Flodén, Martin

Working Paper

Why are capital income taxes so high?

SSE/EFI Working Paper Series in Economics and Finance, No. 623

Provided in Cooperation with:
EFI - The Economic Research Institute, Stockholm School of Economics

Suggested Citation: Flodén, Martin (2006) : Why are capital income taxes so high?, SSE/EFI Working Paper Series in Economics and Finance, No. 623, Stockholm School of Economics, The Economic Research Institute (EFI), Stockholm

This Version is available at:
https://hdl.handle.net/10419/56148

Why Are Capital Income Taxes So High?*

Martin Flodén**
Stockholm School of Economics and CEPR

SSE/EFI Working Paper Series in Economics and Finance
No. 623
March, 2006

Abstract

The Ramsey optimal taxation theory implies that the tax rate on capital income should be zero in the long run. This result holds even if the social planner only cares about workers that do not hold assets, or if the planner only cares about any other group in the economy. This paper demonstrates that although all households agree that capital income taxation should be eliminated in the long run, they do not agree on how to eliminate these taxes. Wealthy households would prefer a reform that is funded mostly by higher taxes on labor income while households with little wealth would prefer a reform that is funded mostly by high taxes on initial wealth. Pareto improving reforms typically exist, but the welfare gains of such reforms are modest.

JEL classification: E60; H21

Keywords: optimal taxation; inequality; redistribution

*Financial support from the Jan Wallander and Tom Hedelius Foundation at Svenska Handelsbanken is gratefully acknoledged.

**Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden, martin.floden@hhs.se.

1 Introduction

According to optimal taxation theory, the tax rate on capital income should be zero in the long run. Chamley (1986) and Judd (1985) first showed this, and the result has subsequently proven robust to a number of extensions and alternative assumptions. In particular, Judd (1985) and Chari and Kehoe (1999) show that this result holds even if the social planner only cares about workers that do not hold assets, or if the planner only cares about any other group in the economy.[1]

In addition to being theoretically robust, the implications of optimal taxation theory seem to be quantitatively important. Cooley and Hansen (1992) find that the welfare gain of eliminating capital taxes can amount to several percent of annual consumption, and Lucas (1990, p. 314) argues that the Ramsey optimal taxation literature has "generated the largest genuinely free lunch I have seen in 25 years in this business".[2] Still, capital income taxes remain high. Carey and Tchilinguirian (2000) document that the average capital income tax rate is 52 percent in the OECD countries if the tax is based on net operating surplus and 27 percent if it is based on gross operating surplus.[3]

The present paper provides some insights to why implementing the optimal tax policies is more difficult than previous studies acknowledge. In particular, I demonstrate that even though all groups agree that capital income taxes should be eliminated in the long run, the distributional effects of optimal tax reform may be important. And households that agree on what taxes should be in the long run, need not agree on how to get from today's tax system to a new steady state.

The idea to quantitatively evaluate the distributional effects of hypothetical tax reforms is not new. Auerbach and Kotlikoff (1987) examined how welfare of different cohorts would be affected if capital income taxed were replaced by higher consumption or labor income taxes in a life-cycle setting. In studies more related to the present, Garcia-Milà et al. (2001), Domeij and Heathcote (2004), and Nishiyama and Smetters (2005) examine the effects of tax reforms in the presence of income and wealth heterogeneity in dynastic settings. Common to all these studies is that they do not consider optimal tax reforms in the sense that the theoretical literature has analyzed. Instead they concentrate on once-and-for all reforms where new constant tax rates are suddenly implemented.[4] In the present paper, I follow the literature on Ramsey optimal tax reforms and solve for policies that maximize the utility of some particular (for example the representative) household in an economy with a realistic distribution of wealth and earnings.

I demonstrate that these policy reforms typically are not Pareto improving. For example, households with a high wealth to earnings ratio suffer a welfare loss equivalent to a 20 percent permanent reduction of consumption if the policy that maximizes the representa-

[1] I will focus on the Ramsey approach to optimal taxation. A growing literature, e.g. Kocherlakota (2005), uses insights from mechanism design theory to allow for more general tax systems where taxes may be nonlinear and conditional on income histories (the Mirrlees approach).

[2] See also Lucas (2003).

[3] The U.S. tax rates are close to the OECD average. Portugal has the lowest tax rates with 22 and 18 percent on net and gross surplus, respectively.

[4] These papers, in particular Garcia-Milà et al., also point out that welfare consequences of tax reform can vary substantially between households with different wealth to wage ratios.

tive household's utility is implemented. Although I rule out confiscatory taxation in the initial period, these wealthy households suffer from an extremely high tax on capital income in the second period.[5] I also consider optimal taxation under the restriction that the capital income tax rate may not be raised from the initial level. A policy that maximizes the representative household's utility then keeps the current tax rate on capital income for 28 years before the tax is eliminated, and the welfare effects of that policy are small except for a small fraction of households in the top and bottom of the wealth to earnings distribution.

The next section presents the theoretical framework. The key ingredients are a neoclassical production function with capital and labor; infinitely lived households that choose consumption and labor supply to maximize utility, and that are heterogenous with respect to initial wealth and skills; and economic policy that must satisfy a dynamic budget constraint. The framework abstracts from uncertainty and the skill heterogeneity is permanent. Section 3 describes the optimal taxation problem, and Section 4 describes how the model is parameterized to be consistent with U.S. data. Section 5 presents the results with an emphasis on distributional implications of tax reforms. The optimal taxation problem is solved for different social welfare functions, and with various restrictions on the tax paths. Finally, Section 6 concludes.

2 The Model

2.1 Households

The economy is populated by a unit mass of infinitely lived households that maximize life-time utility,

$$\sum_{t=0}^{\infty} \beta^t u\left(c_t, h_t\right) \tag{1}$$

where β is the time discount factor, u is the instantaneous utility function, c is consumption, and h is labor supply. Let r denote the interest rate and τ^k the tax rate on capital income, and let $R = 1 + \left(1 - \tau^k\right) r$ denote the gross after tax interest rate. The households' budget constraint is then

$$a_{t+1} = R_t a_t + \left(1 - \tau_t^h\right) w_t z h_t - \left(1 + \tau_t^c\right) c_t \tag{2}$$

where a_{t+1} denotes savings from period t to period $t+1$, τ^h is the labor-income tax rate, w is the wage rate, z is the household's labor productivity, and τ^c is the consumption tax. The per-period budget constraints can also be combined as

$$\sum_{t=0}^{\infty} q_t \left(1 + \tau_t^c\right) c_t = \sum_{t=0}^{\infty} q_t \left(1 - \tau_t^h\right) w_t z h_t + R_0 a_0 \tag{3}$$

[5] Atkeson, Chari, and Kehoe's (1999) survey of the Ramsey optimal taxation literature was given the subtitle "good news for capitalists" in the printed version. If "capitalists" is interpreted as those holding much capital, the title is totally misleading.

where the price of consumption in the first period is normalized to unity, $q_0 = 1$, and $q_{t+1} = q_t / R_{t+1}$.

Households differ with respect to labor productivity z, and initial asset holdings a_0, but have identical preferences. Following Greenwood et al. (1988) I assume that the utility function is

$$u(c,h) = \frac{\left(c - \zeta \frac{h^{1+1/\gamma}}{1+1/\gamma}\right)^{1-\mu}}{1-\mu}. \tag{4}$$

where μ can be thought of as the degree of risk aversion, and γ is the intertemporal labor supply elasticity.

Using the households' first order conditions,

$$\frac{u_{ht}}{u_{ct}} = \frac{-\left(1-\tau_t^h\right) w_t z}{1+\tau_t^c} \tag{5}$$

and

$$u_{ct} = \beta R_{t+1} u_{ct+1} \frac{1+\tau_t^c}{1+\tau_{t+1}^c}, \tag{6}$$

the budget constraint can be rewritten as the implementability constraint

$$\sum \beta^t \left[u_{ct} c_t + u_{ht} h_t\right] = \frac{u_{c0} R_0 a_0}{1+\tau_0^c}. \tag{7}$$

2.2 Production

The representative firm is a price taker and chooses factor inputs K and L on a competitive market to maximize profits,

$$\max F(K,L) - wL - (r+\delta) K$$

where $F(K,L) = K^\theta L^{1-\theta}$ is the production function, K is the aggregate capital stock, L is efficiency units of labor, and δ is the depreciation rate of capital.

2.3 The Government

Government spending is exogenously fixed at the per capita level G, and financed by taxes on labor earnings, capital income, and private consumption. All taxes are proportional and tax rates are identical for all agents. The government's budget constraint is then

$$D_{t+1} = R_t D_t + G - \tau_t^h w_t L_t - \tau_t^k r_t K_t - \tau^c C_t, \tag{8}$$

where D is public debt.[6]

[6] Only policies with a constant consumption tax will be considered, so the time subindex on τ^c will be ignored.

2.4 Equilibrium

Let $s = (z, a_0)$ denote a household's productivity and initial wealth, and let $\lambda(s)$ denote the measure of households over initial states. Following Atkeson et al. (1999) let $\pi_t = \left(\tau_t^h, \tau_t^k, \tau^c\right)$ denote the tax policy in period t, let $x_t = (c_t(s), h_t(s), a_t(s))$ denote household allocations, and let $p_t = (r_t, w_t)$ denote factor prices. Let also $\Pi = \{\pi_t\}_{t=0}^{\infty}$, $X = \{x_t\}_{t=0}^{\infty}$, $P = \{p_t\}_{t=0}^{\infty}$, and $D = \{D_t\}_{t=0}^{\infty}$ denote the paths for policy, allocations, factor prices, and public debt. For future reference, let also $A_t = \int a_t(s)\, d\lambda$ and $C_t = \int c_t(s)\, d\lambda$ denote aggregate asset holdings and consumption in period t.

Before defining a competitive equilibrium in this environment, it will be useful to introduce some further notation. Definition 1 therefore defines factor prices, household decisions, and asset and debt allocations as functions of the tax policy. Definition 2 then provides the definition of a competitive equilibrium, and Definition 3 provides the definition of a feasible government policy.

Definition 1 An allocation rule $\mathbf{X}$, a price rule $\mathbf{P}$, and a debt rule $\mathbf{D}$ *map a policy* Π *into an allocation* $X = \mathbf{X}(\Pi)$*, a price system* $P = \mathbf{P}(\Pi)$*, and a path for public debt* $D = \mathbf{D}(\Pi)$ *such that*

1. *The households' consumption, labor supply, and savings decisions* X *solve the households' optimization problem given the policy* Π*.*
2. *The representative firm's capital and labor input solve the firm's optimization problem in all periods* t*, i.e.*
$$F_K(K_t, L_t) = r_t + \delta$$
and
$$F_L(K_t, L_t) = w_t$$
where the aggregate capital stock is $K_t = A_t - D_t$ *and where aggregate efficiency units of labor supply is* $L_t = \int z h_t d\lambda$*.*
3. *Public debt evolves according to the public budget constraint (8) where initial debt* D_0 *is given.*

Definition 2 *A* competitive equilibrium *consists of a measure* λ *of households over initial states, a policy* Π*, household allocations* $X = \mathbf{X}(\Pi)$*, a price system* $P = \mathbf{P}(\Pi)$*, a path for public debt* $D = \mathbf{D}(\Pi)$*, and a level of government consumption* G*, such that.*

1. *The government's budget constraint is fulfilled and Ponzi schemes are ruled out, i.e.*
$$\sum q_t G = \sum q_t \left(\tau_t^h w_t L_t + \tau_t^k A_t + \tau^c C_t\right).$$
2. *The economy's resource constraint*
$$C_t + G + K_{t+1} = F(K_t, L_t) + (1 - \delta) K_t \tag{9}$$
is fulfilled in all periods t*.*

Definition 3 *A government policy* Π *is* feasible *if* $(\lambda, \Pi, \mathbf{X}(\Pi), \mathbf{P}(\Pi), \mathbf{D}(\Pi), G)$ *constitutes a competitive equilibrium.*

2.5 Aggregation and Disaggregation – The Representative Household

Define $Z = \left(\int z^{1+\gamma} d\lambda\right)^{\frac{1}{1+\gamma}}$ and $A_0 = \int a_0 d\lambda$. The utility function (4) then allows us to capture the economy's aggregate consumption and efficient labor supply by the behavior of a representative agent with productivity Z and initial assets A_0. Propositions 1 and 2 below demonstrate this.

Proposition 1 *A household with productivity* $Z = \left(\int z^{1+\gamma} d\lambda\right)^{\frac{1}{1+\gamma}}$ *supplies* $L_t = \int z h_t d\lambda$ *efficiency units of labor.*

Proof. The intratemporal first order condition (5) implies that

$$h_t(z) = \left[\frac{\left(1-\tau_t^h\right) w_t z}{\zeta\left(1+\tau_t^c\right)}\right]^{\gamma}. \tag{10}$$

A household with productivity Z thus supplies

$$L_t = h_t(Z) Z = \left[\frac{\left(1-\tau_t^h\right) w_t}{\zeta\left(1+\tau_t^c\right)}\right] Z^{1+\gamma}$$

efficiency units of labor. From the definition of Z we thus get

$$L_t = \left[\frac{\left(1-\tau_t^h\right) w_t}{\zeta\left(1+\tau_t^c\right)}\right] \int z^{1+\gamma} d\lambda. \tag{11}$$

We want to show that $L_t = \int z h_t(z) d\lambda$. From (10) we get that

$$\int z h_t(z) d\lambda = \int z \left[\frac{\left(1-\tau_t^h\right) w_t z}{\zeta\left(1+\tau_t^c\right)}\right]^{\gamma} d\lambda = \left[\frac{\left(1-\tau_t^h\right) w_t}{\zeta\left(1+\tau_t^c\right)}\right]^{\gamma} \int z^{1+\gamma} d\lambda$$

which equals L_t according to (11). ■

Proposition 2 *A household with productivity* $Z = \left(\int z^{1+\gamma} d\lambda\right)^{\frac{1}{1+\gamma}}$ *and initial wealth* $A_0 = \int a_0 d\lambda$ *consumes* $C_t = \int c_t(z, a_0) d\lambda$ *and holds wealth* $A_t = \int a_t(z, a_0) d\lambda$.

Proof. The first part of the proof demonstrates that the budget constraint for a household with productivity Z and initial wealth A_0 is identical to the aggregate of all households' budget constraints. The second part of the proof demonstrates that the households' Euler equations imply a path for aggregate consumption that is identical to the path implied by the Euler equation for the household with productivity Z and initial wealth A_0.

Integrate the budget constraint (3) over all households to get

$$\int \sum_{t=0}^{\infty} q_t \left(1+\tau^c\right) c_t d\lambda = \int \sum_{t=0}^{\infty} q_t \left(1-\tau_t^h\right) w_t z h_t d\lambda + \int R_0 a_0 d\lambda.$$

By using $\int z h_t d\lambda = Z h_t\left(Z\right)$ from Proposition 1, this aggregate budget constraint can be rewritten as

$$\sum_{t=0}^{\infty} q_t \left(1+\tau^c\right) C_t d\lambda = \sum_{t=0}^{\infty} q_t \left(1-\tau_t^h\right) w_t Z h_t\left(Z\right) + R_0 A_0$$

which is also the budget constraint for an agent with initial states (Z, A_0).

Using (4) and (10) in the Euler equation (6) gives

$$c_{t+1}\left(z, a_0\right) - \zeta\left(\frac{\left(1-\tau_{t+1}^h\right) w_{t+1} z}{\zeta\left(1+\tau^c\right)}\right)^{1+\gamma} = \left(\beta R_{t+1}\right)^{\frac{1}{\mu}} \left[c_t\left(z, a_0\right) - \zeta\left(\frac{\left(1-\tau_t^h\right) w_t z}{\zeta\left(1+\tau^c\right)}\right)^{1+\gamma}\right].$$

Integrate over all households to get

$$C_{t+1} - \zeta\left(\frac{\left(1-\tau_{t+1}^h\right) w_{t+1} Z}{\zeta\left(1+\tau^c\right)}\right)^{1+\gamma} = \left(\beta R_{t+1}\right)^{\frac{1}{\mu}} \left[C_t - \zeta\left(\frac{\left(1-\tau_t^h\right) w_t Z}{\zeta\left(1+\tau^c\right)}\right)^{1+\gamma}\right]$$

which is also the Euler equation for a household with initial states (Z, A_0). The budget constraint and Euler equation for a household with initial states (Z, A_0) are thus identical to the economy aggregates, and it follows that this household's consumption and wealth paths are identical to the economy's aggregate consumption and wealth paths. ■

As a direct consequence of Propositions 1 and 2, a policy Π is feasible in the heterogenous-agents economy if and only if the policy is feasible in the economy populated by a single representative agent with initial states (Z, A_0). Furthermore, the households' first order conditions (5) and (6), and their implementability constraints (7), provide a mapping from the representative-agent economy to allocations in the disaggregated heterogenous-agents economy. Proposition 3 summarizes these statements.

Proposition 3 *Consider a representative-agent economy with allocations X^{RA} and implied prices P. If X^{RA} and P fulfill the resource constraint (9) and the implementability constraint (7), then (i) there is a unique policy Π such that $X^{RA} = \mathbf{X}\left(\Pi\right)$ and $P = \mathbf{P}\left(\Pi\right)$, and $(\lambda^{RA}, \Pi, X^{RA}, P, \mathbf{D}\left(\Pi\right), G)$ constitutes a competitive equilibrium for the representative-agent economy; and (ii) there is a unique allocation $X = \mathbf{X}\left(\Pi\right)$ such that $(\lambda, \Pi, X, P, \mathbf{D}\left(\Pi\right), G)$ constitutes a competitive equilibrium for the disaggregated economy.*

3 Optimal Tax Policies

I will now consider optimal policies. Throughout I assume that the government has access to a commitment technology so that time-inconsistency problems can be ignored. To find the optimal policy, I use the primal approach and let the government choose an allocation

X^{RA} for the representative agent under the additional constraint that these sequences are consistent with household optimization.[7] As noted in Proposition 3, a policy that is feasible in the representative-agent economy is also feasible in the heterogenous-agents economy, and there is a unique disaggregated allocation that is implied by that policy.

In the baseline policy experiments, the consumption tax rate is fixed at its initial level, and I assume that the capital income tax rate cannot be changed in the first period.[8,9] Assume that tax policy is chosen to maximize the welfare of an agent with initial state s.[10] The Ramsey allocation problem is then

$$\max_{X^{RA}} \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(s\right), h_t\left(s\right)\right)$$

subject to the resource constraint (9) and the implementability constraint for the representative household,

$$\sum \beta^t \left[u_{Ct} C_t + u_{Ht} H_t\right] = \frac{U_{C0} R_0 A_0}{1+\tau_0^c},$$

and where the household choices $c_t\left(s\right)$ and $h_t\left(s\right)$ are part of the allocation X that is implied by X^{RA}.

After the model has been calibrated (next section) the system of first order conditions to this problem is solved numerically.[11]

4 Calibration and the Initial Steady State

Policy variables and parameter values for the baseline model are chosen to match U.S. data. One model period corresponds to one year, the capital share in production is 0.40, the depreciation rate of capital is 0.10, and the discount factor is chosen to obtain a capital to output ratio of 3.0 in the initial steady state. In the utility function, the degree of risk aversion is set to two and the labor-supply elasticity is set to 0.5. The weight on leisure is chosen so that hours worked is 1/3 in the initial steady state. The initial public debt is 60 percent of output, the consumption tax is 6.1 percent, and initial tax rates on capital and labor income are 31.1 and 22.6 percent, respectively.[12] Government spending is chosen so

[7] See Chari and Kehoe (1999) and Ljungqvist and Sargent (2004, chapter 15) for an overview of the primal approach to Ramsey optimal taxation.

[8] There is a continuum of tax policies that implement the optimal allocation if all three tax rates are choice variables.

[9] This assumption is standard in the literature and used to rule out lump sum taxation. There are a number of valid objections to this assumption. For example, high capital income taxes in the second period are close to lump sum taxation. And in the current setting there is no need to rule out lump sum taxation since distributional effects are considered – if lump sum taxation is efficient and all agents agree on this, it should be used.

[10] I only consider policies that maximize welfare of one particular agent. It would be interesting to also consider policies that maximize, for example, the utilitarian welfare function but this turns out to be computationally infeasible. I have, however, examined policies that maximize the utility of small groups of households. The insights are then similar to those with only one optimized household.

[11] See Appendix A for further details on the solution method. The economy is assumed to have reached a new steady state T periods after the policy change. For most specifications, I use $T = 150$.

[12] These tax rates are from table 4 in Carey and Tchilinguirian (2000).

that the government budget balances in the initial steady state. Table 1 summarizes the parameter values used in the baseline specification of the model and calibrated quantities and variables in the initial steady state.

Table 1: Parameter values and initial steady state

Parameters		Policy		Initial values	
μ	2.000	τ^k	0.311	$\frac{K}{Y}$	3.000
γ	0.500	τ^h	0.226	H	0.333
ζ	8.194	τ^c	0.061	r	0.033
β	0.978	$\frac{G}{Y}$	0.184	$\frac{D}{Y}$	0.600
θ	0.400				
δ	0.100				

4.1 Distributions

The government's policies can be found without knowing how labor productivity and initial wealth are distributed in the population, but to evaluate the distributional effects of policy choices, these distributions must be specified. I choose these distributions to match the facts on U.S. inequality reported in Budría Rodríguez et al. (2002). The distribution of initial wealth holdings is approximated by 100 values representing the different percentiles. To choose these values, I interpolate between the 11 observations from the Lorenz curve for wealth reported in Budría Rodríguez et al. (see Table 2).

Budría Rodríguez et al. also report data on average earnings for different wealth groups. One approach to calibrating the productivity distribution would be to calculate productivity for these wealth groups from the average earnings reported in Table 2. That approach, however, implies an earnings distribution that is too compressed (Gini 0.33 rather than 0.61) and too correlated with wealth (correlation 0.95 rather than 0.47) compared to what Budría Rodríguez et al. report. Instead, I allow three different earnings levels for each wealth percentile. These earnings levels and the mass of households allocated to each of them is chosen under the constraint that the average earnings for the different wealth groups equals that in Table 2. Furthermore, I follow an algorithm described in Appendix B to choose the distributions so that the earnings Gini is 0.61, the correlation between earnings and wealth is 0.47, and the mean-to-median ratio for earnings is 1.57, all values being identical to those reported by Budría Rodríguez et al. for U.S. data. Table 3 summarizes some properties of the calibrated wealth and earnings distributions. Note that the calibrated distributions also match the facts reported in Table 2.

Table 2: Distributions

	Percentiles, ranked by wealth										
	1	2–5	6–10	11–20	21–40	41–60	61–80	81–90	91–95	96–99	100
wealth	−0.20	−0.02	0.00	0.00	0.06	0.25	0.61	1.26	2.26	5.78	34.7
earnings	0.90	0.55	0.24	0.37	0.65	0.83	0.99	1.30	1.58	3.15	9.00

The table shows wealth and earnings relative to the average for different wealth percentiles. For example, a typical household in the second wealth percentile has $a = -0.02\bar{a}$ and earnings equal to 55% of the average. Source: Budría Rodriguez et al. (2002).

Table 3: Summary statistics of initial wealth and earnings distributions

	Gini	$\frac{\text{Mean}}{\text{Median}}$	$\frac{\text{Min}}{\text{Mean}}$	$\frac{\text{Max}}{\text{Mean}}$	correlation with wealth	correlation with earnings
wealth	0.80	3.93	−0.20	34.73	1.00	0.47
earnings	0.61	1.57	0.12	18.00	0.47	1.00

5 Distributional Effects of Tax Reforms

In this section, I examine the distributional effects of different tax reforms, with particular focus on Ramsey optimal tax reforms. A household's welfare gain of a policy reform is measured by the constant percentage that consumption must be increased in all periods in the original economy for the household to be as well off as in the reformed economy. Utilitarian welfare gains are similarly measured by the percentage increase in all households' consumption that makes the average life-time utility in the benchmark economy identical to the average life-time utility in the reformed economy.

Let me first fix the consumption tax at its initial level and only consider changes in capital and labor income taxes. Table 4 shows the implications of tax reforms that maximize different households' welfare under different constraints on the policies allowed. Consider first the outcome when the representative household's utility is maximized under the constraint that the capital income tax rate cannot be changed in the first period (the column marked as reform 1 in Table 4). The optimal policy is then to reduce the labor income tax from 22.8 percent to 5.2 percent in the first period, and to raise the capital income tax dramatically, to 1634 percent, in the second period. The labor income tax rate is held almost constant at 16.6 percent from the second period, while the capital income tax slowly falls from 1.4 percent in the third period towards zero. This policy raises the representative household's welfare by 1.5 percent, and a majority (70 percent) of households in the economy benefit from this policy reform. But initially wealthy households are hurt by the high capital tax in the second period. The wealthiest household would be prepared to give up 40.2 percent of its annual consumption to avoid the policy reform.[13]

[13] The 'wealth poor' household in Tables 4–7 has the lowest initial wealth to earnings ratio. This household has wealth from the bottom percentile (−20 percent of the average), and the lowest earnings (12 percent of the average). The 'wealth rich' household has the highest initial wealth to earnings ratio. This

Reforms 2 and 3 optimize with respect to a household that has the representative household's productivity but 70 or 130 percent of average wealth. There are still substantial welfare implications. When optimizing with respect to the household with little wealth, a majority of households still benefit from the implied policy but wealthy households suffer dramatically. Only 31 percent of households in the economy would benefit from a policy that maximizes the utility of the household with 130 percent of average wealth.

Reforms 4 and 5 show the interval of Pareto improving policies. These policies maximize the utility of a household that has 25.6 to 25.9 percent more wealth than the average household. Note that the welfare effects of these policies are modest, although not negligible. Although welfare effects are modest, the tax reforms imply substantial reallocations between capital and labor income and over time. Small deviations from the Pareto improving reforms may therefore have important welfare consequences (see for example reform 3).

Arguably, the policies implied by reforms 1 to 5 are unrealistic in that they allow for very high capital tax rates. Taxes above 100 percent can be avoided if households withdraw capital from the market, and if households have some control of the timing of capital returns, temporary high tax rates below 100 percent may also be infeasible. In reforms 6 to 9, capital income tax rates are restricted not to exceed the initial tax rate.[14] The welfare effects are then small, and in most scenarios the optimal policy is to let the capital tax rate remain at the present level for several decades. For example, when maximizing the representative household's utility, the optimal policy is to keep the capital tax at 31.1 percent for 28 years before it is cut to zero. Committing to policies that reduce taxes far in the future may be difficult in practice.

Theory says that the capital income tax should be zero in the new steady state. Proponents of low capital income taxation sometimes use this theoretical result to argue that capital taxes should be abolished immediately. The final column in Table 4 shows that only 31 percent of households would benefit from such a policy reform, and households with little wealth would suffer substantial welfare losses. Under the Ramsey policy, the government initially taxes capital returns heavily and thereby reduce government debt and accumulate assets. This public wealth enables the government to reduce the tax on labor income. But when initially high capital taxes are not allowed, the eliminated capital tax must be compensated by higher taxes on labor income and this hurts households with a high wage to wealth ratio.

Note that even the representative household dislikes a policy that immediately eliminates capital income taxation. Previous studies report mixed results on this issue. In a representative agent economy, Chari et al. (1994) found a small positive welfare gain in their benchmark economy with log utility, but a small welfare loss under high risk aversion. Domeij and Heatcote (2004) found a clear welfare gain (1.5 percent) when labor supply is exogenous. With endogenous labor supply, they report that only 25 percent of house-

household has wealth that is 3.0 times the average and earnings that are 12 percent of the average. Another wealth rich household has wealth equal to 9.5 times the average and earnings equal to 84 percent of the average. Welfare effects for this household are in general similar.

[14] Domeij and Klein (2005) argue that there may be implementation lags so that tax rates cannot be changed immediately. They demonstrate that the optimal capital tax never exceeds the initial rate if the lag is sufficiently long.

holds benefit from an immediate removal of capital income taxation, but the representative household could possibly belong to that group (since the median household has less wealth than the representative household). In the sensitivity analysis below, I only find a positive welfare effect on the representative household when the labor-supply elasticity is low.

So far, we have considered reforms that maximize welfare for households with the representative household's productivity. Median productivity in the economy is 74 percent of the representative household's productivity. The implications of policies that maximize welfare for households with median productivity have also been examined. For any given initial wealth position, these households have a higher wealth to wage ratio. The implied policies therefore put more emphasis on reducing capital taxation. A household with median productivity but average initial wealth would prefer a policy that initially subsidizes wealth holdings. Wealth poor households would suffer substantially from such policies, but even the representative household would be worse off.

A typical finding in the public finance literature is that consumption taxation is less distortionary and more efficient than income taxation.[15] The reforms considered in Table 5 are identical to the baseline reforms in Table 4, except that the consumption tax rate is raised from 6.1 percent (the U.S. level) to 17.1 percent (the OECD average) at the time of reform. In general, the welfare gains in Table 5 are somewhat higher than those in Table 4. There is thus some support for the claim that consumption taxes are less distortionary than income taxes, but the differences are small and the general conclusions from the baseline experiments still apply. More interestingly the results indicate some scope for a realistic Pareto improving tax reform. All households would benefit by a reform that immediately raises the consumption tax to 17.1 percent and that eliminates capital taxes after three to five years. The surprise increase in consumption taxes reduces the value of previously accumulated wealth and works as a substitute for higher capital income taxes.

The public finance literature also concludes that the deadweight loss of taxation increases more than proportionally in the tax rate, i.e. that additional taxation becomes more distortionary when taxes are already high. It is therefore possible that optimal taxation theory implies larger welfare gains for countries with higher tax rates than the U.S., where taxes are low in an international perspective. Table 6 reports results for optimal tax reforms for a country calibrated where initial tax rates are calibrated to match the European average. Carey and Tchilinguirian (2000) report that the average tax rate is 25.1 percent on capital income, 36.8 percent on labor income, and 18.7 percent on consumption income.[16]

As expected, the welfare effects of tax reforms under European policies reported in Table 6 are typically larger than in the experiments with U.S. policies reported previously. For example, the policy that is optimal for the resentative household now raises this household's welfare by 5.9 percent of annual consumption compared to only 1.5 percent under U.S. policies. But the negative welfare effects are also amplified, and the welfare gains of Pareto improving reforms are still small.

Table 7 summarizes the results of tax reforms that maximize the representative household's

[15] See Krusell et al. (1996) for references, and for an analysis of redistributional effects of different forms of taxation.

[16] Except for tax policies, the economy is still calibrated to match U.S. data.

Table 4: Optimal tax reform, U.S. policy, fixed τ^c, $z = 1$

	$\tau_0^k = \tau_{ss}^k$					$0 \leq \tau_t^k \leq \tau_{ss}^k$				$\tau_t^k \equiv 0$
	$0.7\bar{a}$	$\bar{a}$	$1.256\bar{a}$	$1.259\bar{a}$	$1.3\bar{a}$	$-0.05\bar{a}$	$0.7\bar{a}$	$\bar{a}$	$1.3\bar{a}$	$\bar{a}$
reform #	(2)	(1)	(4)	(5)	(3)	(9)	(7)	(6)	(8)	(10)
welfare gain[a]										
repr. household	1.1	1.5	0.4	0.3	−0.1	0.1	0.2	0.2	−0.1	−0.2
median household	8.2	6.7	0.4	0.3	−1.3	0.1	0.1	0.0	−1.3	−1.6
utilitarian	−4.1	−1.3	0.3	0.3	−0.1	0.1	0.2	0.2	−0.1	−0.2
wealth poor	49.8	37.3	0.6	0.0	−7.7	0.0	−0.4	−1.0	−7.7	−9.1
wealth rich	−55.5	−40.2	0.0	0.7	8.7	0.3	0.9	1.6	8.7	10.2
% gaining	69.1	69.7	100.0	100.0	31.2	100.0	98.9	53.0	31.2	30.9
new equilibrium[b]										
ΔK	23.9	21.9	15.6	15.5	14.0	13.8	14.9	15.1	14.0	13.6
ΔH	8.3	6.6	1.0	0.9	−0.3	−0.3	0.5	0.6	−0.3	−0.7
ΔC	13.9	11.5	3.8	3.6	1.9	1.9	3.0	3.2	1.9	1.4
gvt debt	−202.1	−142.5	37.4	40.5	79.0	76.7	54.1	50.4	79.0	89.4
wealth gini	72.1	75.7	78.5	78.6	78.8	75.1	77.3	78.0	78.8	78.6
policy[c]										
τ_1^k	1841.8	1633.8	403.5	376.3	24.9	31.1	31.1	31.1	24.9	0.0
$\tau_?^k \approx 0$	2.0	2.0	2.0	2.0	2.0	62.0	41.0	29.0	2.0	0.0
τ_∞^h	13.9	16.6	25.1	25.3	27.2	27.0	25.9	25.8	27.2	27.7

Notes: Optimization with respect to household with average productivity and wealth indicated by column head.

[a)] Percent of annual consumption, 'wealth poor' household has the lowest capital/productivity ratio, 'wealth rich' household has the highest capital/productivity ratio, 'median' household is the median with respect to welfare effects.

[b)] Δ indicates percent change from initial to new steady state, debt is in percent of output.

[c)] Tax rates in percent, $\tau_?^k$ is the first period t where $\tau_t^k \approx 0$.

Table 5: Optimal tax reform, U.S. policy, τ^c raised to OECD average (17.1%), $z = 1$

	$\tau_0^k = \tau_{ss}^k$					$0 \leq \tau_t^k \leq \tau_{ss}^k$					$\tau_t^k \equiv 0$
	$0.7\bar{a}$	$\bar{a}$	$1.264\bar{a}$	$1.268\bar{a}$	$1.3\bar{a}$	$0.7\bar{a}$	$\bar{a}$	$1.261\bar{a}$	$1.267\bar{a}$	$1.3\bar{a}$	$\bar{a}$
welfare gain[a]											
repr. household	1.2	1.6	0.4	0.3	0.0	0.5	0.5	0.4	0.3	0.2	0.2
median household	8.5	7.0	0.4	0.3	−1.0	1.3	1.2	0.4	0.3	−0.1	−0.1
utilitarian	−4.2	−1.4	0.4	0.3	0.0	0.4	0.4	0.4	0.3	0.2	0.2
wealth poor	50.9	38.9	0.7	0.0	−6.1	6.0	5.3	0.7	0.0	−2.0	−2.0
wealth rich	−56.4	−41.8	0.0	0.7	7.1	−5.8	−5.0	0.0	0.7	2.8	2.8
% gaining	69.4	69.7	100.0	100.0	31.6	71.8	72.5	100.0	100.0	42.7	42.7
new equilibrium[b]											
ΔK	24.1	22.2	15.6	15.5	14.3	16.2	16.3	15.6	15.5	15.1	15.1
ΔH	8.5	6.8	1.0	0.9	−0.1	1.7	1.7	1.0	0.9	0.6	0.6
ΔC	14.1	11.8	3.8	3.6	2.2	4.7	4.7	3.8	3.6	3.2	3.2
gvt debt	−194.8	−132.0	74.4	798.3	112.1	52.8	51.8	74.6	78.2	90.1	90.1
wealth gini	71.9	76.5	78.5	78.6	78.8	77.5	78.0	78.5	78.6	78.6	78.6
policy[c]											
τ_1^k	1826.5	1583.1	95.1	59.9	−261.2	31.1	31.1	31.1	31.1	0.0	0.0
$\tau_?^k \approx 0$	1.0	1.0	1.0	1.0	1.0	38.0	26.0	5.0	3.0	0.0	0.0
τ_∞^h	4.7	7.6	17.4	17.5	19.1	16.3	16.3	17.4	17.5	18.1	18.1

Notes: The consumption tax rate is raised from 6.1% to 17.1% at the time of reform. See also Table 4.

Table 6: Optimal tax reform, European policy, fixed τ^c, $z = 1$

	$\tau_0^k = \tau_{ss}^k$					$0 \leq \tau_t^k \leq \tau_{ss}^k$					$\tau_t^k \equiv 0$
	$0.7\bar{a}$	$\bar{a}$	$1.3\bar{a}$	$1.856\bar{a}$	1.867	$0.7\bar{a}$	$\bar{a}$	$1.3\bar{a}$	$1.440\bar{a}$	$1.540\bar{a}$	$\bar{a}$
welfare gain[a]											
repr. household	5.6	5.9	5.3	0.4	0.3	0.3	0.3	0.3	0.3	0.2	−1.2
median household	16.0	15.1	12.0	0.4	0.2	0.7	0.6	0.4	0.3	0.2	−2.4
wealth poor	89.7	80.7	59.6	0.8	0.0	3.3	2.4	1.3	0.6	0.0	−10.6
wealth rich	−60.5	−53.0	−37.5	0.0	0.5	−2.1	−1.3	−0.5	0.0	0.3	6.1
utilitarian	0.4	2.1	3.9	0.4	0.3	0.4	0.4	0.3	0.3	0.2	−1.5
% gaining	71.8	71.8	72.5	100.0	100.0	72.5	75.1	82.0	100.0	100.0	25.9
new equilibrium[b]											
ΔK	25.5	24.4	21.5	12.1	11.9	6.1	8.0	9.5	10.0	10.4	5.0
ΔH	12.9	11.8	9.3	0.7	0.5	−4.6	−2.9	−1.6	−1.1	−0.8	−6.2
ΔC	25.6	23.8	19.1	3.7	3.5	−5.8	−2.8	−0.4	0.5	1.1	−8.5
gvt debt	−170.3	−148.5	−94.3	59.3	61.5	134.1	112.1	93.5	86.2	81.6	152.6
wealth gini	65.7	78.4	72.3	77.4	77.4	70.6	72.3	74.0	74.8	75.4	72.4
policy[c]											
τ_1^k	1789.3	1731.6	1486.1	331.4	309.5	25.1	25.1	25.1	25.1	25.1	0.0
$\tau_?^k \approx 0$	1.0	1.0	1.0	1.0	1.0	65.0	55.0	43.0	38.0	33.0	0.0
τ_∞^h	22.8	24.2	27.7	38.6	38.8	44.9	42.9	41.4	40.8	40.4	46.8

Notes: See Table 4.

welfare under a number of alternative model parameterizations.[17] The first five result columns show implications of optimal tax reforms under the constraint that the capital income tax rate is fixed in the first period and the final five columns show implications of policies that immediately abolish capital income taxation. In the first of these columns, the labor supply elasticity is reduced to $\gamma = 0.1$. The most interesting implication of the lower elasticity is that the representative household now benefits from an immediate elimination of capital income taxes. This is consistent with Domeij and Heatcote (2004) who find that the representative agent benefits from an immediate elimination of capital taxation when labor supply is exogenous, and the result is intuitive since labor taxes become less distortionary (and thus more efficient relative capital income taxes) when labor supply is less elastic. The experiments with variations in the labor supply elasticity also show that optimal taxation theory and the potential welfare effects are more important when taxes are more distortionary. Here more distortions are generated by a higher labor supply elasticity. Previously we came to the same conclusion by raising the size of the public sector (see Table 6).

The other robustness checks presented in Table 7 are a lower capital to output ratio; a lower capital share in production; and a lower depreciation rate of capital. Again, these experiments affect the magnitudes of welfare effects but do not change the conclusion that distributional implications are important. A lower capital to output ratio and a lower depreciation rate of capital raise the welfare effects somewhat compared to the baseline specification, and a lower capital share in production reduces the welfare effects.

6 Concluding Discussion

Tax reforms implied by the Ramsey optimal taxation literature may have dramatic distributional effects. Although all households agree that capital income taxation should be eliminated in the long run, they do not agree on how to eliminate these taxes. Wealthy households would prefer a reform that is funded mostly by higher taxes on labor income while households with little wealth would prefer a reform that is funded mostly by high taxes on initial wealth.

If lump-sum taxes and transfers were available, resources could be reallocated between households so that any tax reform that raises the representative household's welfare would be Pareto improving. While lump-sum transfers may be feasible, the absence of lump-sum taxation is at the very heart of the optimal taxation literature. Note that lump-sum transfers would not be particularly useful in this setting. If such transfers were introduced with the tax reform, low-income households would benefit but the representative household and wealthy households would be worse off since the transfer would be financed by distortionary taxes. Lump-sum transfers could be useful if there were reforms that generated substantial welfare benefits for the representative household and for wealthy households, but none of the reforms I have considered have such implications.[18] Without

[17] The results in Table 7 should primarily be compared to those under reforms 1 and 10 in Table 4.

[18] Reform 5 in Table 4 raises welfare by 0.3 percent for the representative household and by 0.7 percent for the wealth rich household. The welfare gain would fall for both these households if lump-sum transfers were introduced.

Table 7: Sensitivity to parameterization, U.S. policy, fixed τ^c, optimize for representative household

	$\tau_0^k = \tau_{ss}^k$					$\tau_t^k \equiv 0$				
	$\gamma = 0.1$	$\gamma = 1.0$	$K/Y = 2.5$	$\theta = 0.36$	$\delta = 0.06$	$\gamma = 0.1$	$\gamma = 1.0$	$K/Y = 2.5$	$\theta = 0.36$	$\delta = 0.06$
welfare gain[a]										
repr. household	0.4	3.6	1.8	0.9	2.2	0.2	−1.0	−0.5	−0.1	−1.1
median household	2.5	10.0	8.1	4.3	10.0	−1.1	−2.3	−3.2	−0.6	−5.7
utilitarian	−0.8	1.3	1.9	−3.2	5.2	0.7	−1.4	−1.8	0.3	$-\infty$
wealth poor	14.7	47.3	60.7	19.5	119.7	−8.3	−9.5	−26.7	−2.9	$-\infty$
wealth rich	−16.3	−47.2	−34.5	−38.9	−31.8	10.1	8.9	14.2	6.0	18.1
% gaining	69.4	71.8	69.7	69.7	69.7	33.9	27.5	30.6	31.2	30.3
new equilibrium[b]										
ΔK	15.1	35.5	34.2	12.8	54.0	14.4	10.0	20.6	8.4	30.5
ΔH	0.6	18.8	9.2	3.8	13.0	−0.1	−3.9	−1.9	−0.2	−4.5
ΔC	3.1	28.2	17.9	6.1	29.8	2.3	−3.1	2.5	0.6	4.5
gvt debt	−39.3	−163.5	−89.7	−154.6	−99.6	78.0	111.2	100.2	77.5	133.0
wealth gini	78.4	71.5	75.1	76.5	73.2	79.1	76.9	77.6	79.2	-
policy[c]										
τ_1^k	943.4	1714.2	886.9	2572.3	737.3	0.0	0.0	0.0	0.0	0.0
$\tau_?^k \approx 0$	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
τ_∞^h	22.3	12.8	15.1	19.0	12.7	27.0	29.6	31.4	25.2	37.7

Notes: See Table 4.

lump-sum taxation, the potential welfare gains from Pareto improving tax reforms are therefore relatively modest.

The Pareto improving Ramsey reforms also rely on unrealistically high initial tax rates on capital income. If capital income taxes cannot be raised above the initial level, it is typically optimal to wait several decades before eliminating capital income taxation. Implementing reforms with such long pre-announcement periods may be difficult, for example because of commitment problems. The potential welfare gains of such reforms are also modest.

The welfare benefits from eliminating capital income taxation are, I argue, less obvious than what has been indicated in the Ramsey optimal taxation literature. In particular, an interesting and relevant theory of optimal taxation must integrate distributional concerns in the analysis. The new optimal taxation theory based on the Mirrlees approach (e.g. Kocherlakota, 2005) may provide a more complete analysis. The conclusion from that analysis may very well be that capital income should not be taxed. But the optimal tax system will then also explicitly handle redistribution between households.

Appendix A Details of the Optimal Taxation Problem

This appendix provides further details to the Ramsey optimal taxation problem formulated in Section 3. The problem is to solve

$$\max_{X^{RA}} \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(s\right), h_t\left(s\right)\right)$$

subject to the resource constraint (9), the implementability constraint for the representative household,

$$\sum \beta^t \left[u_{Ct} C_t + u_{Ht} H_t\right] = \frac{u_{C0} R_0 A_0}{1+\tau^c},$$

and household choices $c_t\left(s\right)$ and $h_t\left(s\right)$ that are part of the allocation X that is implied by X^{RA}.

It is convenient to reformulate the problem as finding allocations both for the representative agent and for the agent whose welfare is maximized. Let now c_t and h_t denote consumption and labor supply for the household with initial state s. The problem is then

$$\max_{X^{RA}, \{c_t, h_t\}} \sum_{t=0}^{\infty} \beta^t \int u\left(c_t, h_t\right) d\lambda$$

subject to the resource constraint (9), the implementability constraint for the representative household,

$$\sum \beta^t \left[u_{Ct} C_t + u_{Ht} H_t\right] = \frac{u_{C0} R_0 A_0}{1+\tau^c}, \tag{A.12}$$

the implementability constraint for the optimized household,

$$\sum \beta^t \left[u_{ct} c_t + u_{ht} h_t\right] = \frac{u_{c0} R_0 a_0}{1+\tau^c}, \tag{A.13}$$

and the constraint that all households face the same tax rates.

I follow Atkeson et al. (1999) and let

$$W\left(c_t, C_t, h_t, H_t, \lambda, \Lambda\right) = u\left(c_t, h_t\right) + \lambda\left(u_{ct} c_t + u_{ht} h_t\right) + \Lambda\left(u_{Ct} C_t + u_{Ht} H_t\right).$$

The optimization problem is then

$$\max \sum_{t=0} \beta^t W\left(c_t, C_t, h_t, H_t, \lambda, \Lambda\right) - \left(\lambda u_{c0} a_0 + \Lambda u_{C0} A_0\right) \frac{R_0}{1+\tau^c}$$

subject to the resource constraint

$$C_t + K_{t+1} + G = F\left(K_t, Z H_t\right) + \left(1-\delta\right) K_t, \tag{A.14}$$

and the constraints on identical tax rates,

$$\frac{u_{ct+1}}{u_{ct}} = \frac{u_{Ct+1}}{u_{Ct}}, \tag{A.15}$$

and

$$\frac{u_{ht}}{zu_{ct}} = \frac{u_{Ht}}{Zu_{Ct}}. \tag{A.16}$$

The Lagrangian to this problem is

$$\begin{aligned}
\pounds &= \sum_{t=0} \beta^t W\left(c_t, C_t, h_t, H_t, \lambda, \Gamma\right) - \left(\lambda u_{c0} a_0 + \Gamma u_{C0} A_0\right) \frac{R_0}{1+\tau^c} + \\
&\sum_{t=0} \beta^t \nu_t \left[F\left(K_t, ZH_t\right) + \left(1-\delta\right) K_t - C_t - K_{t+1} - G\right] + \\
&\sum_{t=0} \beta^t \rho_t \left[u_{ct} u_{Ct+1} - u_{Ct} u_{ct+1}\right] + \\
&\sum_{t=0} \beta^t \xi_t \left[z u_{ct} u_{Ht} - Z u_{Ct} u_{ht}\right]
\end{aligned}$$

where ν, ρ, and ξ are Lagrange multipliers. The first order conditions for c_t, C_t, h_t, and H_t are then (for $t > 0$)

$$W_{ct} + \rho_t u_{cct} u_{Ct+1} - \rho_{t-1} u_{Ct-1} u_{cct}/\beta + \xi_t \left[z u_{cct} u_{Ht} - Z u_{Ct} u_{cht}\right] = 0$$

$$W_{Ct} - \rho_t u_{CCt} u_{ct+1} + \rho_{t-1} u_{ct-1} u_{CCt}/\beta + \xi_t \left[z u_{ct} u_{CHt} - Z u_{CCt} u_{ht}\right] = \nu_t$$

$$W_{ht} + \rho_t u_{cht} u_{Ct+1} - \rho_{t-1} u_{Ct-1} u_{cht}/\beta + \xi_t \left[z u_{cht} u_{Ht} - Z u_{Ct} u_{hht}\right] = 0$$

$$W_{Ht} - \rho_t u_{CHt} u_{ct+1} + \rho_{t-1} u_{ct-1} u_{CHt}/\beta + \xi_t \left[z u_{ct} u_{HHt} - Z u_{CHt} u_{ht}\right] = -\nu_t Z F_{Lt}$$

and (for $t = 0$)

$$W_{c0} + \rho_0 u_{cc0} u_{C1} + \xi_0 \left[z u_{cc0} u_{H0} - Z u_{C0} u_{ch0}\right] = \frac{\lambda u_{cc0} R_0 a_0}{1+\tau^c}$$

$$W_{C0} - \rho_0 u_{CC0} u_{c1} + \xi_0 \left[z u_{c0} u_{CH0} - Z u_{CC0} u_{h0}\right] = \nu_0 + \frac{\lambda u_{CC0} R_0 A_0}{1+\tau^c}$$

$$W_{h0} + \rho_0 u_{ch0} u_{C1} + \xi_0 \left[z u_{ch0} u_{H0} - Z u_{C0} u_{hh0}\right] = \frac{\lambda \left(u_{ch0} R_0 + u_{c0} R_{h0}\right) a_0}{1+\tau^c}$$

$$W_{H0} - \rho_0 u_{CH0} u_{c1} + \xi_0 \left[z u_{c0} u_{HH0} - Z u_{CH0} u_{h0}\right] = -\nu_0 Z F_{L0} + \frac{\lambda u_{c0} R_{H0} a_0 + \Gamma \left(u_{CH0} R_0 + u_{C0} R_{H0}\right) A_0}{1+\tau^c}.$$

The first order conditions for K_{t+1} are

$$\beta \nu_{t+1} \left(F_{Kt+1} + 1 - \delta\right) = \nu_t,$$

while equations (A.12), (A.13), (A.14), (A.15), and (A.16) are the first order conditions for the multipliers (Λ, λ, ν_t, ρ_t, and ξ_t).

To solve this problem, I assume that the economy has reached a new steady state T periods after the tax reform. For most experiments, $T = 150$ turns out to work fine. For some experiments where taxes cannot be raised initially, I use higher T. Then I guess paths for C_t, H_t, K_{t+1}, c_t, h_t, and the multipliers ν_t, ρ_t, ξ_t, and Λ, and λ, and use an equation solver to find the equilibrium.

Appendix B Initial Distributions

When calibrating the model, I assume that households are characterized by a pair (i,j) where $i \in \{1,2,...,100\}$ indicates the household's position in the initial wealth distribution, and $j \in \{1,2,3\}$ indicates the household's labor productivity conditional on i. More specifically, a household of type (i,j) has initial wealth $a_0 = \alpha_i K$ and initial earnings $wzh_0 = e_{i,j}$. Let $\mu_{i,j}$ denote the mass of households of type (i,j).

I construct the grid $\mathbf{A} = \{\alpha_1, \alpha_2, ..., \alpha_{100}\}$ so that the wealth distribution replicates the facts reported in Budría Rodríguez et al. (2002). I use their Table 7 to calculate A_p for $p \in P = \{1,5,10,20,40,60,80,90,95,99,100\}$ and where $A_p \equiv \sum_{i=1}^{p} \alpha_i$. I then use piecewise cubic Hermite interpolation to calculate A_p for percentiles $p \notin P$. From these A_p, I calculate the implied α_i.

For each percentile i, I generate three different earnings levels $e_{i,j}$. I choose these earnings levels and the mass of households allocated to different states to replicate four sets of observations reported by Budría Rodríguez et al. (2002). They report that the Gini coefficient for earnings is 0.61, that the correlation between earnings and wealth is 0.47, that the mean-to-median ratio for earnings is 1.57, and they report how earnings is distributed across the different wealth groups in P. To find $e_{i,j}$ and $\mu_{i,j}$, I use the following algorithm.

1. For the 11 wealth groups p in P, calculate the average earnings (relative to total earnings) E_p from Table 7 in Budría Rodríguez et al. (2002).
2. Specify $\left(\underline{E}, \overline{E}\right) = (0.5 \min_p E_p, 2 \max_p E_p)$.
3. Guess parameters $(\mathtt{a}_s, \mathtt{b}_s) > (0,0)$ for $s = 1,2,3$.
4. Let $X = \{x_k\} = \{0.005, 0.015, ..., 0.995\}$ and let $Y = \{y_k\} = \{1/12, 2/12, ..., 11/12\}$. Let $\mathrm{B}(\cdot,\cdot)$ denote the beta function and let f_s denote the beta probability density function for parameters $(\mathtt{a}_s, \mathtt{b}_s)$, i.e. $f_s(x) = x^{\mathtt{a}_s - 1}(1-x)^{\mathtt{b}_s - 1} / \mathrm{B}(\mathtt{a}_s, \mathtt{b}_s)$.
5. Calculate the weights

$$\omega_{1,k} = \frac{f_1(y_k)}{1 + f_1(y_k)} \text{ for } k = 1,2,...,11$$

and

$$\omega_{2,k} = \frac{f_2(x_k)}{1 + f_2(x_k)} \text{ for } k = 1,2,...,100$$

$$\omega_{3,k} = \frac{f_3(x_k)}{1 + f_3(x_k)} \text{ for } k = 1,2,...,100.$$

6. Let $\mathbf{i}_p$ be a vector with indices to the percentiles included in wealth group $p \in P$, and let n_p denote the length of $\mathbf{i}_p$.[19]

[19] For example, $\mathbf{i}_5 = \begin{bmatrix} 2 & 3 & 4 & 5 \end{bmatrix}$ and $n_5 = 4$.

7. If $n_p = 1$, set $e_{p,1} = E_p$. If $n_p > 1$, calculate

$$\bar{\chi} = \frac{\overline{E} - E_p}{\overline{E} - \underline{E}}$$

and

$$\hat{\chi} = \min\left(\bar{\chi}, 1 - \bar{\chi}\right).$$

Then construct a linearly spaced $1 \times n_p$ vector $\boldsymbol{\chi}$ from $\bar{\chi} - \hat{\chi}\omega_{1,p}$ to $\bar{\chi} + \hat{\chi}\omega_{1,p}$. Let $e_{\mathbf{i}_p(i),1} = \chi_i \underline{E} + (1 - \chi_i)\overline{E}$.

8. For every percentile i, calculate

$$\chi = \frac{e_{i,1} - \omega_{2,i}\overline{E}}{1 - \omega_{2,i}}$$

and let

$$e_{i,2} = \max\left(\underline{E}, \min\left(e_{i,1}, \chi\right)\right)$$

and

$$e_{i,3} = \frac{e_{i,1} - (1 - \omega_{2,i})\, e_{i,2}}{\omega_{2,i}}.$$

9. For every percentile i, let the mass of households allocated to the different productivity levels be

$$\mu_{i,1} = \frac{1 - \omega_{3,i}}{100},$$

$$\mu_{i,2} = \frac{\omega_{3,i}\left(1 - \omega_{2,i}\right)}{100},$$

and

$$\mu_{i,3} = \frac{\omega_{2,i}\omega_{3,i}}{100}.$$

10. Calculate the Gini coefficient for earnings, the correlation between earnings and wealth, and the mean-to-median ration for earnings. If the values differ from those reported by Budría Rodríguez et al., use a minimization algorithm to update $(\mathtt{a}_s, \mathtt{b}_s)$ and repeat from 5.[20,21]

11. Use (10) to transform earnings $e_{i,j}$ to productivity $z_{i,j} = e_{i,j}^{\frac{1}{1+\gamma}}$.

References

Atkeson, Andrew, V. V. Chari, and Partick Kehoe (1999), "Taxing capital income: a bad idea", Federal Reserve Bank of Minneapolis Quarterly Review, summer, 3-17

Auerbach, Alan, and Laurence Kotlikoff (1987), *Dynamic fiscal policy*, Cambridge University Press, Cambridge, U.K.

[20] The algorithm implies that the average earnings in every wealth group is identical to the value reported by Budría Rodríguez et al., so we do not need to check this condition.

[21] I find the solution $(\mathrm{a}_1, \mathrm{b}_1, \mathrm{a}_2, \mathrm{b}_2, \mathrm{a}_3, \mathrm{b}_3) = (4.59, 0.44, 1.59, 0.67, 0.17, 1.40)$.

Budría Rodríguez, Santiago, Javier Díaz-Giménez, Vincenzo Quadrini, and José-Víctor Ríos-Rull (2002), "Updated facts on the U.S. Distributions of Earnings, Income, and Wealth", *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 2002

Carey David, and Harry Tchilinguirian (2000), "Average effective tax rates on capital, labour, and consumption", OECD Economics Department Working Paper No. 258

Chamley, Christopher (1986), "Optimal taxation of capital income in general equilibrium with infinite lives", *Econometrica* 54, 607-622

Chari, V. V., Lawrence Christiano, and Patrick Kehoe (1994), "Optimal fiscal policy in a business cycle model", *Journal of Political Economy* 102, 617-652

Chari, V. V., and Patrick Kehoe (1999), "Optimal fiscal and monetary policy", in: J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, North Holland: Amsterdam

Cooley, Thomas, and Gary Hansen (1992), "Tax distortions in a neoclassical monetary economy", *Journal of Economic Theory* 58, 290-316

Domeij, David and Jonathan Heathcote (2004), "On the distributional effects of reducing capital taxes", *International Economic Review* 45, 523-554

Domeij, David and Paul Klein (2005), "Pre-announced optimal tax reform", *Macroeconomic Dynamics* 9, 150-169

Garcia-Milà, Teresa, Albert Marcet, and Eva Ventura (2001), "Supply side interventions and redistribution", manuscript, Universitat Pompeu Fabra

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman (1988), "Investment, capacity utilization, and the real business cycle", *American Economic Review* 78, 402-417

Judd, Kenneth L. (1985), "Redistributive taxation in a simple perfect foresight model", *Journal of Public Economics* 28, 59-84

Kocherlakota, Narayana (2005), "Zero expected wealth taxes: a Mirrlees approach to dynamic optimal taxation", *Econometrica* 73, 1587-1621

Krusell, Per, Vincenzo Quadini, and José-Victor Ríos-Rull (1996), "Are consumption taxes really better than income taxes?", *Journal of Monetary Economics* 37, 475-503

Ljungqvist, Lars, and Thomas Sargent (2004), *Recursive macroeconomic theory*, second edition, MIT Press: Cambridge

Lucas, Robert E. Jr (1990), "Supply-side economics: an analytical review", *Oxford Economic Papers* 42, 293-316

Lucas, Robert E. Jr (2003), "Macroeconomic priorities", *American Economic Review* 93, 1-14

Nishiyama, Shinichi and Kent Smetters (2005), "Consumption taxes and economic efficiency with idiosyncratic wage shocks", *Journal of Political Economy* 113, 1088-1115