65)	Compul. Ret. (assumed)
Max. Age: J	81 (100)	Certain Death (assumed)
Surv. Prob. ψ_j	Bell and Miller (2002)	Data
Pop. Growth: n	1.1%	Data

3.2 Preferences

Households have time-separable preferences over consumption and leisure and discount the future with factor β. Because our results, and especially the intuition for our results, will point to the labor supply elasticity as an important determinant of our findings we consider two specifications of the period utility function. As benchmark we assume a standard Cobb-Douglas specification:

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

where γ is a share parameter determining the relative importance of consumption, and σ determines the risk aversion of the household.[10] We set $\sigma = 4$ and choose β and γ such that the stationary equilibrium of the economy with benchmark tax system (as described below) features a capital-output ratio of 2.7 and an average share of time worked of one-third of the time endowment (which we normalized to 1).[11] The resulting preference parameters are summarized in Table 2.

Table II: Preferences Parameters

Parameter	Value	Target
β	1.001	$K/Y = 2.7$
σ	4.0	Fixed
γ	0.377	Avg Hours= $\frac{1}{3}$

This preference specification has been criticized as implying a labor supply elasticity that is too high relative to what empirical studies estimate from labor market data (see e.g. Browning et al., 1999). Our benchmark preference specification implies a Frisch labor supply elasticity considered to be high relative to some microeconometric estimates. In the literature the Frisch elasticity is meant to capture the magnitude of the substitution effect. In Blundell and MaCurdy (1999) the Frisch elasticity is defined as the elasticity of labor supply with respect to the wage holding constant the marginal utility of wealth. In our case it takes a value around 1, while in some other applications it is computed as the elasticity of labor supply holding constant the level of consumption (in our case since preferences are non-separable in consumption and leisure it gives a different value, around 2). Usually the microeconometric studies restrict attention to white males of prime age already employed and obtained values smaller than one.

[10] The coefficient of relative risk aversion is given by

$$-\frac{cu_{cc}}{u_c} = \sigma\gamma + 1 - \gamma$$

which should be kept in mind when interpreting our parameter choices.

[11] It is understood that in a general equilibrium model like ours all parameters affect all equilibrium quantities and prices. In our discussion of the calibration we associate a parameter with that equilibrium entity it affects most, in a quantitative sense.

It is not obvious what the relevant labor supply elasticity should be, given that in our theoretical environment the decision unit is a household. It seems reasonable to think that the labor supply elasticity of a household is higher than that of an individual, because of both higher labor supply elasticities of females and the existence of an extensive margin that is not usually considered in the computation of the labor supply elasticities. Heckman (1993) argues that the elasticity of participation decisions is large, in fact most of the movement in aggregate hours worked is due to this margin. Also, Imai and Keane (2004) argue that the individual intertemporal elasticity of substitution in labor supply is higher than usually estimated in a framework with endogenous human capital accumulation (i.e. learning-by-doing), in fact as high as 3.82. Domeij and Floden (2006) have shown both theoretically and empirically that the presence of uninsurable labor income risk and borrowing constraints biases the estimated individual labor supply elasticities downwards. Finally, Kimball and Shapiro (2005) use preferences homothetic in hours worked (rather than in leisure) where the substitution and income effects exactly cancel each other and obtain a Frisch labor supply elasticity around 1, which is the one we have in our benchmark economy.

Notice also that the previous discussion refers to the Frisch labor supply elasticity, that measures only the substitution effect. With our benchmark preferences households with zero wealth would not change hours worked in reaction to changes in the wage (or its marginal tax rate), while the labor supply elasticity would be higher the higher the wealth of the household.

Absent more convincing measurement, we will also consider an alternative preference specification that allows us to precisely match the empirically estimated individual elasticities. In this alternative specification intratemporal preferences are represented by:

$$u(c, 1-\ell) = \frac{c^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-\ell)^{1-\sigma_2}}{1-\sigma_2} \tag{20}$$

We discuss the calibration of the curvature parameters σ_1, σ_2 and the share parameter χ below.

3.3 Labor Productivity Process

Households start their life with no assets beyond the transfers induced by unintended bequests from those deceased at the end of last period. In addition, they are endowed with one unit of time in each period. If households work they have a labor productivity that depends on three components: a deterministic age-dependent component ε_j, a type-dependent fixed effect α_i and a stochastic, persistent, idiosyncratic shock. Thus the natural logarithm of wages of an individual is given by

$$\log(w_t) + \log(\varepsilon_j) + \log(\alpha_i) + \log(\eta)$$

The age-productivity profile $\{\varepsilon_j\}_{j=1}^{jr-1}$ is taken from Hansen (1993). We consider two ability types, with equal population mass $p_i = 0.5$ and fixed effects

$\alpha_1 = e^{-\sigma_\alpha}$ and $\alpha_2 = e^{\sigma_a}$, so that $E(\ln(\alpha_i)) = 0$ and $Var(\ln(\alpha_i)) = \sigma_\alpha^2$. Furthermore, we specify the stochastic process for the idiosyncratic part of log-wages as a discretized version, with seven states, of a simple $AR(1)$ process with persistence parameter ρ and unconditional variance σ_η^2. This choice gives us three free parameters to choose. With their choice we target three statistics from data measuring how cross-sectional labor income dispersion evolves over the life cycle. In particular, Storesletten et al. (2004) document that i) at cohort age 22 the cross-sectional variance of labor income is about 0.2735, ii) at age 60 it is about 0.9 and iii) that it increases roughly linearly in between. In our model labor supply and therefore labor earnings are endogenous, responding optimally to the labor productivity process. We choose the three parameters $(\sigma_\alpha^2, \rho, \sigma_\eta^2)$ so that in the benchmark parameterization the model displays a cross-sectional age-earnings variance profile consistent with the three facts just cited. The implied parameter values for our benchmark preference specification are summarized in Table 3. Note that, evidently, these parameters have to be re-calibrated if the alternative preference specification is being used.

Table III: Labor Productivity

Parameter	Value	Target
σ_α^2	0.14	$Var(y_{22})$
ρ	0.98	Lin. Incr. in $Var(y_j)$
σ_η^2	0.0289	$Var(y_{60})$

3.4 Technology

The production side of our model is completely standard. Therefore the capital share parameter α in the Cobb-Douglas production function is set to the empirical capital share, $\alpha = 0.36$, as standard in the literature, and the depreciation rate is set to match an investment-output ratio of 25.5% in the data (where investment includes nonresidential and residential fixed investment as well as investment into consumer durables). This requires $\delta = 8.3\%$. Technology parameters are summarized in Table 4.

Table IV: Technology Parameters

Parameter	Value	Target
α	0.36	Data
δ	8.33%	$I/Y = 25.5\%$
A	1	Normalization

3.5 Government Policies and the Income Tax Function

The government spends money, collects tax revenues and operates a social security system. The focus of our analysis of the government is the income tax code.

We therefore take the other parts of government activity as exogenously given and calibrate the extent of these activities to observed data. We calibrate government spending G such it accounts for 17% of GDP in the initial stationary equilibrium. Note that we keep G constant across our tax experiments; therefore if an income tax system different from the one specified as benchmark delivers higher output in equilibrium, the corresponding $\frac{G}{Y}$ ratio in that equilibrium is reduced.

Part of tax revenues are generated by a proportional consumption tax, whose size we take as exogenous to our analysis. We set $\tau_c = 5\%$, following Mendoza et al. (1994). In addition to taxes and spending the government runs a pay-as-you-go social security system, defined by a payroll tax. The payroll tax takes a value of 12.4% of labor income up to an upper bound of \$87,000. Benefits are then determined by budget balance of the social security system in the benchmark economy.

We want to determine the optimal income tax function. Ideally one would pose no restrictions on the set of potential tax functions the government can choose from. Maximization over such an unrestricted set is computationally infeasible, however. Therefore we restrict the set of tax functions to a flexible three parameter family. If y is taxable income (either labor income or capital income), then total taxes paid on that income is given by

$$T^{GS}(y; a_0, a_1, a_2) = a_0 \left(y - (y^{-a_1} + a_2)^{-\frac{1}{a_1}} \right) \tag{21}$$

where (a_0, a_1, a_2) are parameters. Note for $a_1 \rightarrow 0$ the tax system reduces to a pure flat tax system, while other parameterizations encompass a wide range of progressive and regressive tax functions.

Without discriminating between capital and labor income Gouveia and Strauss (1994) estimate the parameters (a_0, a_1, a_2) that best approximate actual taxes paid under the actual US income tax system of $a_0 = 0.258$ and $a_1 = 0.768$. We use as benchmark tax system the one implied by their estimates, applied to the sum of labor and capital income. The parameter a_2 is then used to insure government budget balance.[12] The benchmark tax system is summarized in Table 5.

Table V: Policy Parameters

Parameter	Value
τ_c	5%
a_0	0.258
a_1	0.768
τ_{ss}	12.4%

[12] Note that the parameter a_2 is not invariant to units of measurement: if one scales all variables by a fixed factor, one has to adjust the parameter a_2 in order to preserve the same tax function.

4 The Computational Experiment

Once our model is fully parameterized we can determine the optimal tax code. For this we need to specify the set of tax functions considered and the objective function of the government. Define y_l and y_k as taxable labor and capital income, respectively. The set of tax functions we consider is given by

$$\mathcal{T} = \left\{T_l(y_l), T_k(y_k) : T_l(y_l) = T^{GS}(y; a_0, a_1, a_2) \text{ and } T_k(y_k) = \tau_k y_k\right\}$$

and thus by the four parameters (a_0, a_1, a_2, τ_k), out of which we will maximize over three and use a_2 to adjust in order to insure budget balance. That is, we allow for a flexible labor income tax code, but restrict capital taxes to be proportional, an assumption that assures both computational feasibility and makes the comparison to existing studies employing the same assumption easier. Also note that the choices of (a_0, a_1, τ_k) are restricted by the requirement that there has to exist a corresponding a_2 that balances the budget.

The remaining ingredient of our analysis is the social welfare function ranking different tax functions. We assume that the government wants to maximize the ex-ante lifetime utility of an agent being born into a stationary equilibrium implied by the chosen tax function. Formally the government's objective is given by

$$\begin{aligned} SWF(a_0, a_1, \tau_k) &= \int v_{(a_0,a_1,\tau_k)}(a = 0, \eta = \bar{\eta}, i, j = 1) d\Phi_{(a_0,a_1,\tau_k)} \\ &= \sum_{i=1}^{2} v_{(a_0,a_1,\tau_k)}(0, \bar{\eta}, i, 1) \end{aligned}$$

where we used the facts that the two types are of equal mass and everyone starts life with no financial assets and at the average stochastic labor productivity level. Here $v_{(a_0,a_1,\tau_k)}$ and $\Phi_{(a_0,a_1,\tau_k)}$ are the value function and invariant cross-sectional distribution associated with tax system characterized by (a_0, a_1, τ_k).

5 Results

5.1 The Optimal Tax System

We determine as optimal tax system a (marginal and average) tax rate on capital of $\tau_k = 36\%$ and a labor income tax characterized by the parameters $a_0 = 23\%$ and $a_1 \approx 7$. This implies that the labor income tax code is basically a flat tax with marginal rate of 23% and a deduction of about \$6,000 (relative to an average income of \$35,000). Consequently we find that taxing capital is not only a good idea, but taxing it substantially and more heavily than labor income is optimal for a government that is benevolent and maximizes a utilitarian social welfare function.

We performed several exercises trying to evaluate whether it would be welfare enhancing to introduce progressivity of the capital income tax schedule as well, by introducing a deduction. We found optimal that all the progressivity of the tax code be reflected exclusively on the labor income tax schedule.

5.2 Comparison with the Benchmark

In order to assess the relevance of the tax code for equilibrium allocations in our model and to obtain a first understanding for the causes of high capital income taxes we now compare selected equilibrium statistics for the optimal and the benchmark tax system. Table 6 contains a summary of the basic findings.

Table VI: Comparison across Tax Codes

Variable	BENCH.	OPTIMAL
Average Hours Worked	0.333	−0.56%
Total Labor Supply N	−−	−0.11%
Capital Stock K	−−	−6.64%
Output Y	−−	−2.51%
Aggregate Consumption C	−−	−1.63%
Gini Coef. for Wealth	0.636	0.659
Gini Coef. for Consumption	0.273	0.269
ECV	−−	1.33%

We observe that in the optimal tax system capital drops substantially below the level of the benchmark economy. Consequently aggregate output and aggregate consumption fall as well. This is an immediate consequence of the heavy tax on capital income in the optimal tax system, relative to the benchmark (where the highest marginal tax rate is 25.8%). The change in taxes also induces adjustments in labor supply, which are quite small in the aggregate however.

5.2.1 Decomposition of the Welfare Effects

Given the substantial decline in aggregate consumption and the modest decline in average hours worked in the optimal tax system, relative to the benchmark, it is at first sight surprising that the optimal tax system features substantially higher aggregate welfare, equivalent to an increase of 1.33% of consumption at all ages, and all states of the world, keeping labor supply allocations unchanged. Therefore it is useful to decompose these welfare gains into several components. Given the form of the utility function, the steady state welfare consequences of switching from a consumption-labor allocation (c_0, l_0) to (c_*, l_*) are given by

$$CEV = \left[\frac{W(c_*, l_*)}{W(c_0, l_0)}\right]^{\frac{1}{\gamma(1-\sigma)}} - 1$$

where $W(c, l) = SWF(a_0, a_1, \tau_k)$ is the expected lifetime utility at birth of a household, given a tax system (a_0, a_1, τ_k). We can decompose CEV into two components, one stemming from the change in consumption from c_0 to c_*, and one from the change in leisure. Furthermore, the consumption impact on welfare can be further divided into a part that captures the change in average consumption, and one part that reflects the change in the distribution of consumption (across types, across the life cycle and across states of the world). The same is true for labor supply (leisure).[13]

Table presents the results of this decomposition. It shows that, following this distribution, the welfare gains stem from a better allocation of consumption across types and states of the world, and from a reduction of the average time spent working. This more that offsets the lower average level of consumption and the fact that, due to the lower marginal tax rates, labor supply becomes more unevenly distributed.

---

[13] Let CEV_C and CEV_L be defined as

$$\begin{aligned} W(c_*, l_0) &= W(c_0(1+CEV_C), l_0) \\ W(c_*, l_*) &= W(c_*(1+CEV_L), l_0). \end{aligned}$$

Then it is easy to verify that

$$\begin{aligned} 1+CEV &= (1+CEV_C)(1+CEV_L) \text{ or} \\ CEV &\approx CEV_C + CEV_L \end{aligned}$$

We further decompose CEV_C into a consumption level effect CEV_{CL} and a consumption distribution effect CEV_{CD}:

$$\begin{aligned} W(\hat{c}_0, l_0) &= W(c_0(1+CEV_{CL}), l_0) \\ W(c_*, l_0) &= W(\hat{c}_0(1+CEV_{CD}), l_0) \end{aligned}$$

where

$$\hat{c}_0 = (1+g_C)c_0 = \frac{C_*}{C_0} c_0$$

is the consumption allocation resulting from scaling the allocation c_0 by the change in aggregate consumption $\frac{C_*}{C_0}$. A simple caluclation shows that the consumption level effect simply equals the growth rate of consumption:

$$CEV_{CL} = \frac{C_*}{C_0} - 1$$

Similarly, for leisure we define

$$\begin{aligned} W(c_*, \hat{l}_0) &= W(c_*(1+CEV_{LL}), l_0) \\ W(c_*, l_*) &= W(c_*(1+CEV_{LD}), \hat{l}_0). \end{aligned}$$

where $1-\hat{l}_0$ is the leisure allocation derived from l_0 by scaling it by the change in aggregate leisure:

$$1-\hat{l}_0 = \frac{1-L_*}{1-L_0}(1-l_0).$$

Again it is easy to verify that the leisure level effect is given by

$$CEV_{LL} = (1+g_{LE})^{\frac{\gamma}{1-\gamma}} - 1.$$

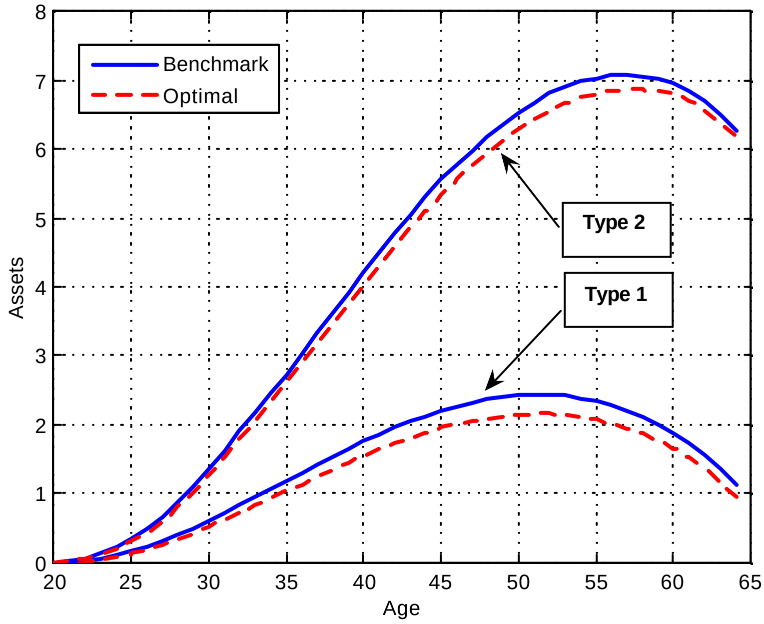

Figure 1: Asset Accumulation over the Life Cycle

Table VII: Decomposition of Welfare

Total Change		1.37%
Consumption	Total	1.19%
	Level	−1.82%
	Distribution	3.07%
Leisure	Total	0.18%
	Level	0.58%
	Distribution	−0.40%

5.2.2 Life Cycle Profiles of Assets, Labor Supply and Taxes

In order to further document who mainly bears the burden of the income tax and how a change in the tax code changes this in this section we discuss life cycle patterns of asset holdings (the relevant tax base for the capital income tax) and labor income (the relevant tax base for labor income taxes).

In figure 1 we display the average asset holdings over the life cycle for both types of households, both for the benchmark and for the optimal tax system. First, we observe the hump-shaped behavior of assets that is typical of any life cycle model. This, in particular, implies that indeed the main burden of the capital income tax is bourne by households aged 40 to 60. Second, it is clearly

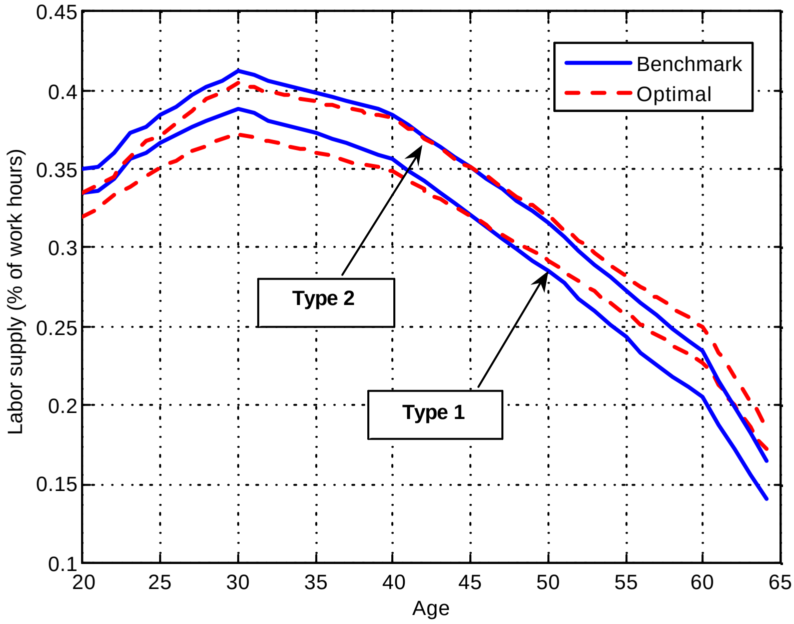

Figure 2: Labor Supply over the Life Cycle

visible how asset accumulation is affected by the higher capital income taxes implied by the optimal, relative to the benchmark tax system, most pointedly for the 40 to 60 year old. This explains the overall decline of assets and thus capital, relative to the benchmark, of 6.6%. *[Adjustments to the figure: units on the y-axes should be in dollar terms comparable to the tax figures, plot should extend above age 65]*

Figure 2 documents the average life cycle pattern of labor supply of both skill groups for the benchmark and the optimal tax code. We observe that the optimal tax code induces the life cycle pattern of labor supply to be tilted towards higher labor supply at ages at which the households are more productive. The lower labor income taxes and the sizeable deduction make an allocation of labor supply that follows more closely the age-efficiency profile optimal, as it alleviates the severity of the borrowing constraint early in life. Especially for the low-skilled group the increase in labor supply at age 50 to 60 is substantial, indicating a high labor supply elasticity with respect to marginal labor income taxes for this group. *I think it is possibly more telling to construct life cycle profiles of labor income, because its a better measure of the labor income tax base.*

Finally, figure 3 displays average taxes paid, both for the benchmark and the optimal tax code, over the life cycle. It demonstrates that the optimal tax code

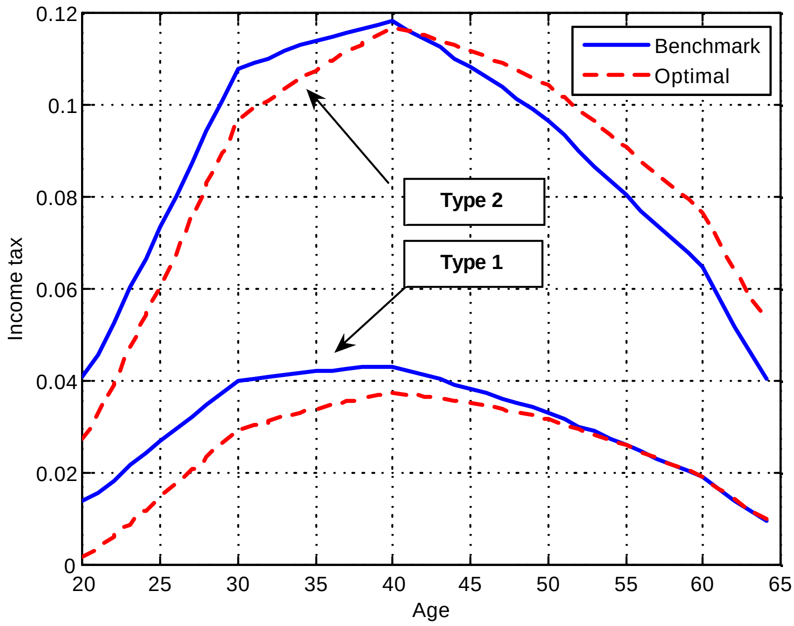

Figure 3: Taxes Paid over the Life Cycle

leads to substantially more redistribution across types, by taxing more heavily the high-skilled, high labor income-earners which also hold a large fraction of financial assets in the economy. The substantially higher capital income taxes of the optimal tax system, relative to the benchmark, explains why these wealthy individuals (see figure 1) pay a larger tax bill in the optimal tax system.

[Again, the units on the y-axis have to be made comparable to the dollar amounts for taxes we report, and the plot should extend beyond age 65]

6 Sensitivity Analysis and Interpretation of the Results

Since our results are computational we will perform several exercises and sensitivity analysis in order to understand the underlying reasons for our high capital income tax result.

6.1 The Case of Separable Preferences

Our previous argument for substantially positive capital income taxes was based on the finding that those individuals contributing most to the tax receipts of the government have a high labor supply elasticity. In this section we want to investigate whether our findings are robust to a different preference specification that allows us to control this labor supply elasticity directly. We employ a utility function of the form given in (20). We choose as parameters a coefficient of relative risk aversion of $\sigma_1 = 2$ and a $\sigma_2 = 3$. This implies a reduction in the Frisch labor supply elasticity by a factor of 3, so that now the labor supply elasticity is below one.[14] For the remaining preference parameters (β, χ) as well as the other model parameters we follow the same calibration strategy as above; Table VII summarizes the new preference parameters.[15]

Table VII: Preferences Parameters

Parameter	Value	Target
β	0.9717	$K/Y = 2.7$
σ_1	2	Fixed
σ_2	3	Fixed
χ	1.92	Avg Hours= $\frac{1}{3}$

Under this new parameterization we find as optimal tax code a marginal capital income tax of $\tau_k = 0.21$ and a marginal labor income tax rate of $a_0 = 0.34$ and $a_1 = 18$, implying again a flat tax rate on labor with deduction of now $\$9,000$. So whereas the main finding of a significant capital income tax and a flat labor income tax with sizeable deduction remains intact, the optimal tax mix shifts towards lower capital taxation and higher labor taxation, at least at the higher end of the labor income distribution .

Clearly, then, the higher the labor supply elasticity the lower the marginal labor income tax and the higher the capital income tax. Nevertheless, the optimal capital income tax is still substantial.

Table 8 repeats the comparison of aggregate statistics under the benchmark and the optimal tax system, now with the alternative preference specification. Note that both the stationary equilibrium with the benchmark as well as the optimal tax system differs from the previous section (of course not along those statistics that we calibrated to, but along most other dimensions).

[14] With this preference specification the Frisch labor supply elasticity is equal to $\frac{1}{\sigma_2} \times \frac{1-\ell}{\ell} = \frac{2}{3}$, while it was 1 in our benchmark economy.

[15] Of the other model parameters, the main changes in parameters occurred for the ones characterizing the labor productivity process; the new choices are $(\sigma^2_\alpha, \rho, \sigma^2_\eta) = (0.19, 0.995, 0.0841)$.

Table VIII: Comparison across Tax Codes

Variable	BENCH.	OPTIMAL
Average Hours Worked	0.333	0.324
Total Labor Supply N	$--$	-2.14%
Capital Stock K	$--$	-7.44%
Output Y	$--$	-4.08%
Aggregate Consumption C	$--$	-3.75%
Gini Coef. for Wealth	0.636	0.699
Gini Coef. for Consumption	0.277	0.271
ECV	$--$	3.4%

Qualitatively, the results are similar to the ones in the previous section. Quantitatively, however, the decline in the capital stock, output, consumption, and particular labor supply is more substantial than with nonseparable preferences. Also, the decline in consumption inequality is much more pronounced now than previously, suggesting that with separable preferences the motives for insurance and redistribution are even more crucial than before. Despite a much more severe drop in aggregate consumption the welfare gains are higher now than with Cobb-Douglas preferences.

6.2 The Role of Insurance and Redistribution

In order to understand what the role of insurance and redistribution is for our high capital income tax result, we redo our exercise eliminating all sources of uncertainty and inequality. In particular, everybody is born identical and both survival uncertainty and labor productivity uncertainty are eliminated. The outcome of this exercise is that now the labor income tax schedule is flat (no progressivity), and capital income taxes are even higher. In particular, with our benchmark preferences we find optimal a flat labor income tax of 11.4% and a capital income tax of 41.8%. Hence, we conclude that the reason for high capital income taxes is not the presence of uninsurable risks as conjectured in Aiyagari (1995).

We interpret our result as consistent with the findings of Davila et al. (2005) relative to the intuition in Aiyagari (1995). In a world with heterogeneity the fact that capital accumulation is higher than in the complete markets economy does not necessarily imply that there exists overaccumulation of capital from a social welfare point of view. In fact, it might be the case that it is optimal to promote capital accumulation in order to increase wages (and hence the welfare of the majority of relatively poor), and that is the rational for having lower capital income taxes in the heterogeneous agent economy.

[It seems important here to divide the results into a part that comes from redistribution and a part that comes from insurance; easiest to do is to document welfare consequences optimal vs. benchmark if idiosyncratic risk is absent, but we still have the two types.]

6.3 Preferences Homothetic in Hours Worked and the Role of Government Debt

Erosa and Gervais (2002) and Garriga (2003) proved that the optimal capital income tax in the steady state of an Overlapping Generation model is zero if taxes can be made age contingent or if preferences are homothetic in hours worked. We then redo our quantitative exercise using this type of preferences in our environment. The only difference, then, is that we have a social security system[16], that we do not allow for government debt and that our objective function (maximization of ex-ante utility of a newborn) does not clearly map into the objective function of a Ramsey problem (welfare of each subsequent generation weighted by some social planner discount factor). Our findings show that the capital income tax is still high, in the order of 25%.

However, if we allow for negative government debt, then we recover the result of a zero capital income tax. In fact, quantitatively, negative government debt has to be of the order of two times GDP. Under such a scenario also labor income taxes are very low, the government accumulates assets and uses the return on those assets to pay for government expenditures. In fact, welfare differences across alternative tax codes become trivially small since most of the government expenditure is already financed through the return on government capital.

Notice that in the quantitative results of Garriga (2003) for the non-separable case of preferences, he showed that for particular values of the social planner discount factor the optimal steady state capital income tax was zero, but then government debt was negative. We view our findings as consistent with these findings.

We conclude from this exercise that the ability of the government to run negative debt is a key ingredient for the optimality of zero capital income taxes. It is important to bear in mind that given our objective function (ex-ante utility of a newborn in the steady state), the need of the government to accumulate assets at the expense of private consumption has no welfare consequences. In Garriga (2003) only for high values of the social planner discount factor such a strategy happens to be optimal from the point of view of the planner, since the increased welfare of future generations dominates the welfare losses associated to the building of government capital during the transition, but it would not be a Pareto improving reform.

7 Conclusion

In this paper we characterize the optimal capital and labor income tax code in a large scale Overlapping Generations model where uninsurable heterogeneity and

[16] We also redid our quantitative exercise in an environment without the PAYGO social security system. In such an environment taxable income is much higher, because of higher labor supply and higher capital accumulation. As a result the optimal capital income tax was lower, but still substantial: 21%.

uncertainty generates a desire for redistribution. We found that a tax system that taxes capital heavily in the long run and taxes labor income according to a flat tax with sizeable deduction is optimal.

The key driving force behind our result is that in a life cycle model like ours labor supply especially of those with high productivity is very elastic with respect to the marginal labor income tax rate, more so than are savings of those that accumulate the most assets. These middle aged individuals save mainly for life-cycle reasons; a higher marginal capital income tax does not affect their savings behavior as drastically, as, say, in an infinite horizon model in the spirit of Aiyagari, where people save purely to smooth out unfavorable productivity shocks.

That the labor supply elasticity of highly productive agents is crucial for our result is demonstrated by our finding that employing a utility specification with lower implied elasticities increases the optimal marginal tax rates for labor income and reduces them for capital income. Nevertheless, it remains optimal to tax capital income highly, significantly different from zero.

Given our findings that the life cycle structure of our model in general, and life cycle savings behavior in particular, appear crucial for our results, future research should investigate how sensitive our findings are to a more detailed modelling of institutions affecting life-cycle savings incentives, especially the social security system and its reform. In a similar vein, so far we have abstracted for any linkage between generations due to bequest motives. It is conceivable, in the light of the classical results on optimal capital taxation in fact likely, that an incorporation of these elements into our model brings its implications for the optimal tax code closer to these classical results. Until then we conclude that taxing capital (heavily) may not be such a bad idea after all.

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] Altig, D., A. Auerbach, L. Kotlikoff, K. Smetters and J. Walliser, 2001, Simulating Fundamental Tax Reform in the U.S., *American Economic Review*, 91, 574-595.

[4] 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.

[5] Bell, F. and M. Miller, 2002, Life Tables for the United States Social Security Area 1900-2100, Office of the Chief Actuary, Social Security Administration

[6] Blundell, R. and T. MaCurdy, 1999, Labor Supply: A Review of Alternative Approaches, in *Handbook of Labor Economics*, 3A, Elsevier, 1559-1695.

[7] Berliant, M. and M. Gouveia, 1993, Equal Sacrifice and Incentive Compatible Income Taxation, *Journal of Public Economics* 51, 219-240.

[8] Bohacek, R. and M. Kejak, 2004, Optimal Government Policies in Models with Heterogeneous Agents, *mimeo*, CERGE-EI.

[9] Browning, M., L. Hansen and J. Heckman, 1999, Micro Data and General Equilibrium Models, in J. Taylor and M. Woodford (eds.) *Handbook of Macroeconomics*, Elsevier Science North-Holland, 543-633.

[10] Castañeda, A., J. Diaz-Gimenez and J.V. Rios-Rull, 1999, Earnings and Wealth Inequality and Income Taxation: Quantifying the Trade-offs of Switching to a Proportional Income Tax in the U.S., *mimeo,* University of Pennsylvania.

[11] Castañeda, A., J. Diaz-Gimenez and J.V. Rios-Rull, 2003, Accounting for Earnings and Wealth Inequality, *Journal of Political Economy* 111, 818-857.

[12] Chamley, C., 1986, Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives, *Econometrica* 54, 607-622.

[13] Chari, V.V. and P. Kehoe, 1999, Optimal Fiscal and Monetary Policy, in *Handbook of Macroeconomics*, vol. 1, edited by J.B. Taylor and M. Woodford, Amsterdam: North Holland.

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

[15] Conesa, J. and D. Krueger, 2006, On the Optimal Progressivity of the Income Tax Code, *Journal of Monetary Economics*, forthcoming.

[16] Davila, J., J. Hong, P. Krusell and J.V. Rios-Rull, 2005, Constrained Efficiency in the Neoclassical Growth Model with Uninsurable Idiosyncratic Risk, *mimeo*, University of Pennsylvania.

[17] Domeij, D. and J. Heathcote, 2004, On the Distributional Effects of Reducing Capital Taxes, *International Economic Review 45, 523-554.*

[18] Domeij, D. and M. Floden, 2006, The Labor Supply Elasticity and Borrowing Constraints: Why Estimates are Biased, *Review of Economic Dynamics* 9, 242-262.

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

[20] Garriga, C., 2003, Optimal Fiscal Policy in Overlapping Generations Models, *mimeo.*

[21] Golosov, M., N. Kocherlakota and A. Tsyvinski, 2003, Optimal Indirect and Capital Taxation, *Review of Economic Studies* 70, 569-587..

[22] Gouveia, M. and R. Strauss, 1994, Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis, *National Tax Journal 47, 317-339.*

[23] Hall, R. and A. Rabushka, 1995, *The Flat Tax, (2nd Ed.),* Stanford University Press, Stanford.

[24] Heckman, J., 1993, What Has Been Learned About Labor Supply in the Past Twenty Years?, *American Economic Review*, 83, 116-131.

[25] Hubbard, G. and K. Judd, 1987, Liquidity Constraints, Fiscal Policy, and Consumption, *Brookings Papers on Economic Activity*, 1986, 1-50

[26] Huggett, M., 1993, The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies, *Journal of Economic Dynamics and Control* 17, 953-969.

[27] Imai, S. and M.P. Keane, 2004, Intertemporal Labor Supply and Human Capital Accumulation, *International Economic Review* 45(2), 601-641

[28] Judd, K., 1985, Redistributive Taxation in a Simple Perfect Foresight Model, *Journal of Public Economics* 28, 59-83.

[29] Kimball, M.S. and M.D. Shapiro, 2005, Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?, mimeo.

[30] Krusell, P. and Smith, A., 1998, Income and Wealth Heterogeneity in the Macroeconomy, *Journal of Political Economy* 106, 867-896.

[31] Mendoza, E., A. Razin and L. Tesar, 1994, Effective Tax Rates in Macroeconomics: Cross-Country Estimates on Factor Income and Consumption, *Journal of Monetary Economics* 34, 297-323.

[32] Mirrless, J., 1971, An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies* 38, 175-208.

[33] Mirrlees, J., 1974, Notes on Welfare Economics, Information and Uncertainty, in *Essays on Economic Behavior under Uncertainty,* edited by M. Balch, D. McFadden and S. Wu, Amsterdam: North Holland.

[34] Nishiyama, S. and K. Smetters, 2005, Consumption Taxes, Risk Sharing and Economic Efficiency, *Journal of Political Economy*, 113, 1088-1115.

[35] Reiter, M., 2004, The Optimal Nonlinear Taxation of Capital in Models With Uninsurable Income Risk, *mimeo*, Universitat Pompeu Fabra.

[36] Saez, E., 2002, Optimal Progressive Capital Income Taxes in the Infinite Horizon Model, *NBER Working Paper 9046.*

[37] Saez, E., 2003, The Effect of Marginal Tax Rates on Income: A Panel Study of Bracket Creep, *Journal of Public Economics*, 87, 1231-1258.

[38] Smyth, S., 2005, A Balancing Act: Optimal Nonlinear Taxation in Overlapping Generations Models, *mimeo*, Harvard University.

[39] Storesletten, K., C. Telmer and A. Yaron, 2004, Consumption and Risk Sharing over the Life Cycle, *Journal of Monetary Economics* 51, 609-633.

[40] Tauchen, G. and R. Hussey 1991, Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models, *Econometrica* 59, 371-396.

[41] Varian, H., 1980, Redistributive Taxation as Social Insurance, *Journal of Public Economics* 14, 49-68.

[42] Ventura, G., 1999, Flat Tax Reform: A Quantitative Exploration, *Journal of Economic Dynamics and Control* 23, 1425-1458.