Why Are Capital Income Taxes So High?*

Martin Flodén**
Stockholm School of Economics and CEPR

June 13, 2007

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 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

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

*I thank two anonymous referees, David Domeij, and seminar participants at the University of Oslo and Uppsala University for helpful comments. Financial support from the Jan Wallander and Tom Hedelius Foundation at Svenska Handelsbanken is gratefully acknowledged.

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

[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).

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) examine 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 a representative-agent setting, Chari et al. (1994) find that most of the welfare gains from Ramsey optimal tax reforms are due to the high initial taxation of capital income. Although they do not directly address distributional implications, this finding indicates that the optimal policy may particularly benefit workers and be costly for capital owners.

In studies more closely related to the present, Garcia-Milà et al. (2001) and Domeij and Heathcote (2004) examine the effects of tax reforms in the presence of income and wealth heterogeneity in dynamic settings. These studies 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. They find that welfare consequences of tax reform can vary substantially between households with different wealth and earnings levels, and in particular Garcia-Milà et al. find that households with low wealth-to-earnings ratios suffer substantial welfare losses if the capital income tax is immediately abolished.[4] Correia (1999) also considers reforms that immediately implement new constant tax rates but in contrast to the studies above, she allows for capital levies in the initial period. She demonstrates that a removal of capital taxation raises inequality if the capital levy is small but reduces inequality and benefits less wealthy households if the levy is sufficiently high.[5]

In the present paper, rather than immediately abolishing capital-income taxation, I follow the literature on Ramsey optimal tax reforms and solve for the time paths of capital and labor-income taxes that maximize a social welfare function in an economy with a realistic distribution of wealth and earnings. I demonstrate that these tax reforms may have dramatic distributional effects and that they typically are not Pareto improving. Wealthy households would prefer a reform that is funded 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. As anticipated by Correia (1999), I therefore find that policies that are

[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]Idiosyncratic income is stochastic in Domeij and Heathcote (2004), implying a less direct relation between the current wealth-to-earnings ratio and welfare effects.

[5]In a setting similar to Domeij and Heathcote's (2004), Nishiyama and Smetters (2005) examine the effects of an increase in the consumption tax, which has similar implications as a tax levy followed by a lower income tax. They also find quantitatively important distributional implications.

optimal for households with low wealth-to-earnings ratios entail a high capital levy, reduce inequality, and imply substantial welfare losses for wealthy households. Even the policy that maximizes the representative household's welfare has such implications; households with a high wealth-to-earnings ratio suffer welfare losses up to the equivalent of a 40 percent permanent reduction of consumption under this policy. Although confiscatory taxation in the initial period is ruled out, these wealthy households suffer from the extremely high tax on capital income in the second period.[6]

Correia (1999) and Bassetto and Benhabib (2006) demonstrate that the median voter theorem holds in settings similar to the present if households only differ in initial wealth holdings, although policy is infinite-dimensional (capital and labor tax rates in infinitely many periods). I show that the median voter theorem also holds in the present setting where both initial wealth and productivity differ between households; the median voter is the household with the median wealth-to-earnings ratio in the initial equilibrium. This household has a much lower wealth-to-earnings ratio than the representative household and would consequently prefer a policy with a high initial tax on capital income. The median voter's policy therefore substantially reduces welfare for wealthy households.

A utilitarian social planner maximizes the average welfare in the economy. I find that the utilitarian policy is rather different from the policy chosen by the median voter. The utilitarian policy has less dramatic welfare effects, but the household with the highest wealth-to-earnings ratio still suffers a welfare loss of 10 percent of annual consumption. The potential utilitarian welfare gain is also quantitatively modest (0.5 percent of annual consumption). I further demonstrate that Pareto-improving reforms exist. Such reforms finance the removal of capital income taxation with a small initial capital levy and a small increase in labor-income taxes. These reforms imply modest but not negligible welfare gains for all households. Small deviations from the Pareto-improving reforms can however imply substantial welfare losses for some households.

The Ramsey approach to optimal taxation assumes that taxes are distortionary and rules out the use of lump-sum taxation. If discriminatory 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.[7] 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 non-discriminatory 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. Without 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

[6] 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.

[7] Garcia-Milà et al. (2001) demonstrate that abolished capital income taxation substantially raises all households' welfare if wealth at the same time can be redistributed from wealthy to less wealthy households.

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. For example, the policy that maximizes the representative household's utility then keeps the current tax rate on capital income for 29 years before the tax is eliminated. Implementing reforms with such long pre-announcement periods may be difficult, for example because of commitment problems. The potential welfare gains are also modest; less than 15 percent of the welfare gain remains for the representative household when tax rates cannot be raised.

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 presents the optimal taxation problem, demonstrates how to find the optimal policy for an arbitrary social welfare function, and demonstrates that the model has a median voter. Section 4 describes how the model is parameterized to be consistent with U.S. data, and 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. 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

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

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

$$\sum_{t=0}^{\infty}\beta^t\left[u_{ct}c_t+u_{ht}h_t\right]=\frac{u_{c0}R_0a_0}{1+\tau_0^c}. \tag{7}$$

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

$$D_{t+1}=R_tD_t+G-\tau_t^h w_tL_t-\tau_t^k r_tK_t-\tau^cC_t, \tag{8}$$

where D is public debt.[8]

[8] 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.*
$$\sum q_t G + R_0 D_0 = \sum q_t \left(\tau_t^h w_t L_t + \tau_t^k r_t K_t + \tau^c C_t\right).$$
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

The utility function (4) implies that the economy Gorman aggregates, i.e. that the aggregate behavior of the heterogenous households can be captured by the behavior of a representative household.[9] 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

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

A household with productivity Z thus supplies

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

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

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

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

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

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'

[9] For further details, see Correia (1999) who demonstrates this more carefully in a similar framework.

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

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

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

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

Integrate over all households to get

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

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 $X^{RA} = \mathbf{X}(\Pi)$ and $P = \mathbf{P}(\Pi)$, and $(\lambda^{RA}, \Pi, X^{RA}, P, \mathbf{D}(\Pi), G)$ constitutes a competitive equilibrium for the representative-agent economy; and (ii) there is a unique allocation $X = \mathbf{X}(\Pi)$ such that $(\lambda, \Pi, X, P, \mathbf{D}(\Pi), G)$ 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.[10] As noted in Proposition 3, a policy that is feasible in the representative-household economy is also feasible in the heterogenous-households economy, and there is a unique disaggregated allocation that is implied by that policy.

The consumption tax rate will be fixed at its initial level, and I assume that the capital income tax rate cannot be changed in the first period.[11,12] The planner chooses a policy that maximizes the welfare of a group of I households. Let s_i denote the initial state of household i, $s_i = (z_i, a_{i0})$, and let ω_i denote the planner's weight on this household's welfare. The Ramsey allocation problem is then

$$\max_{X^{RA}} \sum_{i=1}^{I} \omega_i \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(s_i\right), h_t\left(s_i\right)\right) \tag{12}$$

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

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

and where the household choices $c_t\left(s_i\right)$ and $h_t\left(s_i\right)$ are part of the allocation X that is implied by X^{RA}.[13]

Before turning to the quantitative analysis of this problem, let us further consider some analytical properties of this optimization problem. Consider first policies that maximize the welfare of a single household (i.e. $I = 1$). Although households differ in two dimensions (productivity and initial wealth) and have preferences over multi-dimensional policies (labor and capital tax rates in many time periods), their policy preferences can be ordered in one dimension as demonstrated in Proposition 4.[14]

Proposition 4 *Suppose that the policy Π^* solves the optimization problem for a household with initial state $s = (z, a_0)$. Then, for all α, Π^* also solves the optimization problem for all households with initial states $\hat{s} = \left(\hat{z} = \alpha z, \hat{a}_0 = \alpha^{1+\gamma} a_0\right)$.*

Note that (10) implies that $\hat{z}\hat{h}_t = \alpha^{1+\gamma} z h_t$ so that the proposition states that two households with identical wealth-to-earnings ratios in the initial steady state would prefer identical tax reforms. A consequence of Proposition 4 is that the economy has a median-voter

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

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

[12] This assumption is standard in the literature and used to rule out lump sum taxation. There are, however, 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.

[13] Details of the optimization problem are presented in the Appendix.

[14] The proofs of Propositions 4 and 5 are in the Appendix.

property; the policy preferred by the household with the median wealth-to-earnings ratio in the initial steady state will be chosen by a majority of the households in pairwise comparison to all other policies.

When more than one optimized household is considered in the optimization problem, i.e. when $I > 1$ in (12), solving the problem is computationally challenging. As stated in Proposition 5, however, the policy that maximizes an arbitrary solcial welfare function is also the optimal policy for some single household.

Proposition 5 *Suppose that the policy* Π^* *solves the optimization problem (12) for some welfare weights and initial states* $\{\omega_i, s_i\}_{i=1}^{I}$ *where* $I > 1$. *Then there is an initial wealth position* $\hat{a}_0$ *so that the policy* Π^* *is also optimal for a household with initial state* $\bar{s} = (\bar{z} = 1, \bar{a}_0 = \hat{a}_0)$.

This proposition does not provide a direct procedure for how to find this household (i.e. how to find $\hat{a}_0$). But as a consequence of Propositions 4 and 5, the optimal policy for the group of I households can be solved numerically by considering different candidates $\bar{a}_0$, and noting that the weighted welfare of the I households is single-peaked in $\bar{a}_0$.

We now turn to the quantitative analysis. After the model has been calibrated (next section) the system of first order conditions to this problem is solved numerically and the optimal policies chosen by different households or groups of households are analyzed.[15]

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.[16] Government spending is chosen so 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]

[15] 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 $T = 150$.

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

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]

[Table 3]

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 and Figure 1 show the implications of tax reforms that maximize different social welfare functions when the capital income tax rate cannot be changed in the first period. Consider first the outcome when the representative household's utility is maximized (column 2). 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 tax reform. But initially wealthy households are hurt by the high capital tax in the second period. The household with the highest wealth-to-earnings ratio would be prepared to give up 40.2 percent of its annual consumption to avoid the policy reform.[17]

As demonstrated in Section 3, the median voter in this economy is the household with the median wealth-to-earnings ratio or the median wealth-to-productivity ratio $a_0/z^{1+\gamma}$, which is $0.3746A_0$. The median voter thus has a much lower wealth-to-earnings ratio than the representative household. A clear majority of households benefit from the policy chosen by the median voter, but wealthy households suffer dramatically and the utilitarian welfare measure falls by almost six percent of annual consumption (column 1 in Table 4).

It turns out that a utilitarian social planner would choose the same policy as a planner maximizing the welfare of a single household with $a_0/z^{1+\gamma} = 1.200A_0$. The implications of this policy are shown in column 3 in Table 4. This reform holds the tax on labor income approximately constant and finances the removal of capital-income taxation mostly with high taxation of capital income in the second period. A majority of households benefit from the reform, but the utilitarian welfare only increases by 0.5 percent. Wealth poor households benefit from higher efficiency in terms of higher production and wages as the capital stock increases, but the high capital taxation in the second period still implies that welfare fall by up to 10 percent for wealthy households.

Columns 4 and 5 show the interval of Pareto improving policies. These policies maximize the utility of a household that has 25.1 to 25.4 percent more wealth than the average household and fund the removal of capital taxes both by initially taxing capital heavily and by raising taxes on labor income. Note that the welfare effects of these policies are modest, but 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 as is demonstrated in column 6 where policy maximizes utility for a household that has 30 percent more wealth than the average household. That policy funds the removal of capital taxation entirely through higher taxation of labor income and consequently reduces welfare for wealth-poor households that mostly rely on labor income. Only 31 percent of households in the economy would benefit from that policy. Figure 1 also illustrates the fast fall in the number of households gaining from the reform when initial wealth for the optimized household exceeds 25.4 percent.

[Table 4]

[Figure 1]

[17] The 'wealth poor' household in Tables 4–6 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 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.

Arguably, the policies implied by these experiments 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. Table 5 shows the implications of policy reforms that restrict the capital income tax rate not to exceed the initial tax rate.[18,19] 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 29 years before it is cut to zero. Committing to policies that reduce taxes far in the future may be difficult in practice, in particular when the potential welfare gains are small.

[Table 5]

The reforms considered in Table 5 are similar to reforms that must be pre-announced as in Domeij and Klein (2005). They consider exogenous implementation lags in a representative-agent setting and find that much of the welfare gains remain even if the tax reform must be announced many years in advance. In the present setting, the representative household's preferred policy with the 29 year delay in Table 5 implies a welfare gain that is less than 15 percent of the welfare gain when the implementation lag is just one year as in Table 4. One important reason for the lower welfare gain of delayed reforms in the present setting is that the initial tax on capital is lower (31 percent rather than 51 percent in Domeij and Klein).

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 5 shows that only 31 percent of households would benefit from such a policy reform. Welfare would fall for the representative household 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 low wealth-to-earnings ratio, which was also noted in Garcia-Milà et al. (2001).

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.[20]

[18] The theoretical result that all agents would set the capital income tax to zero hold if the economy eventually settles down in a steady state. As demonstrated in Bassetto and Benhabib (2006), this need not be the case when there is a cap on the capital income tax. In the experiments considered here, however, the capital income tax is always optimally set to zero in finite time.

[19] 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.

[20] They find that the welfare gain under the Ramsey policy is much larger and conclude that most of

Domeij and Heathcote (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 households 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.

A typical finding in the public finance literature is that consumption taxation is less distortionary and more efficient than income taxation.[21] I have also considered reforms like those in Tables 4 and 5 but where the 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 are then somewhat higher, 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 from 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.[22]

Table 6 summarizes the results of tax reforms that maximize the representative household's welfare under a number of alternative model parameterizations. 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 Heathcote (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.

The other robustness checks presented in Table 6 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.

---

the welfare gain comes from the initial capital levy. But ruling out the capital levy need not imply an immediate removal of capital taxation. Column 2 in Table 5 demonstrates that the welfare gains for the representative household are higher when a tax increase is ruled out than when capital taxation is immediately abolished as in column 4.

[21] See Krusell et al. (1996) and Nishiyama and Smetters (2005) for references, and for analyses of redistributional effects of different forms of taxation.

[22] One may thus argue that raises in the consumption tax should be ruled out on the same grounds as raises in the tax rate on first-period capital income.

[**Table 6**]

6 Conclusions

Garcia-Milà et al. (2001) and Domeij and Heathcote (2004) demonstrate that abolished capital income taxation can have important distributional implications when earnings and wealth differ across households. This paper reconfirmed that finding but also took one step further by considering the distributional implications of Ramsey optimal tax reforms. Judd (1985) demonstrates that all household – even households without wealth – would choose to abolish capital income taxation in the long run if they could choose their favorite Ramsey policy. In the present paper I demonstrate that although all households agree that capital income taxes should be eliminated in the long run, they do not agree on the tax policies in the transition. Wealthy households prefer policies that finance the removal of capital income taxes with raised taxes on labor income, while households with little wealth prefer reforms that are financed by high taxes on capital income in the first periods. Distributional effects of the Ramsey optimal reforms can therefore be quantitatively important. Reforms that maximize welfare for the median voter or the representative household would imply substantial welfare losses for wealthy households. Pareto improving reforms typically exist, but the welfare gains of such reforms are modest.

I therefore argue that the welfare benefits from eliminating capital income taxation are less obvious than what has been indicated in the Ramsey optimal taxation literature that has built on representative agents or ignored welfare effects in the transition to a new equilibrium. 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. We first consider optimization with respect to the welfare of a group of households. Then the brief proofs of Propositions 4 and 5 are presented and finally the numerical solution method is described. The problem with constraints on the tax rates and further details on the proofs are available in an appendix on my web page, www.hhs.se/personal/floden.

Appendix A.1 Optimization With Respect to Many Households' Welfare

The problem is to solve

$$\max_{X^{RA}} \sum_{i=1}^{I} \omega_i \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(s_i\right), h_t\left(s_i\right)\right)$$

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 $c_t\left(s_i\right)$ and $h_t\left(s_i\right)$ 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 household and for the households whose welfare is maximized. Let now c_{it} and h_{it} denote consumption and labor supply for the household with initial state s_i. The problem is then

$$\max_{X^{RA},\, \{c_{it}, h_{it}\}} \sum_{i} \omega_i \sum_{t=0}^{\infty} \beta^t u\left(c_{it}, h_{it}\right)$$

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

$$\sum_{t=0} \beta^t \left[u_{Ct} C_t + u_{Ht} H_t\right] = \frac{u_{C0} R_0 A_0}{1+\tau^c}, \tag{A.1}$$

the implementability constraints for the optimized households,

$$\sum_{t=0} \beta^t \left[u_{cit} c_{it} + u_{hit} h_{it}\right] = \frac{u_{ci0} R_0 a_{i0}}{1+\tau^c}, \tag{A.2}$$

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

I follow Atkeson et al. (1999) and let

$$W_i\left(c_{it}, h_{it}, \lambda_i\right) = \omega_i u\left(c_{it}, h_{it}\right) + \lambda_i\left(u_{cit} c_{it} + u_{hit} h_{it}\right)$$

and

$$W\left(C_t, H_t, \Lambda\right) = \Lambda\left(u_{Ct} C_t + u_{Ht} H_t\right).$$

The optimization problem is then

$$\max \sum_i \left[\sum_{t=0} \beta^t W_i\left(c_{it}, h_{it}, \lambda_i\right) - \frac{\lambda_i u_{ci0} R_0 a_{i0}}{1+\tau^c} \right] + \sum_{t=0} \beta^t W\left(C_t, H_t, \Lambda\right) - \frac{\Lambda u_{C0} R_0 A_0}{1+\tau^c}$$

subject to the resource constraint

$$C_t + K_{t+1} + G = F\left(K_t, ZH_t\right) + (1-\delta) K_t, \tag{A.3}$$

and the constraints on identical tax rates,

$$\frac{u_{cit+1}}{u_{cit}} = \frac{u_{Ct+1}}{u_{Ct}}, \tag{A.4}$$

and

$$\frac{u_{hit}}{z_i u_{cit}} = \frac{u_{Ht}}{Z u_{Ct}}. \tag{A.5}$$

The Lagrangian to this problem is then

$$\begin{aligned} \pounds \quad = \quad & \sum_i \left[\sum_{t=0} \beta^t W_i\left(c_{it}, h_{it}, \lambda_i\right) - \frac{\lambda_i u_{ci0} R_0 a_{i0}}{1+\tau^c} \right] + \sum_{t=0} \beta^t W\left(C_t, H_t, \Lambda\right) - \frac{\Lambda u_{C0} R_0 A_0}{1+\tau^c} + \\ & \sum_{t=0} \beta^t \nu_t \left[F\left(K_t, ZH_t\right) + (1-\delta) K_t - C_t - K_{t+1} - G \right] + \\ & \sum_i \sum_{t=0} \beta^t \rho_{it} \left[u_{cit} u_{Ct+1} - u_{Ct} u_{cit+1} \right] + \\ & \sum_i \sum_{t=0} \beta^t \xi_{it} \left[z_i u_{cit} u_{Ht} - Z u_{Ct} u_{hit} \right] \end{aligned}$$

where ν, ρ, and ξ are Lagrange multipliers. The first order conditions for c_{it}, C_t, h_{it}, and H_t are then (for $t > 0$)

$$W_{cit} + \rho_{it} u_{ccit} u_{Ct+1} - \rho_{it-1} u_{Ct-1} u_{ccit}/\beta + \xi_{it}\left[z_i u_{ccit} u_{Ht} - Z u_{Ct} u_{chit} \right] = 0 \tag{A.6}$$

$$W_{Ct} - \rho_{it} u_{CCt} u_{cit+1} + \rho_{it-1} u_{cit-1} u_{CCt}/\beta + \xi_{it}\left[z_i u_{cit} u_{CHt} - Z u_{CCt} u_{hit} \right] = \nu_t \tag{A.7}$$

$$W_{hit} + \rho_{it} u_{chit} u_{Ct+1} - \rho_{it-1} u_{Ct-1} u_{chit}/\beta + \xi_{it}\left[z_i u_{chit} u_{Ht} - Z u_{Ct} u_{hhit} \right] = 0 \tag{A.8}$$

$$W_{Ht} - \rho_{it} u_{CHt} u_{cit+1} + \rho_{it-1} u_{cit-1} u_{CHt}/\beta + \xi_{it}\left[z_i u_{cit} u_{HHt} - Z u_{CHt} u_{hit} \right] = -\nu_t Z F_{Lt} \tag{A.9}$$

and (for $t = 0$)

$$W_{ci0} + \rho_{i0} u_{cci0} u_{C1} + \xi_{i0}\left[z_i u_{cci0} u_{H0} - Z u_{C0} u_{chi0} \right] = \frac{\lambda_i u_{cci0} R_0 a_{i0}}{1+\tau^c} \tag{A.10}$$

$$W_{C0} - \sum_i \left[\rho_{i0} u_{CC0} u_{ci1} + \xi_{i0}\left(z_i u_{ci0} u_{CH0} - Z u_{CC0} u_{hi0} \right) \right] = \nu_0 + \frac{\Lambda u_{CC0} R_0 A_0}{1+\tau^c} \tag{A.11}$$

$$W_{hi0} + \rho_{i0} u_{chi0} u_{C1} + \xi_{i0}\left[z_i u_{chi0} u_{H0} - Z u_{C0} u_{hhi0} \right] = \frac{\lambda_i u_{chi0} R_0 a_{i0}}{1+\tau^c} \tag{A.12}$$

$$\begin{aligned} & W_{H0} - \sum_i \left[\rho_{i0} u_{CH0} u_{ci1} + \xi_{i0}\left(z_i u_{ci0} u_{HH0} - Z u_{CH0} u_{hi0} \right) \right] \\ = \quad & -\nu_0 Z F_{L0} + \frac{\sum_i \lambda_i u_{ci0} R_{H0} a_{i0} + \Lambda\left(u_{CH0} R_0 + u_{C0} R_{H0} \right) A_0}{1+\tau^c} \end{aligned} \tag{A.13}$$

while the first order conditions for K_{t+1} are

$$\beta \nu_{t+1}\left(F_{Kt+1} + 1 - \delta \right) = \nu_t \tag{A.14}$$

and equations (A.1), (A.2), (A.3), (A.4), and (A.5) are the first order conditions for the multipliers (Λ, λ, ν_t, ρ_{it}, and ξ_{it}).

Appendix A.2 Proofs

Proof of Proposition 4: Suppose that the policy Π^* solves (12) when the welfare of a single household ($I = 1$) with initial state $s = (z, a_0)$ and suppose that the solution to (A.1) to (A.14) is $(\Lambda, \lambda, \{C_t, H_t, K_{t+1}, c_t, h_t, \nu_t, \rho_t, \xi_t\}_t)$. Consider now a household with initial state $\hat{s} = \left(\hat{z} = \alpha z, \hat{a}_0 = \alpha^{1+\gamma} a_0\right)$, i.e. a household such that $\hat{a}_0/\hat{z}^{1+\gamma} = a_0/z^{1+\gamma}$. Let $\hat{\Lambda} = \alpha^{(1-\mu)(1+\gamma)}\Lambda$, $\hat{\lambda} = \lambda$, $\hat{c}_t = \alpha^{1+\gamma} c_t$, $\hat{h}_t = \alpha^{\gamma} h_t$, $\hat{\nu}_t = \alpha^{(1-\mu)(1+\gamma)}\nu_t$, $\hat{\rho}_t = \alpha^{1+\gamma}\rho_t$ and $\hat{\xi}_t = \alpha^{\gamma}\xi_t$. Then since $(\Lambda, \lambda, \{C_t, H_t, K_{t+1}, c_t, h_t, \nu_t, \rho_t, \xi_t\}_t)$ solve (A.1) to (A.14) when the optimized household has initial state s, it is straightforward to verify that $\left(\hat{\Lambda}, \hat{\lambda}, \left\{C_t, H_t, K_{t+1}, \hat{c}_t, \hat{h}_t, \hat{\nu}_t, \hat{\rho}_t, \hat{\xi}_t\right\}_t\right)$ solve (A.1) to (A.14) when the optimized household has initial state $\hat{s}$. Since the solutions to the two problems are characterized by the same aggregate variables C, H, and K, they are implemented by the same tax policies.

Proof of Proposition 5: Suppose that the policy Π^* solves (12) and suppose that the solution to (A.1) to (A.14) is $\left(\Lambda, \{C_t, H_t, K_{t+1}, \nu_t\}_t, \{\lambda_i, \{c_{it}, h_{it}, \rho_{it}, \xi_{it}\}_t\}_i\right)$. Define $y_i \equiv (u_{ci0}/u_{c10})^{1/\mu}$ and let $\bar{\lambda} = \left(\sum_i y_i^{\mu-1}\omega_i\right)^{-1}\left(\sum_i y_i^{\mu-1}\lambda_i\right)$. Let also $\bar{h}_t = z_1^{-\gamma} h_{1t}$.

Consider now optimization with respect to the welfare of a single household with initial state $\hat{s} = (\hat{z} = 1, \hat{a}_0)$, and denote the solution by $\left(\hat{\Lambda}, \hat{\lambda}, \left\{\hat{C}_t, \hat{H}_t, \hat{K}_{t+1}, \hat{c}_t, \hat{h}_t, \hat{\nu}_t, \hat{\rho}_t, \hat{\xi}_t\right\}_t\right)$. Note that there is some $\hat{a}_0 = \hat{a}_0^*$ that results in $\hat{\lambda} = \bar{\lambda}$. Find this $\hat{a}_0^*$ and set $\bar{a}_0 = \hat{a}_0^*$ and $\bar{c}_0 = \hat{c}_t^*$. Then calculate $\bar{u}_{c0} = u_c\left(\bar{c}_0, \bar{h}_0\right)$ and define $x_1 = (u_{c10}/\bar{u}_{c0})^{1/\mu}$.

Let now $\bar{c}_t = x_1 c_{1t} + \left(z_1^{-1-\gamma} - x_1\right)\zeta\,(h_{1t})^{1+1/\gamma}/(1+1/\gamma)$, $p = x_1^{\mu-1}\sum_i y_i^{\mu-1}\omega_i$, $\bar{\Lambda} = p^{-1}\Lambda$, $\bar{v}_t = p^{-1}v_t$, $\bar{\rho}_t = p^{-1}x_1^{\mu}\sum_i y_i^{\mu}\rho_{it}$, and $\bar{\xi}_t = p^{-1}x_1^{\mu}\sum_i z_i y_i^{\mu}\xi_{it}$. we can then verify that $\left(\bar{\Lambda}, \bar{\lambda}, \left\{C_t, H_t, K_{t+1}, \bar{c}_t, \bar{h}_t, \bar{\nu}_t, \bar{\rho}_t, \bar{\xi}_t\right\}_t\right)$ satisfy the first-order conditions (A.1) to (A.14) when optimization is respect to a single household with the initial state $\bar{s} = (\bar{z} = 1, \bar{a}_0)$. Since this solutions implies the same path for the aggregate variables C, H, and K as was implied when optimizing with respect to many households, the policy Π^* is also optimal for this stand-in household.

Appendix A.3 Solution Method

Proposition 5 implies that only optimization with respect to one household's welfare need to be implemented numerically. 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 but $T = 300$ has been used in all reported tables and graphs.

The first-order conditions (A.1) to (A.14) then provide $8T + 2$ equations in equally many unknown variables, $(\Lambda, \lambda, \{C_t, H_t, K_{t+1}, c_t, h_t, \nu_t, \rho_t, \xi_t\}_t)$. To solve this problem, I guess Λ and λ and paths for C_t, H_t, K_{t+1}, c_t, h_t, and the multipliers ν_t, ρ_t, and ξ_t, and use an equation solver to find the equilibrium.[23]

[23] Computer code is available on my webpage, www.hhs.se/personal/floden.

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 $\mathbf{i}_p$.[24]

[24] 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

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

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.[25,26]

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 Patrick 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.

[25] 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.

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

Bassetto, Marco, and Jess Benhabib (2006), "Redistribution, taxes, and the median voter", *Review of Economic Dynamics* 9, 211-223

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

Correia, Isabel (1999), "On the efficiency and equity trade-off", *Journal of Monetary Economics* 44, 581-603

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 Quadrini, 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

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				

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

Table 4: Optimal tax reform, $\tau_0^k = \tau_{ss}^k$

	policy optimal for household with $a_0 =$					
	median voter	repr. household	utilitarian	pareto	pareto	wealthy
$a_0 =$	$0.375A_0$	$1.000A_0$	$1.200A_0$	$1.251A_0$	$1.254A_0$	$1.300A_0$
column #	1	2	3	4	5	6
welfare gain[a]						
repr. household	0.6	1.5	0.8	0.4	0.3	−0.2
median voter	8.9	7.0	2.3	0.4	0.3	−1.5
utilitarian	−5.9	−1.3	0.5	0.3	0.3	−0.2
wealth poor	54.2	37.3	10.6	0.6	0.0	−8.8
wealth rich	−61.7	−40.2	−10.5	0.0	0.7	9.9
% gaining	68.8	69.7	71.8	100.0	100.0	30.9
new equilibrium[b]						
ΔK	24.7	22.0	17.4	15.6	15.5	13.8
ΔH	9.0	6.6	2.6	1.0	0.8	−0.5
ΔC	14.7	11.5	6.0	3.8	3.6	1.6
gvt debt	−223.3	−142.9	−12.3	36.8	40.2	84.5
wealth gini	69.3	75.7	78.1	78.5	78.6	78.9
policy[c]						
τ_1^k	1849.8	1633.8	807.9	403.8	377.4	−29.1
τ_∞^h	12.9	16.6	22.8	25.1	25.3	27.4

Notes: Optimization with respect to household with average productivity and initial wealth indicated by column head. [a)] Percent of annual consumption, the 'wealth poor' household has the lowest wealth/earnings ratio, the 'wealth rich' household has the highest wealth/earnings ratio.
[b)] Δ indicates percent change from initial to new steady state, debt is in percent of output.
[c)] Tax rates in percent.

Table 5: Optimal tax reform, restrictions on the capital income tax

	$0 \le \tau_t^k \le \tau_{ss}^k$			$\tau_t^k \equiv 0$
	policy optimal for household with $a_0 =$			
	median voter	repr. household	wealthy	repr. household
$a_0 =$	$0.375A_0$	$1.000A_0$	$1.300A_0$	$1.000A_0$
column #	1	2	3	4
welfare gain[a]				
repr. household	0.2	0.2	−0.2	−0.2
median voter	0.1	0.0	−1.6	−1.6
utilitarian	0.2	0.2	−0.2	−0.2
wealth poor	−0.2	−1.0	−9.1	−9.1
wealth rich	0.6	1.6	10.2	10.2
% gaining	100.0 [d)]	53.0	30.9	30.9
new equilibrium[b]				
ΔK	14.7	15.2	13.7	13.7
ΔH	0.3	0.6	−0.7	−0.7
ΔC	2.7	3.2	1.4	1.4
gvt debt	61.0	50.0	89.2	89.2
wealth gini	76.5	78.1	78.6	78.6
policy[c]				
τ_1^k	31.1	31.1	0.0	0.0
$\tau_?^k \approx 0$	51.0	29.0	0.0	0.0
τ_∞^h	26.3	25.8	27.7	27.7

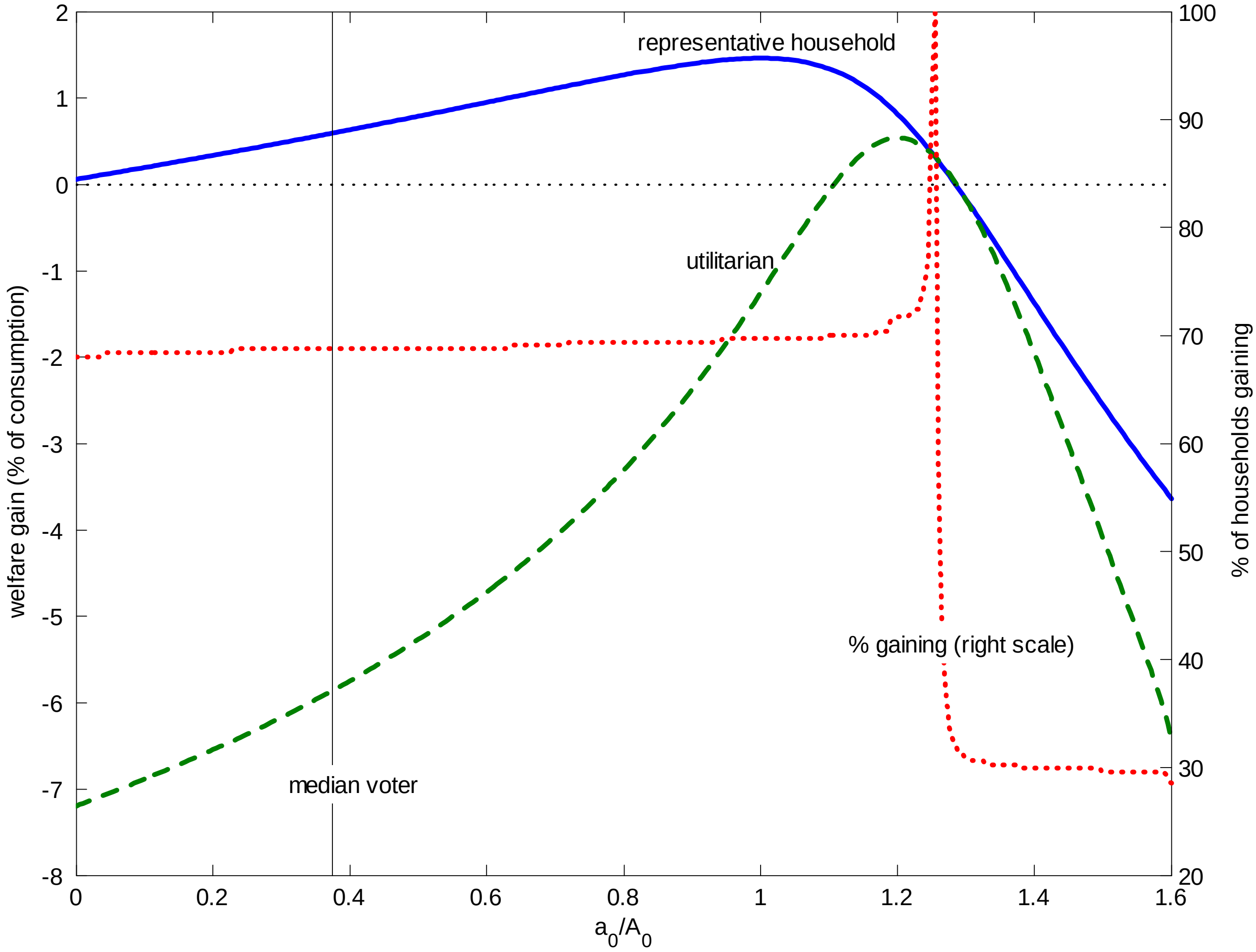
Notes: Optimization with respect to household with average productivity and initial wealth indicated by column head. [a)], [b)], see Table 4. [c)] Tax rates in percent, $\tau_?^k$ is the first period t where $\tau_t^k \approx 0$. [d)] All except the wealth poor household (with mass 3×10^{-5}) benefit.

Table 6: Sensitivity to parameterization, policy optimal 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 voter	2.6	10.3	8.5	4.4	10.5	−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	35.4	60.7	19.5	119.7	−8.3	−9.5	−24.4	−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	36.1	34.2	12.8	54.6	14.4	9.9	20.7	8.4	30.6
ΔH	0.6	18.9	9.2	3.8	13.1	−0.1	−3.9	−1.9	−0.2	−4.5
ΔC	3.1	28.5	17.9	6.1	30.0	2.3	−3.1	2.5	0.6	4.5
gvt debt	−39.4	−165.2	−89.9	−154.8	−100.3	78.0	111.2	100.2	77.3	132.8
wealth gini	78.4	71.7	75.1	76.5	73.2	79.1	76.9	77.6	79.2	-
policy[c]										
τ_1^k	943.4	1714.4	886.9	2572.3	737.3	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 Tables 4 and 5.

Figure 1: Welfare effects of optimal policies



Note: The figure shows the welfare implications of a policy that is optimal for a household with initial state $s = (z = 1, a_0)$.