Taxing Capital? Not a Bad Idea After All!

Juan Carlos Conesa

Department of Economics and Economic History

Universitat Autonoma de Barcelona

Edifici B - Campus de Bellaterra

08193 Bellaterra, Cerdanyola del Vallès, Spain

Sagiri Kitao

Marshall School of Business

University of Southern California

3670 Trousdale Parkway

Los Angeles, CA 90089

Dirk Krueger

Department of Economics

University of Pennsylvania

3718 Locust Walk

Philadelphia, PA 19104

April 9, 2008

Abstract

We quantitatively characterize the optimal capital and labor income tax in an overlapping generations model with idiosyncratic, uninsurable income shocks and permanent productivity differences of households. The optimal capital income tax rate is significantly positive at 36%. The optimal progressive labor income tax is, roughly, a flat tax of 23% with a deduction of $7,200 (relative to average household income of $42,000). The high optimal capital income tax is mainly driven by the life cycle structure of the model whereas the optimal progressivity of the labor income tax is attributable to the insurance and redistribution role of the tax system.

Keywords: Progressive Taxes, Capital Taxation, Optimal Taxation

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

1 Introduction

Should the government tax capital income in the long run? The seminal contributions of Judd (1985) and Chamley (1986) answer this question negatively. Jones et al. (1997), Chari and Kehoe (1999) and Atkeson et al. (1999) show that this result is robust to a relaxation of a number of assumptions made by Chamley and Judd.

The literature has identified (at least) two modelling choices that may invalidate the zero capital income tax result in the long run. First, Hubbard and Judd (1986), Aiyagari (1995) and Imrohoroglu (1998) have emphasized that if households face tight borrowing constraints and/or are subject to uninsurable idiosyncratic income risk, then the optimal tax system will in general include a positive capital income tax. Second, Alvarez et al. (1992), Erosa and Gervais (2002) and Garriga (2003) show that in life cycle models the optimal capital income tax in general is different from zero, at least if the tax code cannot explicitly be conditioned on the age of the household.[1] It is an open question, however, how large the optimal capital income tax is, relative to the optimal labor income tax, in a realistically calibrated life cycle model in which households face borrowing constraints and idiosyncratic income risk.

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: incomplete capital markets and an explicit life cycle structure. In addition, we allow agents to be heterogenous with respect to their innate ability to generate income, modelled as a fixed effect in labor productivity. If society values an equitable distribution of welfare this model element induces a positive redistributive role for taxes. In this paper (in contrast to much of the literature) we allow the government

[1]Golosov et al. (2003) argue, in a Mirrleesian dynamic private information model 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. In a recent paper Weinzierl (2007) studies age-dependent labor income taxation in such a model. Huffman (2007) shows that the fiscal authority may find it optimal to both tax capital income and grant investment tax credits in a model with endogenous investment-specific technical change. Albanesi and Armenter (2007) give a sufficient condition such that (with optimal policy) equilibria in a class of models with infinitely lived households do not feature intertemporal distortions in the long run.

to use progressive taxes, and we will demonstrate that the government's desire to tax capital may depend on whether it has access to progressive labor income taxes.

In order to determine the optimal tax system we need to take a stand on the social welfare function to evaluate policies. The welfare criterion we employ is ex-ante (before ability is realized) expected (with respect to idiosyncratic shocks) lifetime 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 across households with different ability, since transferring an extra dollar from the highly able to the less able, ceteris paribus, increases social welfare since the value function characterizing lifetime utility is strictly concave in the ability to generate income.[2] Such insurance and redistribution can be achieved by progressive labor income taxes or taxation of capital income, or both. The policy maker then has to trade off this concern against the standard distortions these taxes impose on labor supply and capital accumulation decisions.

We find that the optimal capital income tax is significantly positive at a rate of 36%. The associated progressive labor income tax code is, to a first approximation, a flat tax of 23% with a deduction of $7,200 (relative to a GDP per capita of $42,000). What explains this tax structure to be optimal? As Erosa and Gervais (2002) and Garriga (2003) show theoretically, in life cycle models with endogenous

[2] One interpretation of our social welfare function is ex-ante (before the realization of permanent productivity differences) lifetime utility of a household to be born into the steady state. This interpretation avoids the problem of comparing lifetime utility across different households. What we call redistribution is, under this interpretation, insurance against low ability.

labor supply it is typically optimal to tax labor at different ages at different rates. In the absence of age-dependent labor income taxes a positive capital income tax allows the government to achieve the same, as can a progressive labor income tax.[3] Furthermore, in the presence of upward sloping life cycle earnings profiles and tight borrowing constraints it may be suboptimal to tax labor earnings of the young too heavily, as Hubbard and Judd (1986) and Imrohoroglu (1998) suggest. Again, this can be circumvented by either using the capital income tax more heavily or making the labor income tax suitably progressive. With a sequence of thought experiments that sequentially shut down certain elements of the model we show, in section 6, that endogenous labor supplied differentially over the life cycle is crucial in driving the high capital income tax results whereas market incompleteness and distributional concerns are mainly responsible for shaping the progressive labor income tax.

In an extensive sensitivity analysis we document that our results are qualitatively, and to a large part quantitatively, robust to a lower labor supply elasticity, to allowing the capital income tax code to be progressive, to alternative values of the intertemporal elasticity of substitution and to alternative specifications of the social welfare function. The introduction of government debt deserves a more qualified statement. We show in section 6.3.5 that only in the rather extreme case in which the government can accumulate so substantial *negative* government debt that it can finance almost all government outlays by interest earned on these assets, the optimal

[3] For the preferences used in this paper the desire to tax age-dependently can in turn be related to age-varying labor supply elasticities. We discuss this further in section 6.2.

capital income tax is zero (and the optimal labor income tax is close to zero as well).

Our study is related to the literature on the optimal progressivity of the income tax code. Mirrlees (1971), Mirrlees (1974), Varian (1980), Benabou (2002) and Reiter (2004) study the trade-off between providing efficient labor supply incentives on one hand and generating an equitable after-tax income distribution or providing income insurance on the other hand. We follow the tradition of this literature, but take a quantitative approach as Altig et al. (2001), Ventura (1999), Castañeda et al. (1999), Domeij and Heathcote (2001) and Nishiyama and Smetters (2005) in their positive analyses of tax reforms. On the normative side, 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.[4] Therefore these papers cannot contribute to the discussion about the optimal capital income tax.

Section 2 describes the model and section 3 its calibration. In section 4 we explain the tax experiments, with results presented in section 5. In section 6 we interpret these results and provide extensive sensitivity analysis. Section 7 concludes.

[4]Conesa and Krueger (2006) find an optimal tax code that is roughly a flat tax with sizeable deduction, as proposed by Hall and Rabushka (1995). Saez (2002) studies the optimal size of the deduction in a representative agent model. Smyth (2005) characterizes the optimal capital and labor income tax in a life cycle model that maximizes a weighted sum of lifetime utility of *all agents alive* in the steady state. Since in his model households are identical at birth his analysis also does not capture redistributive motives for taxation.

2 The Economic Environment

2.1 Demographics

Time is discrete and the model is populated by J overlapping generations. In each period a continuum of new households is born whose mass grows at rate n. Each household faces a positive probability of death at each age. Let ψ_j denote the conditional survival probability from age j to age $j+1$. At age J agents die with probability one, $\psi_J = 0$. There are no annuity markets and therefore a fraction of households leaves unintended bequests, denoted by Tr_t, that are redistributed in a lump-sum manner across individuals currently alive. At an exogenous age j_r, agents retire and start to receive social security payments SS_t, which are financed by proportional payroll taxes $\tau_{ss,t}$, paid up to an income threshold $\bar{y}$.

2.2 Endowments and Preferences

Households are endowed with one unit of 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 market or consuming leisure.

Households are heterogeneous along three dimensions that affect their labor productivity. First, they differ by age in their average labor productivity ε_j, which governs the average wage of an age cohort. Retired agents (those with $j \geq j_r$) are not productive at all, $\varepsilon_j = 0$. Second, we introduce permanent differences in productivity, standing in for differences in education and innate abilities. We assume that households are born as one of M possible ability types $i \in \mathbf{I}$, and that this ability does not change over a household's life cycle. 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 the same age and ability face idiosyncratic risk with respect to their individual labor productivity. Let $\eta \in \mathbf{E}$ denote a generic realization of this idiosyncratic labor productivity uncertainty in the current period. 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.

$$Q_t(\eta, E) = Prob(\eta' \in E|\eta) = Q(\eta, E). \tag{1}$$

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 $\bar{\eta} \in \mathbf{E}$ and $\Pi(\eta)$ is the probability of η under the stationary distribution. Different realizations of the stochastic process then give rise to cross-sectional productivity distributions that become more dispersed as a cohort ages. In the absence of explicit insurance markets for labor productivity risk a progressive tax system is an effective policy to share this idiosyncratic risk across agents.

At any given time households are characterized by (a, η, i, j), where a are current holdings of one period, risk-free bonds, η is stochastic labor productivity status, i is ability type and j is age. A household of type (a, η, i, j) working l_j hours commands pre-tax labor income $\varepsilon_j \alpha_i \eta \ell_j w_t$, where w_t is the wage per efficiency unit of labor at time t. Let $\Phi_t(a, \eta, i, j)$ denote the measure of agents of type (a, η, i, j) at date t.

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

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

where β is the time discount factor. Expectations are taken with respect to the stochastic processes governing idiosyncratic labor productivity and mortality.

2.3 Technology

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

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

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 constant A normalizes units in our economy[5], and the depreciation rate for physical capital is denoted by δ.

2.4 Government Policy

The government engages in three activities: it spends resources, 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 $\bar{y}$, as in the actual U.S. system. 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 take as exogenously given in our analysis. Second, the government taxes capital income of households, $r_t(a + Tr_t)$ according to a constant marginal capital tax rate $\tau_{K,t}$.[6] Here r_t denotes the risk free interest rate, a denotes asset held by the household, and Tr_t denotes transfers from accidental bequests. Finally, the government can tax each

[5] We abstract from technological progress, since we will consider 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 with microeconometric evidence, as well as a relative risk aversion that is widely used in the literature.

[6] Section 6.3.2 will explicitly study under what conditions the fiscal authority might find optimal to introduce progressivity in capital income taxes.

individual's taxable labor income according to a potentially progressive labor income tax schedule T. Define as $yp_t = w_t \alpha_i \varepsilon_j \eta l_t$ 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 current U.S. tax law. Thus we define as taxable labor income[7]

$$y_t = \begin{cases} yp_t - ess_t & \text{if } j < j_r \\ 0 & \text{if } j \geq j_r \end{cases} \tag{4}$$

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, in contrast, can be made an arbitrary function of individual taxable labor income in a given period. We denote the tax code by $T(y_t)$. Our investigation of the optimal tax code then involves finding the labor income tax *function* T and the capital tax rate τ_K that maximize a social welfare function, defined below.

[7] Social security benefits are not taxable in our model. Such a tax would constitute a lump-sum tax and here we follow our general approach and take the structure of the social security system as exogenously given to focus on the optimal structure of the distortionary income tax system.

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 absent. However, agents trade one-period risk free bonds to self-insure against labor productivity risk. The possibility of self-insurance is limited, however, by the imposition of 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.

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 $\mathbf{S} = \mathbf{R}_+ \times \mathbf{E} \times \mathbf{I} \times \mathbf{J}$. 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

[8] If households were allowed to borrow, 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. In order to keep the asset market structure tractable we decided to prevent agents from borrowing altogether, in line with much of the incomplete markets literature; see e.g. Aiyagari (1994) or Krusell and Smith (1998).

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 government expenditures $\{G_t\}_{t=1}^{\infty}$ and consumption tax rates $\{\tau_{c,t}\}_{t=1}^{\infty}$ and initial conditions K_1 and Φ_1, a competitive equilibrium is a sequence of functions for the household, $\{v_t, c_t, a'_t, l_t\}_{t=1}^{\infty}$, 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 income 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):*

$$v_t(a, \eta, i, j) = \max_{c, a', \ell}\{u(c, l) + \beta\psi_j \int v_{t+1}(a', \eta', i, j+1)Q(\eta, d\eta')\} \tag{5}$$

subject to

$$(1+\tau_{c,t})c+a' = w_t\varepsilon_j\alpha_i\eta\ell-\tau_{ss,t}\min\{w_t\varepsilon_j\alpha_i\eta l, \bar{y}\}+(1+r_t(1-\tau_{K,t}))(a+Tr_t)-T_t[y_t], \text{for } \ j < j_r, \tag{6}$$

$$(1+\tau_{c,t})c + a' = SS_t + (1 + r_t(1 - \tau_{K,t}))(a + Tr_t), \text{for } \ j \geq j_r, \tag{7}$$

$$a' \geq 0, c \geq 0, 0 \leq l \leq 1. \tag{8}$$

2. Prices w_t and r_t satisfy:

$$r_t = \alpha A \left(\frac{N_t}{K_t}\right)^{1-\alpha} - \delta, \tag{9}$$

$$w_t = (1-\alpha) A \left(\frac{K_t}{N_t}\right)^{\alpha}. \tag{10}$$

3. The social security policies satisfy

$$\tau_{ss,t} \int \min\{w_t \alpha_i \varepsilon_j \eta l_t, \bar{y}\} \Phi_t(da \times d\eta \times di \times dj) = SS_t \int \Phi_t(da \times d\eta \times di \times \{j_r, ..., J\}). \tag{11}$$

4. Transfers are given by:

$$Tr_{t+1} \int \Phi_{t+1}(da \times d\eta \times di \times dj) = \int (1-\psi_j) a'_t(a, \eta, i, j) \Phi_t(da \times d\eta \times di \times dj) \tag{12}$$

5. Government budget balance:

$$\begin{aligned} G_t \;\; = \;\; & \int \tau_{K,t} r_t (a + Tr_t) \Phi_t(da \times d\eta \times di \times dj) + \\ & \int T_t[y_t] \Phi_t(da \times d\eta \times di \times dj) + \\ & \tau_{c,t} \int c_t(a, \eta, i, j) \Phi_t(da \times d\eta \times di \times dj) \end{aligned} \tag{13}$$

6. Market clearing:

$$K_t = \int a \Phi_t(da \times d\eta \times di \times dj) \tag{14}$$

$$N_t = \int \varepsilon_j \alpha_i \eta l_t(a, \eta, i, j) \Phi_t(da \times d\eta \times di \times dj) \tag{15}$$

$$\int c_t(a, \eta, i, j)\Phi_t(da \times d\eta \times di \times dj) + K_{t+1} + G_t = AK_t^{\alpha} N_t^{1-\alpha} + (1-\delta)K_t \tag{16}$$

7. *Law of Motion:*

$$\Phi_{t+1} = H_t(\Phi_t) \tag{17}$$

where the function $H_t : \mathbf{M} \to \mathbf{M}$ *can be written explicitly as follows. For all* $\mathcal{J}$ *such that* $1 \notin \mathcal{J}$*:*

$$\Phi_{t+1}(A \times E \times \mathcal{I} \times \mathcal{J}) = \int P_t((a, \eta, i, j); A \times E \times \mathcal{I} \times \mathcal{J})\Phi_t(da \times d\eta \times di \times dj) \tag{18}$$

where

$$P_t((a, \eta, i, j); A \times E \times \mathcal{I} \times \mathcal{J}) = \begin{cases} Q(e, E)\psi_j & \text{if } a'_t(a, \eta, i, j) \in A, i \in \mathcal{I}, j+1 \in \mathcal{J} \\ 0 & \text{else} \end{cases} \tag{19}$$

For $\mathcal{J} = \{1\}$

$$\Phi_{t+1}((A \times E \times \mathcal{I} \times \{1\}) = (1+n)^t \begin{cases} \sum_{i \in \mathcal{I}} p_i & \text{if } 0 \in A, \bar{\eta} \in E \\ 0 & \text{else} \end{cases} \tag{20}$$

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 ag-*

gregate variables grow at the constant growth rate of the population n.

3 Functional Forms and Calibration

In order to carry out the numerical determination of the optimal tax code we first choose a model parameterization which we now describe.

3.1 Demographics

In our model households are born at age twenty (model age 1). They 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 I summarizes our choices for all parameters (also the ones to come).

3.2 Preferences

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

$$u(c, 1-l) = \frac{(c^{\gamma}(1-l)^{1-\gamma})^{1-\sigma}}{1-\sigma} \tag{21}$$

where γ is a share parameter determining the relative importance of consumption, and σ determines the risk aversion of the household.[9] 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.[10] The calibrated values of σ and γ imply that the intertemporal elasticity of substitution is approximately 0.5.

Microeconometric studies tend to restrict attention to white males of prime age already employed and obtain values for the Frisch elasticity smaller than one. We take as decision unit in our model the household. It therefore seems reasonable that the labor supply elasticity might be higher than the low estimates implied by traditional microeconometric studies, due to both higher labor supply elasticities of females and the existence of an extensive margin that is not usually considered in the empirical estimation of labor supply elasticities.[11]

---

[9] The coefficient of relative risk aversion in consumption is given by $-\frac{cu_{cc}}{u_c} = \sigma\gamma + 1 - \gamma$.

[10] It is understood that in any general equilibrium model all parameters affect all equilibrium entities. We associate a parameter with the equilibrium entity it affects quantitatively most.

[11] Heckman (1993) suggests that the elasticity of participation decisions is large. Most of the movement in aggregate hours worked is due to this extensive margin. 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, possibly as high as 3.82. Domeij and Floden (2006) show 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 that are homothetic in *hours worked* where the substitution and income effects exactly cancel each other and obtain a Frisch labor supply elasticity around 1, equal to the one implied in our benchmark economy.

Table I: Calibrated Parameters

Parameter	Value	Target
Demographics		
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
Preferences		
Discount Factor β	1.001	$K/Y = 2.7$
Risk Aversion σ	4.0	IES = 0.5
Consumption Share γ	0.377	Avg. Hours= $\frac{1}{3}$
Labor Productivity Process		
Variance Types σ_α^2	0.14	$Var(y_{22})$
Persistence ρ	0.98	Lin. Incr. in $Var(y_j)$
Variance Shock σ_η^2	0.0289	$Var(y_{60})$
Technology		
Capital Share α	0.36	Data
Depreciation Rate δ	8.33%	$I/Y = 25.5\%$
Scale Parameter A	1	Normalization
Government Policy		
Consumption Tax τ_c	5%	Mendoza et al. (1994)
Marginal Tax κ_0	0.258	Gouveia and Strauss (1994)
Tax Progressivity κ_1	0.768	Gouveia and Strauss (1994)
Payroll Tax τ_{ss}	12.4%	Data

Given the difficulty to empirically pin down the labor supply elasticity for our model, we conduct sensitivity analysis with respect to the labor supply elasticity by considering an alternative preference specification that allows us to choose a lower elasticity than in our benchmark. This alternative is given by

$$u(c, 1-l) = \frac{c^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-l)^{1-\sigma_2}}{1-\sigma_2} \tag{22}$$

We discuss the calibration of the parameters $(\sigma_1, \sigma_2, \chi)$ in section 6.3.1.

3.3 Labor Productivity Process

A household's labor productivity depends on three components: a deterministic age-dependent component ε_j, a type-dependent fixed effect α_i and a persistent, idiosyncratic shock η. The natural logarithm of wages of a household is given by

$$\log(w_t) + \log(\varepsilon_j) + \log(\alpha_i) + \log(\eta) \tag{23}$$

The age-productivity profile $\{\varepsilon_j\}_{j=1}^{j_r-1}$ 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 $\alpha_2 = e^{\sigma_\alpha}$, so that $E(\log(\alpha_i)) = 0$ and $Var(\log(\alpha_i)) = \sigma_\alpha^2$. 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 the three free parameters $(\sigma_\alpha^2, \rho, \sigma_\eta^2)$

to choose. With their choice we target three statistics from data measuring how cross-sectional labor income dispersion evolves over the life cycle. Storesletten et al. (2004) document that i) at cohort age 22 the cross-sectional variance of *household* 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 household age-earnings variance profile consistent with these facts. The implied parameter values are summarized in Table I.

3.4 Technology

The capital share parameter α is set to the empirical capital share, $\alpha = 0.36$, a standard value in the literature.[12] The depreciation rate is set to match an investment-output ratio of 25.5% in the data (investment includes nonresidential and residential fixed investment and purchases of consumer durables). This requires $\delta = 8.3\%$.[13]

3.5 Government Policies and the Income Tax Function

The government in our model (meant to stand in for all levels, federal, state and local in the real world) consumes resources, collects tax revenues and operates a

[12] For example, Castañeda et al. (1999) choose $\alpha = 0.376$ and Domeij and Heathcote use $\alpha = 0.36$.

[13] Note that our parameter choices yield a benchmark real interest rate of 5%, and with population growth of 1.1% the economy is deep in the dynamically efficient region.

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 choose government spending G so that it accounts for 17% of GDP in the initial stationary equilibrium. G is kept constant across our tax experiments; therefore if an income tax system delivers higher output in equilibrium, the corresponding $\frac{G}{Y}$ ratio declines.

Part of tax revenues are generated by a proportional consumption tax, whose size we take as exogenous and set to $\tau_c = 5\%$, following Mendoza et al. (1994). Furthermore the government runs a pay-as-you-go social security system, defined by a payroll tax. This tax takes a value of 12.4% of labor income up to a limit of 2.5 times average income,[14] with benefits determined by budget balance of the system.

We want to determine the optimal income tax function. Ideally one would impose no restrictions on the set of 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, total taxes are given by

$$T^{GS}(y; \kappa_0, \kappa_1, \kappa_2) = \kappa_0 \left(y - (y^{-\kappa_1} + \kappa_2)^{-\frac{1}{\kappa_1}} \right), \tag{24}$$

where $(\kappa_0, \kappa_1, \kappa_2)$ are parameters. This functional form, proposed by Gouveia and Strauss (1994), has been employed in the quantitative public finance literature by

[14] The limit of earnings subject to the payroll tax changes every year, which is $102,000 in 2008.

Castañeda et al. (1999), Smyth (2005) and Conesa and Krueger (2006). Roughly speaking, κ_0 controls the level of the average tax rate and κ_1 determines the progressivity of the tax code. For $\kappa_1 \to 0$ the tax system reduces to a pure flat tax, while other values encompass a wide range of progressive and regressive tax functions.

Without discriminating between capital and labor income Gouveia and Strauss (1994) estimate the parameters $(\kappa_0, \kappa_1, \kappa_2)$ that best approximate taxes paid under the actual US income tax system and find $\kappa_0 = 0.258$ and $\kappa_1 = 0.768$. We use their estimated tax system (applied to the sum of labor and capital income) as benchmark, for calibration and comparison purposes. The parameter κ_2 adjusts to ensure government budget balance.

4 The Computational Experiment

Define y_l and y_k as taxable labor and capital income, respectively. The set of tax functions the government optimizes over is given by

$$\mathcal{T} = \left\{T_l(y_l), T_k(y_k) : T_l(y_l) = T^{GS}(y_l; \kappa_0, \kappa_1, \kappa_2) \text{ and } T_k(y_k) = \tau_k y_k\right\} \qquad (25)$$

and thus by the four parameters $(\kappa_0, \kappa_1, \kappa_2, \tau_k)$, one of which (we take κ_2) is determined by budget balance. Thus we allow for a flexible labor income tax code, but restrict capital taxes to be proportional, an assumption that assures computational feasibility and makes the comparison to existing studies employing the same assumption easier. Also note that the choices of $(\kappa_0, \kappa_1, \tau_k)$ are restricted by the requirement that there has to exist a corresponding κ_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 born into the stationary equilibrium implied by the chosen tax function. The government's objective is thus given by

$$SWF(\kappa_0, \kappa_1, \tau_k) = \int v_{(\kappa_0,\kappa_1,\tau_k)}(a = 0, \eta = \bar{\eta}, i, j = 1) d\Phi_{(\kappa_0,\kappa_1,\tau_k)} \tag{26}$$

Given that *all* newborn households start with zero assets and average labor productivity, social welfare is simply equal to average expected lifetime utility across the two ability groups.[15]

5 Results

5.1 The Optimal Tax System

The optimal tax system is given by a tax rate on capital of $\tau_k = 36\%$ and a labor income tax characterized by the parameters $\kappa_0 = 23\%$ and $\kappa_1 \approx 7$. Therefore the labor income tax code is basically a flat tax with marginal rate of 23% and a deduction

[15] Here $v_{(\kappa_0,\kappa_1,\tau_k)}$ and $\Phi_{(\kappa_0,\kappa_1,\tau_k)}$ are the value function and invariant cross-sectional distribution associated with tax system characterized by $(\kappa_0, \kappa_1, \tau_k)$.

of about $\$7,200$ (relative to average income of $\$42,000$).[16]

5.2 Comparison with the Benchmark

In order to assess the importance of the tax code for equilibrium allocations Table II compares equilibrium statistics for the optimal and the benchmark tax system.

Table II: Changes in Aggregate Variables in the Optimal Tax System

Variable	
Average Hours Worked	-0.56%
Total Labor Supply N	-0.11%
Capital Stock K	-6.64%
Output Y	-2.51%
Aggregate Consumption C	-1.63%
ECV	1.33%

We observe that under 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

[16] In section 6.3.2 we show that given our social welfare function it is not welfare enhancing to introduce progressivity of the *capital* income tax schedule. According to our results all progressivity of the tax code should be embedded in the *labor* income tax schedule.

labor supply. While average hours worked drop by 0.56%, labor efficiency units drop by only 0.11%; thus labor supply shifts from less to more productive households.

5.2.1 Decomposition of the Welfare Effects

Given the substantial decline in aggregate consumption and the only modest decline in average hours worked in the optimal tax system, relative to the benchmark, it is surprising that the optimal tax system features substantially higher aggregate welfare, equivalent to a 1.33% increase in consumption at all ages and all states of the world, keeping labor supply allocations unchanged.[17] Given the form of the utility function, the welfare consequences of switching from a steady state consumption-labor allocation (c_0, l_0) to (c_*, l_*) are given by

$$CEV = \left[\frac{W(c_*, l_*)}{W(c_0, l_0)}\right]^{\frac{1}{\gamma(1-\sigma)}} - 1 \tag{27}$$

where $W(c, l)$ is the expected lifetime utility at birth of a household, given a tax system. We can decompose CEV into a component stemming from the change in consumption from c_0 to c_*, and one from the change in leisure. The consumption impact on welfare can itself be divided into a part that captures the change in average consumption, and a 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

[17] Note that if one imposes zero capital income taxes and let the government optimize over labor income taxes, the welfare *losses* for not using capital income taxes are substantial: 2% relative to the unconstrained optimum and 0.7% relative to the benchmark system.

for labor supply (leisure).[18]

Table III shows that 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 than offsets the lower average level of consumption and the less favorable, in utility terms, distribution of leisure over the life cycle.

Table III: Decomposition of Welfare

Total Change		1.33%
Consumption	Total	1.29%
	Level	−1.63%
	Distribution	2.97%
Leisure	Total	0.04%
	Level	0.41%
	Distribution	−0.37%

[18] 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 $1+CEV = (1+CEV_C)(1+CEV_L)$ or $CEV \approx CEV_C + CEV_L$. We further decompose CEV_C into a level effect CEV_{CL} and a 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 calculation shows that the level effect equals the growth rate of consumption $CEV_{CL} = \frac{C_*}{C_0} - 1$. A similar decomposition applies to leisure.

5.2.2 Life Cycle Profiles of Assets, Labor, Consumption and Taxes

As we argue below a nontrivial life cycle profile of hours worked and consumption is crucial for our capital income tax result. In this subsection we therefore document the life cycle pattern of asset holdings (the relevant tax base for the capital income tax), labor income (the relevant tax base for labor income taxes), consumption and taxes paid. In the upper-left panel of figure 1 we display average asset holdings by age for both productivity types of households, for the benchmark and the optimal tax system. We observe the hump-shaped behavior of assets that is typical of any life cycle model. This profile implies that the main burden of the capital income tax is borne by households aged 40 to 70. In addition, the negative impact on asset accumulation of higher capital income taxes in the optimal, relative to the benchmark tax system, is clearly visible. Therefore aggregate assets and capital decline by 6.6%.

The upper right panel of Figure 1 documents the life cycle pattern of labor supply. First we note that, independent of the tax system, labor supply tends to fall over the life cycle despite the fact that labor productivity only peaks at age 50. This fact, crucial for explaining the optimal capital income tax result in section 6, is mainly driven by the fact that with a calibrated $\beta > 1$ and a substantially positive (after tax) return it is beneficial for households to postpone leisure to older ages. Second, we observe that the optimal tax code, relative to the benchmark, induces households to work more at ages at which they are more productive. Lower labor income taxes and the sizeable deduction induce 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 elasticity of hours with respect to marginal labor income taxes for this group. Overall the optimal tax system induces a flatter life cycle profile of labor supply and thus leisure.

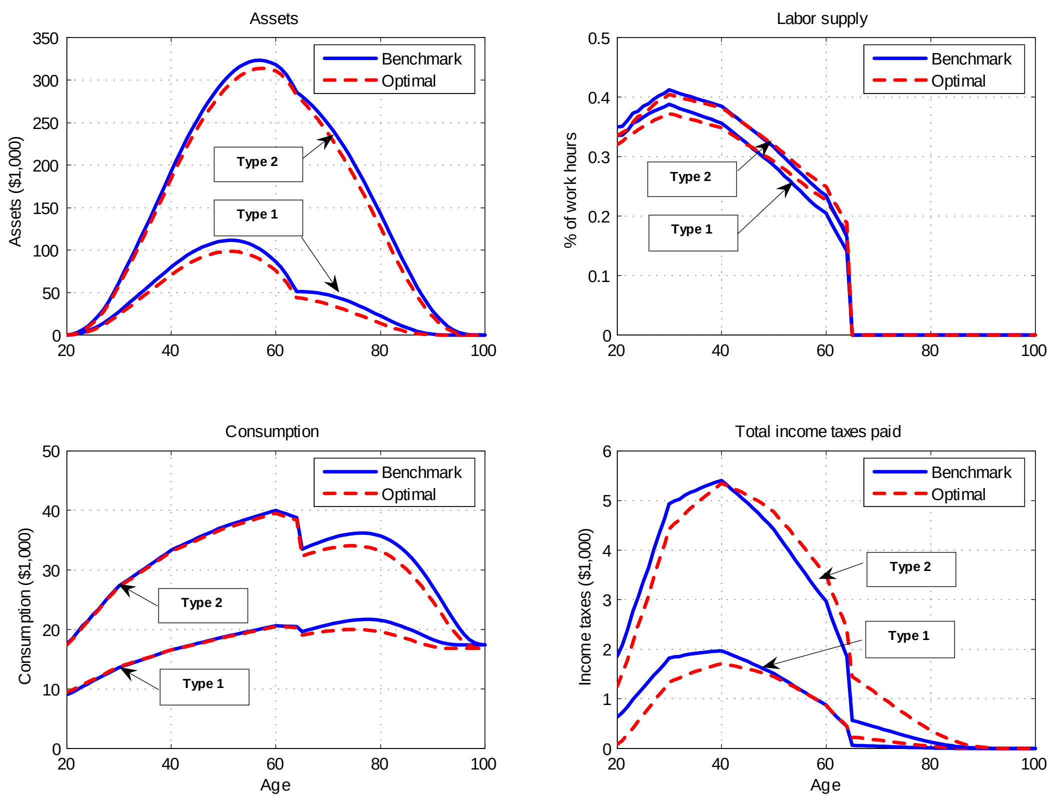

Figure 1: Life Cycle Profiles of Assets, Labor Supply, Consumption and Taxes

The lower left panel of Figure 1 documents the life cycle consumption pattern; it displays an empirically plausible hump over the life cycle and discrete fall at the time of retirement, due to the nonseparability between consumption and leisure. Relative to the benchmark tax system a larger capital income tax makes future consumption more expensive and thus flattens its profile. Finally, the lower right panel of Figure 1 displays the life cycle profile of taxes paid. The figure first demonstrates that the optimal tax code leads to substantially more redistribution across types, by taxing more heavily the high-skilled, high labor income-earners who also hold a large fraction of financial assets in the economy, especially at ages 40 to 60. Second, since under the optimal tax code households aged 40 to 60 work more than under the benchmark, they pay higher labor income taxes (despite the fact that their marginal tax rates have been reduced). Finally, the higher capital income taxes of the optimal system explain why retired capital holders pay a larger tax bill under this system.

6 Interpretation and Sensitivity of the Results[19]

To isolate the driving forces for our two quantitative results, a significantly positive capital income tax and a labor income tax schedule that is progressive due to a substantial deduction we now show which model elements are responsible for these findings. The crucial model elements include i) an *endogenous labor-leisure choice*, formally represented by a $\gamma < 1$ in the utility function, ii) a *tight borrowing constraint*

[19] **The proofs of various claims in this section are contained in a separate appendix available at http://www.econ.upenn.edu/~dkrueger/research/taxapp.pdf**

$a' \geq 0$, iii) *ex-ante heterogeneity* in labor productivity, $\sigma_\alpha^2 > 0$, iv) *idiosyncratic income risk*, driven by the Markov chain for labor productivity η, that is $\sigma_\eta^2 > 0$, and v) model elements that let households undergo a meaningful *life cycle*, such as idiosyncratic mortality risk, $\psi_j < 1$, a nontrivial life cycle profile of wages $\varepsilon_j \neq 1$, and a PAYGO social security system, $\tau_{ss} > 0$ and $SS > 0$.

Table IV: Summary of Quantitative Results[20]

Model	End.[21] Lab.	BC	Type	Idio.	Life Cyc.	β	r	τ_k	τ_l	Prog.
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	Yes	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
Bench	Yes	Yes	Yes	Yes	Yes	1.001	5.6%	36%	23%	Yes

[20] In the table, τ_l gives the marginal tax rate on labor as labor earnings tend to infinity. In a flat tax system (with or without deduction) it is the common marginal tax rate. The column *Prog.* simply indicates whether the optimal labor tax system is significantly progressive.

[21] *End. Lab.* stands for model element i) labor being chosen endogenously, *BC* indicates ii) the presence of a tight borrowing constraint, *Type* implies iii) the presence of ex-ante productivity heterogeneity, *Idio.* iv) the presence of idiosyncratic income risk, and *Life Cyc.* indicates v) the presence of the life cycle model elements listed in the main text.

Table IV summarizes the optimal tax code in various versions of our model where various combinations of these elements are shut down. The benchmark model analyzed so far, the last row of the table, includes all five model elements. In all models parameters have *always* been re-calibrated to match (with the Gouveia-Strauss tax function) the same targets as in the benchmark. We observe the following. First, without labor being supplied elastically no *robust* argument can be made for significantly positive capital income taxes. Second, the size of the optimal capital income tax when households choose how much to work depends crucially on the presence of realistic life cycle elements of the model. Third, while type heterogeneity and idiosyncratic earnings risk are key determinants of the progressivity of *labor income taxes*, the finding that a large capital income tax is optimal does not hinge on these model elements. Finally, we find that borrowing constraints are not crucial for our optimal capital income tax results *if* labor income taxes are permitted to be progressive. We now discuss these findings in more detail.

6.1 Inelastic Labor Supply

To argue that endogenous labor supply being chosen over a realistically modeled life cycle is crucial for our optimal capital income tax results we briefly present optimal tax results for versions of the model with exogenous labor supply. In these model variants the labor income tax is a lump-sum tax. Thus it is not surprising that we find that *no robust case for positive capital income taxes can be made with exogenous* *labor supply.*

In the most basic OLG model in which households live for J periods, only value consumption in their time separable lifetime utility function and face no risk or borrowing constraints, the population grows at rate n, capital depreciates at rate δ and (per capita of the youngest generation) output is produced according to a neoclassical production function $f(K)$ with standard properties, and used for private consumption, investment and government spending G, a social planner whose objective is to maximize steady state lifetime utility of a newborn agent will choose the capital stock to satisfy the golden rule $f'(K^*) = \delta + n$. Furthermore the socially optimal allocation can be implemented as a competitive equilibrium in which the government chooses $\tau_k = 0$, as long as there are no restrictions on government debt. However, if government debt is restricted to zero (that is, period by period budget balance is imposed) there is no guarantee that with $\tau_k = 0$ private asset demand equals the capital stock K^*. Therefore, in the absence of (negative) government debt, even in this simple model the optimal capital income tax is not necessarily equal to zero. When we recalibrate parameters to the same empirical targets as in the benchmark[22] we find that in the model without government debt the optimal capital income tax is 10% and the associated labor income tax equals $\tau_l = 19\%$ as shown in Table IV (row M1). We return to the restrictions the absence of government debt imposes on the optimal tax code in section 6.3.5.

[22] The parameters are σ=2, β=0.9825, α=0.36, δ =0.0833 and n=0.011.

However, the *absence of government debt alone does not generate a robust reason for positive capital income taxes in models with exogenous labor supply.* Adding all life cycle elements to the model yields an optimal capital income tax of approximately $\tau_k = 0$ if the labor income tax is restricted to be proportional.[23] If the government can tax labor progressively, it does so drastically, as the 100% marginal tax rate in row M2 in Table IV shows. The optimal capital income tax becomes substantially negative, -24%. For a given consumption profile, shifting after tax labor income towards younger ages through highly progressive taxes increases aggregate saving and thus the capital stock and the *level* of aggregate consumption.

The model with exogenous labor supply can also be used to most clearly show how effectively progressive labor income taxes can deal with the problem of potentially binding borrowing constraints. Hubbard and Judd (1986) and Imrohoroglu (1998) show that in life cycle models where households face upward-sloping labor earnings profiles and tight borrowing constraints the government should not rely on labor income taxes alone, since high labor income taxes translate directly into low consumption for young households at the constraint.[24] However, if the labor income tax code is allowed to be progressive it is possible to tax young, borrowing constrained households with lower labor earnings at lower rates than older households. Adding tight borrowing constraints and idiosyncratic risk (as well as type heterogeneity) to

[23] If β is kept the same as before, then optimal capital income tax with proportional labor income tax roughly remains at its previous optimum of 10%.

[24] This effect can be especially severe, as Imrohoroglu (1998) argues, if households in addition face idiosyncratic income shocks.

the model confirms our previous findings. 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 higher than in the model without borrowing constraints, as suggested by the results of Imrohoroglu (1998). However, as soon as we allow the labor income tax to be progressive, Table IV (row M3) shows that the optimal capital income tax is -34%, financed by a high marginal labor income tax with substantial deduction. The finding that *tight borrowing constraints do not provide a strong rationale for capital income taxes if labor income taxes are allowed to be progressive* continues to hold with endogenous labor supply, as our results below indicate.

6.2 Elastic Labor Supply

The main lesson we draw from the previous subsection is that labor supply endogenously chosen over the life cycle is a necessary ingredient of our optimal tax result. In this subsection we first review what can be said theoretically about the optimal tax structure in life cycle models with endogenous labor supply and then decompose our quantitative results further.

6.2.1 Theory

Atkeson et al. (1999), Erosa and Gervais (2002) and Garriga (2003) analyze theoretically the optimal tax structure in OLG models without idiosyncratic risk, type heterogeneity and the restriction to proportional, albeit potentially age-dependent tax schedules. To provide the cleanest intuition for our quantitative results, suppose that households live only two periods, value consumption and leisure in both periods, and have labor productivity of 1 when young and ε when old. The production technology is given by $F(K, L) = rK + L$ so that the marginal product of capital is constant at r and the marginal product of labor is constant at 1. The benevolent government maximizes social welfare[25] by choosing (potentially age-dependent) proportional labor and capital income taxes and uses government debt to satisfy its sequence of budget constraints.

With these assumptions several analytical results can be derived.[26] First, suppose that preferences are separable between consumption and leisure and obey the functional form specified in (22), the case we analyze quantitatively in section 6.3.1. Then, *if* labor income taxes can be conditioned on age, in the steady state the optimal capital income tax equals zero and the optimal age-dependent labor income tax rates satisfy $\tau_{l1} > \tau_{l2}$ if and only if $l_1 > l_2$ (if labor supply is falling over the life cycle). That is, labor is taxed at a higher rate when it is high. Since with these preferences the Frisch labor supply elasticity is given by $\varepsilon_l = \frac{1-l}{\sigma_2 l}$, this can be restated as labor being taxed more heavily when it is supplied less elastically.

[25] Social welfare is a weighted sum of lifetime utilities of current and future generations.

[26] Similar results for more general life cycle environments are contained in Atkeson et al. (1999), proposition 7, Erosa and Gervais, section 4.2 and Garriga (2003), proposition 4. The purpose of our discussion is not to claim originality, but to provide the clearest possible intuition for our quantitative results.

If labor income taxes cannot be conditioned on age, the optimal long-run capital income tax is zero only if labor supply does not undergo a life cycle, that is, if $l_1 = l_2$. Denoting by U_{l_i} the marginal disutility from work the intertemporal optimality condition governing labor supply reads as

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

The same intertemporal wedge can be generated with a combination of ($\tau_{l1} > \tau_{l2}$ and $\tau_k = 0$) and with ($\tau_{l1} = \tau_{l2}$ and $\tau_k > 0$). Thus *a positive capital income tax mimics a labor income tax that is falling with age.* This role of capital income taxes has to be traded off against the distortion for the intertemporal consumption allocation.

If the period utility function is of the Cobb-Douglas form (21), 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. In particular, if labor supply falls over the life cycle (as it does in our quantitative model), then the optimal capital income tax is positive (as long as $\sigma > 1$ which we assume in the quantitative model). 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. Note that the previous argument 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 continues to apply.

To summarize, endogenous labor supply coupled with life cycle model elements that generate a non-constant age labor supply profile implies a robust role for positive capital income taxation, as long as the government cannot condition the tax code on age (and in the nonseparable case, even with age-dependent labor income taxes). *The capital income tax implicitly 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 makes use of the capital income tax for this reason.

6.2.2 Quantitative Findings

The theoretical discussion has argued that in a life cycle model in which household labor supply changes with age, 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 literature on optimal taxation in the Ramsey tradition restricts attention to proportional taxes. 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 taxing households of different ages at different rates. But if labor supply is endogenous progressive taxes have adverse incentive effects. Quantitative analysis is required to determine the extent to which capital income tax and the progressive labor income tax are used to tax labor at different ages at different rates.[27]

In the simplest quantitative model with endogenous labor supply that abstracts from idiosyncratic risk, household heterogeneity and life cycle elements (apart from

[27] Gervais (2004) studies the desirability of a progressive income tax system implied by a log-linear tax function using a life-cycle model similar to ours.

finite life), the capital income tax is quantitatively significant at 20% (see row M4 in Table IV). Its size is, however, only about half of that in the benchmark model, for reasons explained below. The calibrated parameters imply that for all τ_k in equilibrium $\beta(1 + r(1 - \tau_k)) > 1$. Therefore consumption and leisure increase over the life cycle. As our theoretical discussion above suggests, labor at older ages should be taxed less heavily than labor at younger ages. In the absence of age-dependent labor income taxes a positive capital income tax achieves this.[28]

Adding life cycle elements into the model (social security, mortality risk, and age-dependent labor productivity) strongly affects the optimal life cycle profile of consumption and leisure, and incentives to save.[29] In order to achieve the same empirical target the newly calibrated time discount factor increases to $\beta = 1.009$, see row M5 in table IV. This in turn generates consumption and leisure profiles that are more strongly upward sloping, relative to model M4. As a consequence the optimal capital income tax, inducing indirect age-dependent leisure taxation, rises to 34%, with a labor income tax that is proportional with a marginal rate of 14%. Thus *a fully articulated life cycle model with endogenous labor supply but no intracohort heterogeneity implies an optimal capital income tax almost as high as in the benchmark model, but an essentially proportional optimal labor income tax.*

[28] With a flat labor efficiency profile and declining labor supply over the life cycle earnings decline with age. Thus a progressive labor income tax allows the government to tax older households that supply less labor at lower rates. The optimal tax code is therefore very slightly progressive.

[29] Mortality risk implies that the time discount factor is adjusted by the conditional probability of survival. This adjustment is quantitatively important for elderly households and helps to generate the falling portion of the hump-shaped life cycle consumption profile.

In model M6 we add ex-ante heterogeneity among households, in the form of permanent labor productivity differences. This model element strongly affects the optimal labor income tax code, but leaves the capital income tax virtually unaffected. Given our social welfare function the government has a motive to redistribute between households of different ability types. A progressive labor income tax that taxes high-earnings households at higher rates is exactly the appropriate tool for this type of redistribution. Thus the labor income tax schedule becomes substantially progressive. One way to see this from Table IV is to notice that κ_0, the marginal tax rate for highest earnings households, increases from 14% to 18%, mainly to compensate for the lower tax revenues collected from low-earnings households. *Further introducing intracohort heterogeneity in the form of uninsurable idiosyncratic labor productivity risk strengthens the case for progressive labor income taxes* to provide partial insurance against idiosyncratic labor income risk. Row M7 of Table IV shows that, as a result marginal tax rates at low income levels fall, and they increase at higher earnings levels (the highest marginal rate is now $\kappa_0 = 23\%$) and for capital income (the optimal capital income tax now is 35%).[30]

The introduction of tight borrowing constraints leads us back to our benchmark

[30] Aiyagari (1995) argues that uninsurable idiosyncratic risk 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 his model the government also chooses optimally the level of government consumption which enters the households' utility function additively. When we repeat our public finance analysis in Aiyagari's (1995) model we find consistently high capital income *subsidies* as optimal. The optimal capital subsidy is about 45% in the Aiyagari model, calibrated to the same targets as in our benchmark. These results are consistent with the theoretical analysis in Davila et al. (2007) and demonstrate that Aiyagari's (1995) results depend crucially on government consumption G being endogenous in his model.

model, with results summarized in the last row of Table IV. As can be seen the effect of borrowing constraints on the optimal tax code is relatively minor. The reason for this finding is two-fold. First, as explained above the use of progressive labor income taxes allows to tax the young and poor at lower rates than older households. Second, in crucial difference to Hubbard and Judd (1986) the presence of idiosyncratic risk induces precautionary saving even early in the life cycle. As the asset life cycle profile in Figure 1 shows most young households find it optimal to hold a small but positive amount of assets. Consequently the borrowing constraint is not binding for most households, and relaxing it has quantitatively noticeable, but no dramatic effects on our findings (compare rows M7 and Bench in Table IV).

6.3 Sensitivity Analysis

In this section we document how sensitive our results are with respect to the assumed labor supply elasticity of households. Then we investigate whether it is optimal to employ a progressive capital income tax schedule. We then study the dependence of our results on the intertemporal elasticity of substitution, and on the form of the social welfare function. Finally we discuss what happens if we allow the government to accumulate (negative) government debt.

6.3.1 Elasticity of Labor Supply

Our previous results for the optimal mix of capital and labor income taxes in our model were based on the idea that older, high earnings households have a higher labor supply elasticity than younger households. In this section we want to investigate whether our findings are robust to a different preference specification that implies a lower (overall) labor supply elasticity. The functional form of the utility function is given in (22). We choose as parameters a coefficient of relative risk aversion of $\sigma_1 = 2$, and $\sigma_2 = 3$. The latter choice implies a Frisch labor supply elasticity that is now substantially below one.[31] For the remaining preference parameters (β, χ) as well as the other model parameters we follow the same calibration strategy as above; this yields $\beta = 0.972$ and $\chi = 1.92$ as new parameters.[32]

Under this new parameterization we find as optimal tax code a marginal capital income tax of $\tau_k = 21\%$ and a marginal labor income tax rate of $\kappa_0 = 34\%$ and $\kappa_1 = 18$, implying again a flat tax rate on labor with deduction of now $\$10,800$. So whereas the main qualitative findings of a significantly positive capital income tax and a flat labor income tax with sizeable deduction remain intact, quantitatively a reduction in the labor supply elasticity shifts the optimal tax mix towards lower capital taxation and higher labor taxation.

Table V summarizes the changes in the aggregate variables under the optimal tax

[31] With this preference specification the the Frisch (constant marginal utility of wealth) labor supply elasticity is equal to $\frac{1-l}{\sigma_2 l} = \frac{2}{3}$, while it was 1 in our benchmark.

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

system, relative to the benchmark. Qualitatively, the results are similar to the ones in the previous section. Quantitatively, however, the decline in the capital stock, output, consumption, and in particular labor supply is more substantial than with nonseparable preferences. Despite a much more severe drop in aggregate consumption the welfare gains are higher now than with Cobb-Douglas preferences, due to a stronger reduction in hours worked and a more equitable consumption distribution.

Table V: Changes in Aggregate Variables in the Optimal Tax System

Variable	
Average Hours Worked	-2.70%
Total Labor Supply N	-2.14%
Capital Stock K	-7.44%
Output Y	-4.08%
Aggregate Consumption C	-3.75%
ECV	3.4%

6.3.2 Progressivity of Capital Income Tax

Our previous analysis restricted the capital income tax to be proportional. We now document that this is not a binding restriction by allowing it to be progressive as well. We initially specified Gouveia-Strauss functional forms for both labor and capital income taxation, but since we found that the optimal tax functions are always very well approximated by a flat tax with deduction, we now document whether the government should choose a positive deduction for capital income or not.

Our quantitative (but not qualitative) results depend somewhat on the chosen social welfare function, a point we will return to in section 6.3.4. In addition to our benchmark ex-ante expected utility of a newborn agent we consider a Rawlsian social welfare function that maximizes the minimum of lifetime utility across the two different types (i.e. maximizes lifetime utility of the low ability type). We also consider a social welfare function that maximizes lifetime utility of the high ability type.

Under our benchmark welfare criterion of ex-ante utility of a newborn, allowing for progressivity of the capital income tax does not improve welfare and the optimal is given by the same proportional tax on capital income that we obtained above.[33] The optimal tax system under the Rawlsian welfare function is a combination of a 30% marginal tax on labor income with a deduction of \$11,500 and a 28% capital income tax, while it is a 14% labor tax with a small deduction of \$500 and a 37% capital tax when only the welfare of the high type is valued. For both alternative social welfare criteria it is not optimal for the government to make capital income taxes progressive. The degree of progressivity in the labor income tax, in contrast, depends crucially on which group of the population is given more weight by the government.[34]

[33] As typical in this class of models our model understates wealth concentration at the top of the distribution, relative to the data. This may weaken the case for progressive capital income taxes.

[34] We have examined various other social welfare functions. The one case where we found a capital income deduction (of roughly $\$5,000$) to be optimal is when the government maximizes the sum of steady state lifetime utilities of all generations. Note that this criterion (in addition to comparing welfare across households of different ages) essentially double-counts old ages, once directly because

6.3.3 Intertemporal Elasticity of Substitution

In our benchmark economy the coefficient of relative risk aversion in consumption (its inverse is the intertemporal elasticity of substitution, IES) is given by $-\frac{cu_{cc}}{u_c} = \sigma\gamma + 1 - \gamma$. Our benchmark choice, $\sigma = 4$, implies an IES of about 0.5. Now we perform sensitivity analysis with respect to our choice of σ, in order to assess how the elasticity of consumption growth (and thus savings) with respect to changes in the after tax real interest rate affects our results.

We restrict the tax code to be flat with deduction on labor and flat on capital, to make the results comparable to those in the previous subsection.[35] If we lower the value of σ to 2 (that is, increase the IES to about 3/4), the optimal capital income tax is characterized by a marginal labor income tax of 18.9%, a deduction on labor of around \$5,400 and a marginal capital income tax of 22.9%, substantially smaller than in the benchmark.[36] Increasing σ to 8 yields a substantially higher capital income tax of 52.4%, in conjunction with a marginal labor income tax of 19.1% and deduction of around \$10,300. The results are consistent with the findings in Imrohoroglu (1998), and confirm the intuition that increasing the IES (lowering σ) shifts the optimal tax system towards labor and away from capital income.

---

of old households and once because of the years currently young households will spend in old age. Since old households derive most of their income from capital, not surprisingly the optimal capital income tax deduction is positive.

[35] This also applies to subsection 6.3.4.

[36] In this section the parameters are not recalibrated. If we recalibrate, the changes in results are qualitatively similar, but differences to our benchmark results become somewhat smaller.

6.3.4 Social Welfare Function

In this section we document that our quantitative results are robust to different specifications of the social welfare function. We have already shown in 6.3.2 that a Rawlsian maximin rule or a social welfare function that gives all the weight to the high productivity types also produce substantially positive capital income tax rates.

We also compute the tax code that maximizes ex-ante expected utility, subject to the constraint that none of the two groups is worse off than in the status quo (i.e. under the Gouveia-Strauss tax function). Compared to the unconstrained optimum, lifetime expected utility of the high type has to be raised as the constraint is binding for this group.[37] The optimal way to do so is for the government to lower the marginal labor income tax to 19%, with a deduction of $9,700. The optimal capital income tax increases slightly to 44%. Although high type households are subject to a higher capital income tax under this tax system, their after tax income increases due to lower labor income taxes.

A Side Remark on Welfare along the Transition We characterize the optimal tax system in the steady state. In many optimal policy analyses focusing on steady states is problematic since it ignores the transitional welfare costs or gains associated with the economy converging to the new steady state. Here we would like to point out that the substantially positive capital income tax that we found optimal leads to

[37] Without these constraints welfare gains for the low types are in the order of 2.5% and welfare losses for the high types amount to about 0.5% in terms of CEV.

a capital stock that is significantly below the status quo. Thus, along the transition path the capital stock falls and can be partly consumed by transitional generations.

While it is computationally infeasible to compute the entire optimal tax transition we conducted some experiments to mimic such an exercise. Restricting the tax code to a proportional capital income tax and a proportional labor income tax plus deduction, we ask what is the optimal once and for all tax reform. That is, starting at the status quo the constant marginal labor and capital income tax rate (τ_l, τ_k) are chosen, and the deduction in every period along the transition adjusts to guarantee budget balance. In order to define optimality we again have to define a social welfare function that now incorporates 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* in the initial steady state. We find an optimal capital income tax of $\tau_k = 65\%$, with the labor income tax at $\tau_l = 10\%$, which together implies a sizeable deduction. Taxing income from capital that has already been accumulated is nondistortionary. In addition, 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 available for consumption (since net investment is negative along the transition). The resulting tax system is therefore even more strongly geared towards capital income taxes.[38]

[38] See Greulich and Marcet (2008) for a recent analysis of optimal tax reforms in Chamley-Judd type models.

6.3.5 Government Debt

As discussed above Erosa and Gervais (2002) and Garriga (2003) prove theoretically that the optimal capital income tax in the steady state of an OLG model without idiosyncratic risk and type heterogeneity is zero if the tax schedule can differ by household age (and preferences are weakly separable between consumption and leisure and homothetic in consumption). In his quantitative work Garriga (2003) demonstrates, for our *non-separable* benchmark preference specification, that for particular values of the social discount factor of the Ramsey government the optimal steady state capital income tax is zero, but with implied large negative government debt positions. In this section, we therefore would like to discuss how our quantitative results are affected by relaxing the balanced budget assumption.

Government debt enters the steady state government budget constraint and the asset market clearing condition. We calibrate government debt such that, under the benchmark tax function, the debt/GDP ratio is given as specified in the first column of the table. As in the previous subsections the government chooses the optimal tax code by maximizing over a flat capital income tax τ_k, and a progressive labor income tax function defined by flat marginal tax τ_l above the deduction d_l.

When debt is negative, the government owns part of the physical capital stock in the economy and can use its interest income to partially finance government expenditures. With the additional revenue, the government can reduce tax distortions by lowering marginal tax rates or improve insurance and redistribution by increasing the exemption level d_l. In addition to this direct revenue effect, lower tax rates on capital and labor income induce more savings and higher labor supply, as shown in the last two columns of Table VI. This in turn enlarges the tax base and enables the government to further decrease the tax rates required to finance its expenditures. The results of the table indicate that optimal capital income taxes fall as government debt declines, but remains significantly positive even if the government holds massive amounts of positive assets. To obtain zero capital income taxes as optimal under our benchmark social welfare function the government would need to essentially own the entire capital stock in the economy (in which case it can finance most of its spending through returns on its assets).

Table VI: Government Debt and the Optimal Tax System[39]

Debt/GDP ratio	τ_k	τ_l	d_l	r	w	K	hours
-100%	24%	20%	\$13,400	3.43%	1.113	1.364	1.019
-20%	39%	20%	\$10,800	5.29%	1.025	1.073	1.004
0% (benchmark)	43%	20%	\$10,200	5.59%	1.000	1.000	1.000
20%	43%	20%	\$9,500	6.37%	0.982	0.938	0.985
100%	55%	21%	\$4,000	9.33%	0.885	0.703	0.984

[39] Wages, capital and hours are normalized by their values in the zero-debt optimum. To simplify comparisons across economies with different Debt/GDP ratios in this table we maximize over flat capital income taxes and flat labor income taxes with deduction, rather than the full Gouveia-Strauss tax function. This explains that the optimal capital income tax in this table differs from that reported in Table IV.

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 income risk generate a desire for redistribution and social insurance. We find that a system that taxes capital heavily and taxes labor income according to a flat tax with sizeable deduction is optimal in the long run.

We have argued that the key driving force behind the capital income tax result is the life cycle structure of our model in conjunction with endogenously chosen labor supply. We also show that the assumed labor supply elasticity is important for the large size of the optimal capital income tax, but not its existence. With the alternative preference specification it remains significantly different from zero.

Given our findings that the life cycle structure of our model is crucial for our results, future research should investigate how sensitive our findings are to a more detailed modelling of institutions affecting life-cycle labor supply and savings incentives, in particular the social security and Medicare system. Similarly, so far we have abstracted from any linkage between generations due to one- or two-sided altruism (see Fuster et al., 2007 for such a model and application to social security reform). In light of the classical results on zero optimal capital taxation in dynastic models it is conceivable, and subject to future research, that an incorporation of these elements into our model brings its implications for the optimal tax code somewhat closer to the classical results on optimal capital taxation.

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] Albanesi, S. and R. Armenter, 2007, Intertemporal Distortions in the Second Best, *Working Paper.*

[4] 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.

[5] Alvarez, Y., J. Burbidge, T. Farrell and L. Palmer, 1992, Optimal Taxation in a Life-Cycle Model, *Canadian Journal of Economics* 25, 111-122.

[6] 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.

[7] 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.

[8] Benabou, R., 2002, Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency, *Econometrica* 70, 481-517.

[9] Bohacek, R. and M. Kejak, 2004, Optimal Government Policies in Models with Heterogeneous Agents, *Working Paper*.

[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., *Working Paper*.

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

[12] Chari, V.V. and P. Kehoe, 1999, Optimal Fiscal and Monetary Policy, in J. Taylor and M. Woodford (eds.) *Handbook of Macroeconomics*, vol. 1, Elsevier.

[13] Conesa, J. C. and D. Krueger, 2006, On the Optimal Progressivity of the Income Tax Code, *Journal of Monetary Economics* 53, 1425-1450.

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

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

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

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

[18] Fuster, L., A. Imrohoroglu and S. Imrohoroglu, 200, Elimination of Social Security in a Dynastic Framework, *Review of Economic Studies* 74, 113-145.

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

[20] Gervais, M., 2004, On the Optimality of Age-Dependent Taxes and the Progressive U.S. Tax System, *Working paper.*

[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] Greulich, K. and A. Marcet, 2008, Pareto-Improving Optimal Capital and Labor Taxes, *Working paper.*

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

[25] Hansen, G., 1993, The Cyclical and Secular Behavior of the Labor Input: Comparing Efficiency Units and Hours Worked, *Journal of Applied Econometrics* 8, 71-80.

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

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

[28] Huffman, G., 2007, An Analysis of Fiscal Policy with Endogenous Investment-Specific Technological Change, *Working Paper.*

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

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

[31] Jones, L., R. Manuelli and P. Rossi, 1997, On the Optimal Taxation of Capital Income, *Journal of Economic Theory* 73, 93-117.

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

[33] Kimball, M. and M. Shapiro, 2005, Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?, *Working Paper*.

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

[35] 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.

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

[37] 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.

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

[39] Reiter, M., 2004, The Optimal Nonlinear Taxation of Capital in Models With Uninsurable Income Risk, *Working Paper*.

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

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

[42] Smyth, S., 2005, A Balancing Act: Optimal Nonlinear Taxation in Overlapping Generations Models, *Working paper.*

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

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

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

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

[47] Weinzierl, M., 2007, The Surprising Power of Age-Dependent Taxes, *Working Paper.*