Optimal Government Policies in Models with Heterogeneous Agents

Radim Boháček and Michal Kejak*
CERGE-EI, Prague, Czech Republic

August 13, 2004

Abstract

In this paper we develop a new methodology for finding optimal steady state government policies in economies with heterogeneous agents. The methodology is solely based on three classes of equilibrium conditions from government's and individual agent's optimization problems: 1) the first order conditions; 2) the stationarity condition on the distribution function; and, 3) the aggregate market clearing conditions. These conditions form a system of functional equations which we solve numerically. The solution takes into account simultaneously the effect of the government policy on individual allocations and (from the government's point of view) optimal distribution of agents in the steady state. This general methodology is applicable to a wide range of optimal government policies in models with heterogeneous agents. We illustrate it on a steady state Ramsey problem with heterogeneous agents, finding the optimal tax schedule.

JEL Keywords: Optimal macroeconomic policy, optimal taxation, computational techniques, heterogeneous agents, distribution of wealth and income

---

*Contact: CERGE-EI, Politickych veznu 7, 111 21 Prague 1, Czech Republic. Email: radim.bohacek@cerge-ei.cz, michal.kejak@cerge-ei.cz. First version: September 2002. For helpful comments we thank Juan-Carlos Conesa, Jim Costain, Max Gillman, Marek Kapička, Josep Pijoan-Mas, Galyna Vereshchagina, and the participants at the CNB/CERGE-EI Macro Workshop 2004, SED 2004 conference, and the Macroeconomic Seminar at the Federal Reserve Bank of Minneapolis. We are especially grateful to Michele Boldrin and Tim Kehoe for their support and advice. Anton Tyutin and Jozef Zubricky provided excellent research assistance. All errors are our own.

1 Introduction

This paper provides a new methodology for computing equilibria which allows the stationary distribution of wealth and income to be a part—perhaps the most important—of a government optimization problem. This general solution method is applicable to a wide range of optimal government policies in models with heterogeneous agents. We formulate the optimal government policy problem as an "operator" problem subject to a system of constraints: 1) the first order conditions from the individual agent's problem; 2) the stationarity condition on the distribution function; and, 3) the aggregate market clearing conditions. The first order conditions of the government operator problem forms a system of functional equations in individual agents' and government's policies and in the distribution function over agents' individual state variables. We solve this system numerically by the projection method.

It should be emphasized that our approach does not use any additional restrictions or assumptions on the equilibrium allocations but is strictly derived from the first order and envelope conditions and from the stationarity of the endogenous distribution in the steady state. Our main contribution is in the formulation of the functional equations for the government problem and for the stationary distribution over individual state variables. In this way, we are able to solve simultaneously for the government optimal policy, for the optimal individual allocations, and for the (from a government's point of view) optimal distribution of agents in the steady state. To our knowledge, this paper is the first one that provides a solution method for this kind of problem.

We illustrate this methodology on a steady state Ramsey problem with heterogeneous agents. We recast the original Ramsey (1927) and Lucas (1990) normative question for an economy with heterogeneous agents: What choice of a tax schedule will lead to maximal social welfare in the steady state, consistent with given government consumption and with market determination of quantities and prices? What is the welfare differential with respect to social welfare resulting from the existing progressive tax schedule in the U.S. economy and as well as from the usual flat-tax reform?

For a Ramsey problem with distortionary taxation of the total income from labor and capital, we find a welfare maximizing tax schedule which takes into account simultaneously its effects on agents' allocations and on the stationary distribution of agents in a steady state. Previous models analyzing the effects of government policies in this class of models were limited to a sub-optimal policy reform exogenously imposed on the model. Within the context of optimal taxation, several papers have analyzed the steady state implications (and transition paths) resulting from an *ad hoc* flat-tax reform or from an *ad hoc* removal of double taxation of capital income. In this paper, we solve for the *optimal* steady state tax schedule chosen by a welfare-maximizing government constrained not to use the first-best, lump sum taxes.

Our approach is very different from the current contributions to the optimal taxation literature. Based on the original Mirrlees (1971) and Mirrlees (1976) work, Kocherlakota (2003), Golosov, Kocherlakota, and Tsyvinski (2003) or Albanesi and Sleet (2003) study optimal social planner policies with asymmetric information. These incentive constraints are necessary for characterization of optimal policies. Compared to these papers, we solve for the optimal tax schedule within the standard neoclassical, general equilibrium, full information and full commitment economy with heterogeneous agents and incomplete markets. Moreover, we characterize the set of admissible tax functions that satisfy the definition of stationary recursive competitive equilibria and are consistent with each agent's maximization problem.

We limit our analysis to the optimal tax schedule on the total income from labor and capital that is needed to raise a given fraction of GDP. There are two reasons why we choose this setup. First, the tax on the total income enables us to study a distortionary tax system with a non-degenerate distribution of agents in a steady state. If the government had an access to a lump-sum, first best taxation the model would collapse to a representative agent one. The second reason for a simple tax on the total income is the complexity of the problem we solve. Our government imposes the same tax on the same total current income regardless whether it is composed of high labor income and low capital income and vice versa.

By focusing only on a steady state analysis and by imposing a single tax rate on labor and capital income we are for now avoiding two important issues related to optimal taxation: the issue of time-consistency and the issue of the optimal capital income tax rate in the steady state. With respect to the former, our government is fully and credibly committed, the tax schedule is constant over time.[1] The latter issue is to some extent mitigated by the findings of Aiyagari (1995) who showed that for our class of models with incomplete insurance markets and borrowing constraints, the optimal tax rate on

[1] For the time-consistency problem see Kydland and Prescott (1977) and a recent contribution by Klein and Rios-Rull (2004).

capital income is positive even in the long run.[2]

In order to evaluate the benefits of the optimal tax schedule, we compare the steady state aggregate levels, welfare, efficiency and distribution of resources associated with the optimal tax schedule to a simulated steady state of the U.S. economy with the existing progressive tax schedule and to a steady state resulting from a standard flat-tax reform.

The optimal tax schedule we find is not monotone. It is a positive, U-shaped function, taxing the lowest income at 45%, decreasing to a minimum of 19% and rising to 62% at the highest level of total income. It provides incentives for agents to accumulate high level of individual assets while preserving the equality measures in the economy. Its impact on aggregate levels and welfare is large. Compared to the progressive tax schedule steady state, capital stock increases by 49%, output by 15.8%, consumption by 5.8% and welfare by 4.8% (relative to the flat-tax steady state, capital stock increases by 15%, output by 4.5%, consumption by 1.1%, and welfare by 0.8%).

The distributional effects of the optimal tax are the main mechanism behind these large changes. The optimal tax schedule concentrates the agents around the mean at high levels of wealth, something what a social planner with an access to lump-sum transfers would do. The high tax rate at low income levels provides incentives for these agents to save more and move to higher income levels. The even higher tax rate on high income discourages further savings by the wealthiest agents. In the middle of the total income levels, the tax rate is lower than that found for the flat-tax reform. In this way, the optimal tax schedule solves the tradeoff between efficiency and equality by altering individual real incomes. For comparison, the flat tax also increases aggregate levels but does not take into account the distribution of agents. On the other hand, the progressive tax schedule provides too much short-run insurance at the cost of the long-run average levels.

Finally, in order to evaluate the short run costs of the optimal tax reform, we compute transitions from the progressive and the flat-rate tax schedule steady states to the steady state of the optimal tax schedule. We discuss these results and other distributional results in a great detail later.

We are aware of the other issues important for the analysis of optimal taxation. We do not consider technology or population growth, we do not model a life-cycle

[2]The optimal capital tax in an economy with two types of agents is discussed in Chamley (1986) and Judd (1985).

earnings process like Ventura (1999). The tax revenues are not given back to the agents, we abstract from public goods. We study the simplest utility maximization problem on the consumption-investment margin. However, our methodology can be applied to any optimal government policy problem, including fiscal environments with endogenous labor supply or separate taxation of labor and capital income. Finally, in the future research we also plan to analyze a much more difficult problem, that of a stationary competitive equilibrium which is the limit of the optimal dynamic tax schedule.

The paper is organized as follows. The following section describes the economy with heterogeneous agents, defines the stationary recursive competitive equilibrium and the Ramsey problem. Section 3 formulates the equilibrium as a system of functional equations and defines the operator Ramsey problem. Section 4 characterizes the optimal tax schedule by the first order necessary conditions for the operator Ramsey problem. The computational projection method is described in Section 5. Section 6 presents the results and Section 7 concludes. Appendix contains the proof of the main proposition.

2 The Steady State Ramsey Problem

This Section describes the economy, defines the stationary recursive competitive equilibrium and formulates the Steady State Ramsey Problem. The economy is populated by a continuum of infinitely lived agents on a unit interval. Each agent has preferences over consumption given by a utility function

$$E\left[\sum_{t=0}^{\infty} \beta^t u(c_t)\right],$$

where $\beta \in (0,1)$ and $u : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously differentiable, strictly increasing and strictly concave function. We assume that the utility function satisfies the Inada conditions.

In each period, each agent receives an idiosyncratic labor productivity shock which takes on values in a finite set of real numbers, $z \in Z = \{z_1, z_2, \ldots, z_J\}$. The shock is measured in efficiency units and follows a first-order Markov chain with a transition function $Q(z, z') = \text{Prob}(z_{t+1} = z' | z_t = z)$. We assume that Q is monotone, satisfies the Feller property and the mixing condition defined in Stokey, Lucas, and Prescott (1989). As the labor productivity shock is independent across agents there is no uncertainty at the aggregate level.

All agents are initially endowed with a nonnegative stock of capital. In each period, each agent supplies his realized labor endowment and accumulated capital stock k to competitive firms operating a constant returns to scale production technologies. We restrict the accumulated capital to be nonnegative, $k \in B = [k_L, \infty)$, where $k_L = 0$. The capital stock depreciates at a rate $\delta \in (0, 1)$.

Finally, there is a government that finances its expenditures by taxing the agents in the economy. We assume that the government expenditures are a fixed percentage of the total output, they are not returned to the agents and the government cannot use the first best, lump-sum taxation. We assume that the government has access only to a proportional taxation of the total income from labor and capital.

The ultimate goal is to analyze the economy as a steady-state Ramsey problem with heterogeneous agents. In particular, we seek an optimal, time invariant tax schedule in a stationary recursive competitive equilibrium that will lead to maximal average utility in the steady state, consistent with given government consumption and with market determination of quantities and prices. In the whole paper, the value function, the policy functions, the distribution function, aggregate levels and prices all depend on the tax schedule.

In the competitive equilibrium, given any time-invariant tax schedule τ, agents accumulate an aggregate capital stock K used by a representative firm together with inelastically supplied aggregate effective labor L in a production technology $F(K, L) = AK^\alpha L^{1-\alpha}$, with technology parameters $A > 0$ and $\alpha \in (0, 1)$. Profit maximization implies the following factor prices

$$r = F_K(K, L) - \delta \quad \text{and} \quad w = F_L(K, L). \tag{1}$$

In each period, an agent inelastically supplies his labor endowment at wage w, rents capital stock at interest rate r, and maximizes his or her utility by choosing consumption and a level of capital stock for the next period. We preserve the heterogeneity in the economy by closing insurance markets so that the agents can only imperfectly insure against idiosyncratic labor productivity shocks by precautionary savings.

Given competitive factor prices (r, w), the government finances its expenditures by a proportional tax on each agent's total income y from labor and capital,

$$y(k, z) = rk + wz.$$

Then the tax schedule on total income is a function

$$\tau : \mathbb{R}_+ \rightarrow \mathbb{R},$$

so that an agent with a total income y faces a tax rate $\tau(y)$ and receives an after tax income $(1-\tau(y))y$.

We will model the economy in a stationary recursive competitive equilibrium. For a time-invariant tax schedule τ, such an equilibrium exhibits constant factor prices, constant levels of aggregate variables, and a stationary distribution of agents over their individual states. An agent's individual state is a pair $(k, z) \in B \times Z$ denoting his or her accumulated stock of capital and the realized labor productivity shock, respectively. Taking the factor prices (r, w) and the tax schedule τ as given, an agent (k, z) solves the following dynamic programming problem

$$v(k, z) = \max_{c,k'} \left\{ u(c) + \beta \sum_{z'} v(k', z')\, Q(z, z') \right\}, \tag{2}$$

subject to a budget constraint

$$c + k' \leq (1 - \tau(y))y + k, \tag{3}$$

and a no borrowing constraint,

$$k' \geq 0, \tag{4}$$

where a taxable income is defined as

$$y = rk + wz.$$

For a time-invariant tax schedule τ, a probability measure λ defined on subsets of the state space describes the heterogeneity of the agents over their individual state (k, z). Let $(B, \mathcal{B})$ and $(Z, \mathcal{Z})$ be measurable spaces, where $\mathcal{B}$ denotes the Borel sets that are subsets of B and $\mathcal{Z}$ is the set of all subsets of Z. Let $(B \times Z, \mathcal{B} \times \mathcal{Z}, \lambda)$ be a probability space. We interpret λ as a probability measure describing the fractions of agents with the same individual state.

The policy function for next-period capital $k'(k, z)$ and the Markov process for the productivity shock generate a law of motion

$$\lambda'(B', z') = \sum_{z} \int_{\{(k,z) \in B \times Z :\, k'(k,z) \in B'\}} Q(z, z')\, \lambda(k, z)\, dk, \tag{5}$$

for all $(B', z') \in \mathcal{B} \times \mathcal{Z}$. According to this law of motion, the fraction of agents that will begin next period with capital stock in the set B' and a productivity shock z' is given by all those agents that transit from their current shock z to a shock z' and whose optimal decision for capital accumulation belongs to B'.

The government budget constraint is, for government expenditures G expressed as a fixed fraction g of the total output Y,

$$g \leq \frac{G}{Y} = \frac{\sum_z \int \tau(y(k,z))y(k,z)\,\lambda(k,z)\,dk}{F(K,L)}. \tag{6}$$

Definition 1 (Stationary Recursive Competitive Equilibrium) *Given a time-invariant tax schedule τ, a stationary recursive competitive equilibrium is a value function v, policy functions (c, k'), a probability measure λ, and prices (r, w), such that*

1. *given prices and the tax schedule, the policy functions solve each agent's optimization problem (2)-(4),*
2. *firms maximize profit (1),*
3. *the probability measure (5) is time invariant,*
4. *the government budget constraint (6) holds at equality,*
5. *the capital and labor markets clear,*

$$K = \sum_z \int k'(k,z)\,\lambda(k,z)\,dk, \tag{7}$$

$$L = \sum_z \int z\,\lambda(k,z)\,dk, \tag{8}$$

6. *and the allocations are feasible,*

$$\sum_z \int \left[c(k,z) + k'(k,z)\right]\,\lambda(k,z)\,dk + G = F(K,L) + (1-\delta)K. \tag{9}$$

Note that the aggregate feasibility constraint is implied from the other market clearing conditions by the Walras' law. Since labor supply is inelastic, the labor market clears by construction. We want to emphasize here that our methodology is applicable to a case in which the government consumption G were a given number.

2.1 The Existence of Stationary Recursive Competitive Equilibrium

Given a tax schedule τ and prices, the optimal policy function $k'(k,z)$ for all $(k,z) \in B \times Z$, can be derived from the first order condition,

$$u'(c) \quad \geq \quad \beta \sum_{z'} u'(c') \left[(1 - \tau(y') - \tau'(y')y') \, r + 1 \right] Q(z,z'),$$

where $c' = (1-\tau(y'))\, y' + k' - k''$ and $y' = rk' + wz'$. Because we study a stationary recursive competitive equilibrium, the two period ahead saving function $k'' = k'(k'(k,z),z')$. Note that the agent also takes into account the effect of his current savings decisions on the marginal tax tomorrow, $\tau'(y')$.

Because the tax schedule is an arbitrary function, we must ensure that the first order approach is valid. In order to characterize the admissible tax functions and to prove the Schauder Theorem for economies with distortions, we follow the notation in Stokey, Lucas, and Prescott (1989), Chapter 18. For each agent $(k,z) \in B \times Z$ with a taxable income $y(k,z) = rk + wz$, denote the after-tax gross income as

$$\psi(k,z) \equiv (1 - \tau(y(k,z)))\, y(k,z) + k. \tag{10}$$

Using $\psi(k,z)$, rewrite the Euler equation[3] as

$$u'(\psi(k,z) - k'(k,z)) = \beta \sum_{z'} u'(\psi(k'(k,z),z') - k'(k'(k,z),z'))\psi_1(k'(k,z),z')Q(z,z'),$$

where

$$\psi_1(k'(k,z),z') = (1 - \tau(y(k'(k,z),z')) - \tau'(y(k'(k,z),z'))\, y(k'(k,z),z'))\; r + 1$$

is the marginal after-tax return on an extra unit of investment. In the following theorem we establish the validity of the first order approach and the existence of the competitive equilibrium.

Theorem 1 *If for each $(k,z) \in B \times Z$, given a tax schedule $\tau : \mathbb{R}_+ \to \mathbb{R}$,*

1. $\psi_1(k,z) > 0$, *and*

[3] We consider here only the interior solution. We analyze the case of borrowing constrained agents in the following Sections.

2. ψ is quasi-concave,

then the solution to each agent's maximization problem and the stationary recursive competitive equilibrium exist.

The proof of Theorem 1 is in the Appendix.

The following corollary characterizes the set of admissible tax schedules that satisfy the conditions of Theorem 1. For that purpose, define $\overline{k}$ as the maximal sustainable capital for any agent (for a detailed definition see the proof of Theorem 1). Let $\underline{w}$ and $\overline{r}$ denote some exogenously imposed lower and upper bounds for equilibrium wage and interest rate, respectively, and let $\underline{z}$ and $\overline{z}$ stand for the lowest and highest productivity shocks. Finally, $\varepsilon_y^\tau(y) \equiv \frac{\tau'(y)}{\tau(y)} y$ is the elasticity of the tax rate to the after-tax income.

Corollary 1 (Admissible Tax Schedule Functions) *Let $C^2(\mathbb{R}_+)$ be a set of continuously differentiable functions from $\mathbb{R}_+$ to $\mathbb{R}$. If a tax schedule function $\tau \in C^2(\mathbb{R}_+)$ belongs to the set of admissible tax schedules Υ,*

$$\Upsilon = \left\{ \tau \in C^2(\mathbb{R}_+) : \tau(y) \left(1 + \varepsilon_y^\tau\right) < 1 + \frac{1}{\overline{r}} \right\} \tag{11}$$

for all $y \in [\underline{wz}, \overline{r}\overline{k} + \overline{wz}]$, then it satisfies the conditions of Theorem 1.

The above statement follows directly from the fact that $\psi_1(k, z) > 0$ and that ψ is quasi-concave. The corollary implies that there exists an upper bound on the left hand side of equation (11). The upper bound is not likely to bind for a wide range of realistic tax schedules.[4] We want to stress here that while numerically solving for the optimum tax schedule we do not impose any of these exogenous bounds but we check the admissibility of the tax schedule ex post.

The goal of our paper is to solve for the following steady state Ramsey problem.

Definition 2 (The Steady State Ramsey Problem) *A solution to the Ramsey problem for a stationary economy with heterogeneous agents is an admissible time-invariant tax schedule $\tau \in \Upsilon$ that maximizes social welfare in the steady state*

$$\max_{\tau} \sum_{z} \int v(k, z)\, \lambda(k, z)\, dk, \tag{12}$$

[4]For an equilibrium interest rate $r = 0.05$, the upper bound on the left hand side of (11) is equal to 21. In an increasing part of the tax schedule, e.g., for a tax rate of 42% the elasticity of the tax schedule at that level of income would have to be equal to an unrealistic 49 to violate the constraint.

consistent with a given government consumption and with allocations satisfying the definition of the stationary recursive competitive equilibrium, where $v : B \times Z \to \mathbb{R}$ is the value function of individual agents and $\lambda : B \times Z \to [0,1]$ is the stationary distribution.

It is easy to show that the above specification of the Steady State Ramsey Problem is equivalent to maximizing the average current period utility,

$$\max_{\tau} \sum_{z} \int u(c(k,z))\, \lambda(k,z)\, dk.$$

In the following Sections, we will characterize the optimal tax schedule using this latter specification.

3 The Operator Steady State Ramsey Problem

In this Section we study the Steady State Ramsey Problem of choosing the tax schedule to maximize the utilitarian social welfare subject to the constraints that the government's budget is balanced and that the resulting allocation is a stationary competitive equilibrium. Since the problem is to find an admissible time-invariant function $\tau : \mathbb{R}_+ \to \mathbb{R}$ describing the optimal tax schedule, we transform the Steady State Ramsey Problem into an operator problem characterizing the stationary competitive equilibrium.

First, we specify the operator equation for the Euler equation from the individual household optimization problem (2)-(4). For all $(k,z) \in B \times Z$,

$$u'(c) \;\; \geq \;\; \beta \sum_{z'} u'(c')\, [(1 - \tau(y') - \tau'(y')y')\, r + 1]\, Q(z,z'), \tag{13}$$

where $c' = (1 - \tau(y'))\, y' + k' - k''$ and $y' = rk' + wz'$. Again, the solution of the Euler equation is a time invariant policy function for the next period capital, $k'(k,z)$, and $k'' = k'(k'(k,z),z')$.

In principle, there exist two types of agents characterized by $(k,z) \in B \times Z$: agents who are unconstrained in their savings decision (i.e. $k'(k,z) > k_L$ and the Euler equation (13) holds with equality) and agents who are borrowing constrained (i.e. $k'(k,z) = k_L$ and equation (13) is satisfied with inequality). An example with only two shock levels, $Z = \{\underline{z}, \overline{z}\}$, is depicted in Figure 1. In this figure, agents with the low shock and accumulated assets $k \in [k_L, \underline{k}(\underline{z})]$ are borrowing constrained.

In general, for all $z \in Z$ there exists a minimal asset level $\underline{k}(z)$ above which agents are not borrowing constrained,

$$\underline{k}(z) \equiv \min_{x \in [k_L, \infty)} \{k'(x, z) \geq k_L\}.$$

If for a given $z \in Z$ the minimal asset level $\underline{k}(z) > k_L$, then agents with accumulated assets in $k \in [k_L, \underline{k}(z)]$ are borrowing constrained. If on the other hand $\underline{k}(z) = k_L$, all agents for that shock level are unconstrained in their saving decision.

In order to express the Euler equation in the operator form, we introduce an operator $\mathcal{F}$ defined on two functions: the tax schedule $\tau : \mathbb{R}_+ \to \mathbb{R}$ and the next period capital policy function $k' : B \times Z \to B$. The operator $\mathcal{F}$ is a mapping from a two-dimensional space of continuous functions into a one-dimensional space of continuous functions. First, for the agents with unconstrained savings decision, i.e. those with $(k, z, z') \in [\underline{k}(z), \infty) \times Z \times Z$, the operator is given by

$$\mathcal{F}(\tau, k') \equiv u'(c[k'(k, z), z; \tau]) - \beta u'(c[k'(k'(k, z), z'), z'; \tau]) R[k'(k, z), z'; \tau]\, Q(z, z'), \quad (14)$$

where

$$\begin{aligned}
c[k'(k, z), z; \tau] &= (1 - \tau(y(k, z)))\, y(k, z) + k - k'(k, z), \\
c[k'(k'(k, z)), z'; \tau] &= (1 - \tau(y(k'(k, z), z')))\, y(k'(k, z), z') + k'(k, z) - k'(k'(k, z), z'), \\
y(k'(k, z), z') &= rk'(k, z) + wz', \\
R[k'(k, z), z'; \tau] &= (1 - \tau(y(k'(k, z), z')) - \tau'(y(k'(k, z), z'))\, y(k'(k, z), z'))\ r + 1.
\end{aligned}$$

In the stationary recursive competitive equilibrium the operator Euler equation must be equal to zero, $\sum_{z'} \mathcal{F}(\tau, k') = 0$.

Second, for the agents who are borrowing constrained, i.e. for those with $(k, z, z') \in [k_L, \underline{k}(z)] \times Z \times Z$ and the next period savings k' equal to k_L, the operator $\mathcal{F}$ is equal to $k' - k_L$, which is zero.

Next we turn to the operator equation for the stationary distribution, $\mathcal{L}$, defined on three functions: the tax schedule $\tau : \mathbb{R}_+ \to \mathbb{R}$, the next period capital policy function $k' : B \times Z \to B$, and the probability measure function $\lambda : B \times Z \to [0, 1]$. In the stationary recursive competitive equilibrium, $\sum_z \mathcal{L}(\tau, k', \lambda) = 0$.

The operator $\mathcal{L}$ is a mapping from a three-dimensional space of continuous functions into a one-dimensional space of continuous functions. Because the stationary distribution is derived from agents' savings decisions, we have to again distinguish between the constrained and unconstrained agents. First, for the agents with unconstrained savings decision the operator $\mathcal{L}$ for equation (5) is given by

$$\mathcal{L}(\tau, k', \lambda) \equiv \lambda(x, z') - \lambda((k')^{-1}(x, z), z)\, Q(z, z'), \tag{15}$$

for all $(x, z, z') \in B \times Z \times Z$ and $x \geq k'(\underline{k}(z), z)$. As we show bellow, the savings function is a monotone function in k over the whole interval $[\underline{k}(z), \infty]$ and thus for any z, there exists an inverse function k^{-1} assigning the current value of capital k to the value of the next period capital x according to $k = (k^{-1})(x, z)$.

Second, for the agents borrowing constrained with the next period capital equal to k_L, the operator $\mathcal{L}$ has to be defined as

$$\mathcal{L}(\tau, k', \lambda) \equiv \lambda(k_L, z') - \int_{k_L}^{\underline{k}(z)} \lambda(k, z)\, Q(z, z')\, dk, \tag{16}$$

where $\lambda(k_L, z')$ is the mass of agents with the next period capital k_L.

Definition 3 (The Operator Stationary Recursive Competitive Equilibrium) *Given an admissible time-invariant tax schedule $\tau \in \Upsilon$, an operator stationary recursive competitive equilibrium is prices (r, w), a policy function $k' : B \times Z \to B$, a probability measure $\lambda : B \times Z \to [0, 1]$, and operators $(\mathcal{F}, \mathcal{L})$, such that*

1. *given prices and the tax schedule, the policy functions solve each agent's optimization problem*

$$\sum_{z'} \mathcal{F}(\tau, k') = 0, \tag{17}$$

2. *firms maximize profit (1),*

3. *the probability measure is time invariant*

$$\sum_{z} \mathcal{L}(\tau, k', \lambda) = 0, \tag{18}$$

4. *the government budget constraint (6) holds at equality,*

5. *the capital and labor markets clear, (7)-(8),*

6. *and the allocations are feasible, (9).*

Now we can specify an operator version of the Steady State Ramsey Problem.

Definition 4 (The Operator Steady State Ramsey Problem) *A solution to the Operator Steady State Ramsey Problem for a stationary economy with heterogeneous agents is an admissible time-invariant tax schedule $\tau \in \Upsilon$ that maximizes average steady state social welfare,*

$$\arg\max_{\tau} \sum_{z} \int \mathcal{W}(\tau, k', \lambda)\, dk \equiv \arg\max_{\tau} \sum_{z} \int u(c[k'(k,z), z; \tau])\, \lambda(k,z)\, dk, \tag{19}$$

subject to a system of operator equations (17)-(18), consistent with equilibrium prices (1) and the market clearing conditions (7)-(8) in Definition 1.

4 Characterization of the Optimal Tax Schedule

Before we derive the first order conditions for the Operator Steady State Ramsey problem we need to be more clear about what we understand by 'derivatives' of the operators with respect to the unknowns functions, in our case with respect to the tax schedule, the next-period capital policy, and the distribution function. We will define these derivatives[5]as in the calculus of variations in the following way. Assume that the triple of the tax, the saving policy and the distribution function (τ, k', λ) describe the optimum solution of the Operator Steady State Ramsey Problem. Then define 'perturbation' functions

$$\tilde{\tau}(y) \equiv \tau(y) + \varepsilon h_{\tau}(y),$$

$$\tilde{k}'(k,z) \equiv k'(k,z) + \varepsilon h_{k'}(k,z),$$

$$\tilde{\lambda}(k,z) \equiv \lambda(k,z) + \varepsilon h_{\lambda}(k,z),$$

where $\varepsilon \in \mathbb{R}$ and functions $(h_{\tau}, h_{k'}, h_{\lambda})$ are arbitrary continuously differentiable functions. The derivative of an operator $\mathcal{X} \in (\mathcal{W}, \mathcal{F}, \mathcal{L})$ with respect to function $\varphi \in (\widetilde{\tau}, \widetilde{k}', \widetilde{\lambda})$ is defined as

$$\frac{\partial \mathcal{X}}{\partial \varphi} \equiv \left(\frac{\partial \mathcal{X}}{\partial \varepsilon} \right)_{\varepsilon = 0}.$$

[5]In principle, these derivatives, sometimes defined as the Fréchet derivatives, are very close to a notion of the variation in the calculus of variations (see Kamien and Schwartz (1991), Part I).

In a similar way, we define so called 'sensitivity' functions, $\mathcal{D}k'$ and $\mathcal{D}\lambda$, which capture the effect of marginal changes in the tax function on the policy function and on the distribution function, respectively.

The first order necessary conditions for an interior solution to the Operator Steady State Ramsey Problem in Definition 4 are stated in the following Proposition.

Proposition 1 (The FOC for the Operator Steady State Ramsey Problem) *The first order necessary conditions for the Operator Steady State Ramsey Problem form the following system of operator equations:*

$$\sum_z \int \left(\frac{\partial \mathcal{W}(\tau, k', \lambda)}{\partial \tau} + \frac{\partial \mathcal{W}(k'(\tau, k', \lambda)}{\partial k'} \mathcal{D}k' + \frac{\partial \mathcal{W}(k'(\tau, k', \lambda)}{\partial \lambda} \mathcal{D}\lambda \right) dk = 0, \tag{20}$$

$$\sum_{z'} \mathcal{F}(\tau, k') = 0, \tag{21}$$

$$\sum_z \mathcal{L}(\tau, k', \lambda) = 0 \tag{22}$$

consistent with equilibrium prices (1), balanced budget, and the market clearing conditions (7)-(8) in Definition 1. The above first order necessary conditions (20)-(22), together with two additional operator equations

$$\sum_{z'} \left(\frac{\partial \mathcal{F}(\tau, k')}{\partial \tau} + \left[\frac{\partial \mathcal{F}(\tau, k')}{\partial k'} + \frac{\partial \mathcal{F}(\tau, k')}{\partial (k'(k'))} \frac{\partial k'(k')}{\partial k'} \right] \mathcal{D}k' + \frac{\partial \mathcal{F}(\tau, k')}{\partial (k'(k'))} \mathcal{D}k'(k') \right) = 0, \tag{23}$$

and

$$\sum_z \left(\frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial \tau} + \frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial k'} \mathcal{D}k' + \frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial \lambda} \mathcal{D}\lambda \right) = 0, \tag{24}$$

form a system of operator equations in the unknown time-invariant functions $k', \lambda, \mathcal{D}k', \mathcal{D}\lambda,$ *and* τ.

The proof of Proposition 1 is in the Appendix.

Intuitively, the first order condition for the Operator Steady State Ramsey Problem in equation (20) resembles a total derivative of $\mathcal{W}$ with respect to τ equal to zero, i.e., a first order condition for an unconstrained optimization problem. This comes from the fact, formally stated in the proof of Proposition 1 and discussed already in Theorem 1, that the stationary recursive competitive equilibrium is properly defined for any admissible tax function from the set Υ.[6] Therefore, the remaining two first order conditions (21)-(22) are, as we proved in the previous section, the necessary as well as sufficient conditions. Finally, the equations (23)-(24) are not the first order conditions but serve rather as additional conditions on the unknown sensitivity functions.

The exact formulas for terms $\partial\mathcal{X}/\partial\varphi$, for $\mathcal{X} \in \{\mathcal{W}, \mathcal{F}, \mathcal{L}\}$ and $\varphi \in \{k', \lambda, \tau\}$ are derived in Lemma 1 in the Appendix. Here we will discuss the results in a simplified notation only.

4.1 Effects of τ on Social Welfare

Equation (20) describes the effect of the income tax schedule on social welfare. There is a direct effect of the income tax schedule given by the first term and two indirect effects via the next period capital decision and the distribution of capital. Using (40) and (47) in Lemma 1 in the Appendix, the direct effect of tax on social welfare can be rewritten as

$$\frac{\partial\mathcal{W}}{\partial\tau} = u'(c)\left((1-\tau)\frac{\partial y}{\partial\tau} - y - \frac{\partial\tau}{\partial y}y\frac{\partial y}{\partial\tau}\right)\lambda.$$

The direct cumulative (integrated over the distribution of capital and shocks) effect of the tax change comes via the direct change in current consumption weighted by the marginal utility. The direct change in current consumption can be decomposed in two parts: the first captures a decline in disposable income proportional to the pre-tax income y; the second describes the effect of a tax change via the pre-tax income $(1-\tau)\frac{\partial y}{\partial\tau}$ (mainly through returns to capital and labor—see equations (47) and (47) in the Appendix). Interestingly, there is also an additional effect, which we call a *feedback effect* that must be taken into account when we consider a nonlinear income tax schedule contingent on the current pre-tax income: the tax change which influences income y also influences the tax rate via the change in income, thus the effect is $\frac{\partial\tau}{\partial y}y\frac{\partial y}{\partial\tau}$. The unknown

[6] In our numerical simulations we have also searched for a tax schedule that maximizes social welfare without using the first order condition approach. However, we were not able to find a tax policy for which social welfare would be superior to the solution found by using the Operator Steady State Ramsey Problem. We were exploring even the space of functions outside of the set of admissible tax functions. It is necessary to say here that for some tax functions from outside of the admissible set Υ the conditions of the stationary recursive competitive equilibrium were not satisfied. This verification can be considered as a numerical proof that the first order conditions in the Operator Steady State Ramsey Problem are necessary and sufficient, at least for this calibration of the economy.

function τ is determined from conditions (20)-(24) in Proposition 1. Knowing τ, we can obtain $\frac{\partial \tau}{\partial y}$. According to Lemma 1 in the Appendix, $\frac{\partial y}{\partial \tau} = \frac{\partial r}{\partial \tau} k + \frac{\partial w}{\partial \tau} z$, $\frac{\partial r}{\partial \tau} = F_{KK}(K) \frac{\partial K}{\partial \tau}$, $\frac{\partial w}{\partial \tau} = F_{KL}(K) \frac{\partial K}{\partial \tau}$, where K is the aggregate capital and $\frac{\partial K}{\partial \tau} = \sum_z \int k \mathcal{D} \lambda dk$.

The indirect effect of the income tax schedule on social welfare via the next period capital decision is from equation (38) equal to

$$\frac{\partial \mathcal{W}}{\partial k'} \mathcal{D} k' = -u'(c) \lambda \mathcal{D} k'.$$

The negative sign in the formula comes from $\frac{\partial c}{\partial k'} = -1$. The effect of changes in the tax schedule on the next-period capital decision is captured by the unknown sensitivity function $\mathcal{D}k'$. However, this function is characterized by the implicit functional equation (23).

The second indirect effect of the income tax schedule on the social welfare is via the distribution of capital. Using (39), it is simply equal to

$$\frac{\partial \mathcal{W}}{\partial \lambda} \mathcal{D} \lambda = u(c) \mathcal{D} \lambda.$$

The unknown function $\mathcal{D}\lambda$ contains the effects of the tax schedule on the distribution of capital. We characterize this function by the implicit functional equation (22).

4.2 Effects of τ on the Euler Equation

The total effect of the tax schedule on an individual agent's Euler equation (23), which must equal to zero by definition, can be decomposed into three effects: a *direct effect*, an *indirect effect via the next period capital*, and an *indirect effect via the two-period ahead capital*. By Lemma 1, the *direct effect* can be expressed as

$$\begin{aligned}
\frac{\partial \mathcal{F}}{\partial \tau} &= u''(c) \left((1-\tau) \frac{\partial y}{\partial \tau} - y - \frac{\partial \tau}{\partial y} \frac{\partial y}{\partial \tau} y \right) \\
&- \beta u''(c') \left((1-\tau') \frac{\partial y'}{\partial \tau} - y' - \frac{\partial \tau'}{\partial y'} \frac{\partial y'}{\partial \tau} y' \right) \left[(1 - \tau' - \frac{\partial \tau'}{\partial y'} y') r + 1 \right] Q(z, z') \\
&- \beta u'(c') \left[\left(1 - \tau' - \frac{\partial \tau'}{\partial y'} y' \right) \frac{\partial r}{\partial \tau} - r - \left(2 \frac{\partial \tau'}{\partial y'} + \frac{\partial^2 \tau'}{\partial y'^2} y' \right) \frac{\partial y'}{\partial \tau} r \right] Q(z, z').
\end{aligned}$$

In the formula above the first term measures the direct effect of a tax change in the current period due to the change in consumption weighted by the marginal change of the marginal utility $u''(c)$. The second term captures the direct effect of a tax change in the next period due to the change in consumption weighted by the marginal change of the marginal utility of the next period consumption $u''(c')$. The third term is the direct effect on the return to capital next period weighted by the marginal utility of the next period consumption $u'(c')$. To obtain the total effect, these partial effect are summed over all possible future values of shocks z'.

We can see a similar pattern of the direct effect of a change in the tax schedule on the current consumption c, the next period consumption c', and the after tax return to capital $(1-\tau)r$. There is a decline in the disposable income proportional to the pre-tax income y, or in the return to capital r. Similarly, there is the effect of a tax change via the pre-tax income $(1-\tau)\frac{\partial y}{\partial \tau}$ and the pre-tax return to capital $(1-\tau)\frac{\partial r}{\partial \tau}$. And finally, there is the feedback effect $\frac{\partial \tau}{\partial y}\frac{\partial y}{\partial \tau}y$ and $\frac{\partial \tau}{\partial y}\frac{\partial y}{\partial \tau}r$.

An additional part of the effect on the next-period after-tax return on capital comes from the existence of the term of $-\frac{\partial \tau'}{\partial y'}y'$ in the return to capital in the individual Euler equation capturing the fact that the positive slope of the tax schedule makes an additional incentive to the individual to earn lower income next period to be taxed by the lower tax rate. This means that the two additional terms, the first order and the second-order feedback effects, $2-\frac{\partial \tau'}{\partial y'}\frac{\partial y'}{\partial \tau}-\frac{\partial^2 \tau'}{\partial y'^2}\frac{\partial y'}{\partial \tau}y'$ capture the effect of tax schedule on this disincentive to earn too high income. The first term says that the disincentive depends on how the next period income changes with the tax schedule. The second term takes into account how the slope of the tax schedule changes with the change of the tax schedule.

The unknown function τ can be determined from conditions (20)-(24) of Proposition 1. Once we know τ, we also know $\frac{\partial \tau}{\partial y}$ and $\frac{\partial^2 \tau}{\partial y^2}$; τ', $\frac{\partial \tau'}{\partial y'}$, and $\frac{\partial^2 \tau'}{\partial y'^2}$ are just functions τ, $\frac{\partial \tau}{\partial y}$, and $\frac{\partial^2 \tau}{\partial y^2}$ applied to the next period, i.e. $\tau'=\tau(y')$, $\frac{\partial \tau'}{\partial y'}=\frac{\partial \tau(y')}{\partial y}$, and $\frac{\partial^2 \tau'}{\partial y'^2}=\frac{\partial^2 \tau(y')}{\partial y^2}$. According to Lemma 1, $\frac{\partial y'}{\partial \tau'}\equiv\frac{\partial r}{\partial \tau}k'+\frac{\partial w}{\partial \tau}z'$.

The term on the second line of equation (23) is the *indirect effect via next period capital* on the individual Euler equation. We rewrite it in a simplified notation as

$$\left[\frac{\partial \mathcal{F}}{\partial k'}+\frac{\partial \mathcal{F}}{\partial k''}\frac{\partial k''}{\partial k'}\right]\mathcal{D}k'. \tag{25}$$

eThe function τ affects the Euler equation directly via k' (the first term) and indirectly via $k''=k'(k')$ (the second term). The expressions $\frac{\partial \mathcal{F}}{\partial k'}$, $\frac{\partial \mathcal{F}}{\partial k''}$, and $\frac{\partial k''}{\partial k'}$ are described below. The term $\mathcal{D}k'$, how the change in the tax function changes the next period capital decision, is the unknown sensitivity function.

The first, indirect effect in equation (25) is equal to

$$\begin{aligned}\frac{\partial \mathcal{F}}{\partial k'} &= -u''(c) - \beta u''(c')\left((1-\tau')r - \frac{\partial \tau'}{\partial y'} y' r\right)\left[(1-\tau' - \frac{\partial \tau'}{\partial y'} y')r + 1\right] Q(z, z') \\ &- \beta u'(c')\left(-2\frac{\partial \tau'}{\partial y'} - \frac{\partial^2 \tau'}{\partial y'^2} y'\right) r^2\, Q(z, z').\end{aligned}$$

It can be again decomposed into three parts: the effect through the current consumption $-u''(c)$, the next period consumption, and the feedback effect on the tax schedule $\left(2\frac{\partial \tau'}{\partial y'} + \frac{\partial^2 \tau'}{\partial y'^2} y'\right) r^2$ since $\frac{\partial y'}{\partial k'} = r$.

The second indirect effect in equation (25),

$$\frac{\partial \mathcal{F}}{\partial k''} = \beta u''(c')\left[(1-\tau' - \frac{\partial \tau'}{\partial y'} y')r + 1\right] Q(z, z'),$$

is the effect of the two period ahead capital decision function $k'' = k'(k')$, using the fact that $\partial c'/\partial k'' = -1$.

The second term in equation (25) also contains the effect of the policy function k' on k''. In the full notation,

$$\frac{\partial k''}{\partial k'} = \frac{\partial k'(k'(k,z), z')}{\partial k}.$$

Exploiting the time invariant structure of the model, knowing the decision function k' also allows us to construct its derivative with respect to k, $\frac{\partial k'}{\partial k}$, and thus $\frac{\partial k''}{\partial k'}$.

The last line in the individual Euler equation (23), the *indirect effect via two period ahead capital* on the individual Euler equation can be expressed as

$$\frac{\partial \mathcal{F}}{\partial k''} \frac{\partial k''}{\partial \tau}.$$

This effect captures the effect of k'' on the Euler equation and the effect of the tax schedule on k''. The first part $\frac{\partial \mathcal{F}}{\partial k''}$ was already discussed above. The direct effect of the tax schedule on the two period ahead capital decision function can again be clearly understood in the full notation,

$$\frac{\partial k''}{\partial \tau} = \mathcal{D}k'(k'(k,z), z').$$

Again, the knowledge of $\mathcal{D}k'$ is sufficient for determining $\frac{\partial k''}{\partial \tau}$.

4.3 Effects of τ on the Stationary Distribution

The total effect of the tax schedule on the stationary distribution in equation (22), which must be, by definition, equal to zero, can be decomposed into three effects: a *direct effect*, an *indirect effect via next period capital*, and an *indirect effect via the stationary distribution*.[7]

The *indirect effect via the next period capital* from the second term of equation (24) is by Lemma 1 equal to

$$\frac{\partial \mathcal{L}}{\partial k'}\mathcal{D}k' = \frac{\partial \lambda}{\partial k}\frac{\partial k}{\partial \tau}\,Q(z,z').$$

The indirect effect of the tax schedule via the next period capital associated with a shock z today is simply composed of the effect of current capital on the distribution, $\frac{\partial \lambda}{\partial k}$, and of the effect of the tax schedule on the current level of capital. In the Appendix we show that the effect of tax on the current period capital $\frac{\partial (k')^{-1}(k',z)}{\partial \tau}$, given a pair (k',z), can be expressed as $\frac{\partial (k')^{-1}(k',z)}{\partial \tau} = \mathcal{D}k'\left[\frac{\partial k'}{\partial k}\right]^{-1}$ where $\mathcal{D}k'$ is the unknown function. Note that if we know the policy function k' then $\frac{\partial k'}{\partial k}$ can be determined too. The total indirect effect via the stationary distribution is the sum of the indirect effects summed over all values of shocks z.

Finally, the last term in equation (24) describes *the indirect effect of tax schedule on the stationary distribution*. It is equal to the difference between the indirect effect on the distribution next period and the sum of the indirect tax effects on the current distribution over all possible current shocks z,

$$\frac{\partial \mathcal{L}}{\partial \lambda}\mathcal{D}\lambda = \mathcal{D}\lambda' - \mathcal{D}\lambda\,Q(z,z').$$

Using the unknown function $\mathcal{D}\lambda$ we can determine $\mathcal{D}\lambda' = \mathcal{D}\lambda(k',z')$ and $\mathcal{D}\lambda = \mathcal{D}\lambda((k')^{-1}(k',z),z)$ by evaluating $\mathcal{D}\lambda$ at (k',z') and $((k')^{-1}(k',z),z)$, respectively.

5 The Least Squares Projection Method

The solution to the Operator Steady State Ramsey Problem from Proposition 1 can be found numerically by using the least squares projection method. In this Section

[7] For simplicity, we will discuss here only the case when the lower bound constraint is not binding. In this case there is no direct effect of taxes on the stationarity condition and $\frac{\partial \mathcal{L}}{\partial \tau} = 0$. This is not the case when the lower bound is binding. For details, see Lemma 1 in the Appendix.

we outline its application to our problem and the approximation of the optimal tax schedule.[8]

The solution to the Operator Steady State Ramsey Problem are the zeros of the given operator equations. We first approximate the unknown functions by combinations of polynomials from a polynomial base. Therefore, approximated solutions are specified by unknown parameters transforming the original infinitely dimensional problem into a finite dimensional one. After substituting the approximated functions into the original operator equations we create the so called *residual equations*. Ideally, the residual functions should be uniformly equal to zero. In practical situations, however, this is not achievable and we limit the problem to a finite number of conditions, the so called *projections*, whose satisfaction guarantees a reasonably good approximation. There are many possibilities how to define the projections.[9] We have chosen here the least squares projection method for its good convergence properties. We search for parameters approximating the functional equations that minimize the squared residual functions.

In the system of operator equations given by (20)-(22) and (23)-(24), there are five unknown classes of functions $\{k', \lambda, \mathcal{D}k', \mathcal{D}\lambda, \tau\}$. Since we assume that the shocks are discrete, $z \in Z = \{z_1, z_2, \dots, z_J\}$ and $J > 1$, we define the following family of policy and distribution functions, and their derivatives $\{k'_i(k), \lambda_i(k), \mathcal{D}k'_i(k), \mathcal{D}\lambda_i(k)\}_{i=1}^{J}$, for each shock value $z_1, z_2, \dots, z_J$. We interpret the policy function k'_i as the next-period capital function of an agent who was hit by the shock level z_i. Analogously, the distribution function λ_i is the distribution of agents with the shock z_i, etc. Similarly, we assign the Euler and distribution function operators to every shock level, $\mathcal{F}_i$ and $\mathcal{L}_i$, respectively. We approximate all unknown functions by the orthogonal Chebyshev polynomial base $\{T_i(x)\}_{i=0}^{\infty}$ defined for $x \in [-1, 1]$.

As we have to define our approximation on a finite interval, we set the highest capital level to a value k_H, greater than the endogenous upper bound on the stationary distribution. Let the interval of approximation be $[k_L, k_H]$ and the degrees of approximation for $\{k'_i(k), \lambda_i(k), \mathcal{D}k'_i(k), \mathcal{D}\lambda_i(k)\}$ be $M, N, O, P, Q \geq 2$, respectively.[10]

[8]For a detailed explanation of projection methods to stationary equilibria in economies with a continuum of heterogenous agents see Bohacek and Kejak (2002). We use the least square projection method for its advantage in solving systems of nonlinear operator equations.

[9]For an excellent survey and description of these methods see Chapter 11 in Judd (1998).

[10]The details on Chebyshev polynomials can be found in Judd (1992), Judd (1998) or in any book on numerical mathematics. The linear transformation $\xi : [k_L, k_H] \to [-1, 1]$ is necessary if we want

Thus, we obtain

$$\widehat{k'}_i(k;a^i) \equiv \sum_{j=1}^{M} a_j^i \phi_j(k), \tag{26}$$

$$\widehat{\lambda}_i(k;a^{i+J}) \equiv \sum_{j=1}^{N} a_j^{i+J} \phi_j(k), \tag{27}$$

$$\widehat{\mathcal{D}k'}_i(k;a^{i+2J}) \equiv \sum_{j=1}^{O} a_j^{i+2J} \phi_j(k), \tag{28}$$

$$\widehat{\mathcal{D}\lambda'}_i(k;a^{i+3J}) \equiv \sum_{j=1}^{P} a_j^{i+3J} \phi_j(k), \tag{29}$$

$$\widehat{\tau}(k;a^{4J+1}) \equiv \sum_{j=1}^{Q} a_j^{4J+1} \phi_j(k). \tag{30}$$

for any $k \in [k_L, k_H]$ where $\phi_j(k) \equiv T_{j-1}(\xi(k))$, a's are the unknown parameters and $i = 1, \ldots, J$.

Now we have to define residual functions as approximations to the original operator functions (20)-(24). Substituting the approximated functions above for the unknown functions, we have

$$R^W(k;\mathbf{a}) = \sum_{z_j} \widehat{\int} \left(\frac{\partial \mathcal{W}(\widehat{\mathbf{k}'}, \widehat{\mathbf{\Lambda}}, \widehat{\tau})}{\partial \tau} + \frac{\partial \mathcal{W}(\widehat{\mathbf{k}'}, \widehat{\mathbf{\Lambda}}, \widehat{\tau})}{\partial k'} \widehat{\mathcal{D}k'}_j + \frac{\partial \mathcal{W}(\widehat{\mathbf{k}'}, \widehat{\mathbf{\Lambda}}, \widehat{\tau})}{\partial \lambda} \widehat{\mathcal{D}\lambda'}_j \right) dk, \tag{31}$$

$$R_i^F(k;\mathbf{a}) = \sum_{z'_j} \mathcal{F}_i(\widehat{\tau}, \widehat{\mathbf{k}'}), \tag{32}$$

$$R_i^L(k;\mathbf{a}) = \sum_{z_j} \mathcal{L}_i(\widehat{\tau}, \widehat{\mathbf{k}'}, \widehat{\mathbf{\Lambda}}), \tag{33}$$

$$R_i^{FT}(k;\mathbf{a}) = \sum_{z'_j} \left(\frac{\partial \mathcal{F}(\widehat{\tau}, \widehat{\mathbf{k}'})}{\partial \tau} + \frac{\partial \mathcal{F}(\widehat{\tau}, \widehat{\mathbf{k}'})}{\partial k'(k')} \widehat{\mathcal{D}k'}_j(\widehat{k'_i}) + \left[\frac{\partial \mathcal{F}(\widehat{\tau}, \widehat{\mathbf{k}'})}{\partial k'} + \frac{\partial \mathcal{F}(\widehat{\tau}, \widehat{\mathbf{k}'})}{\partial (k'(k'))} \frac{\partial \widehat{k'_j}(\widehat{k'_i})}{\partial k'} \right] \widehat{\mathcal{D}k'}_i \right), \tag{34}$$

$$R_i^{LT}(k;\mathbf{a}) = \sum_{z_j} \left(\frac{\partial \mathcal{L}(\widehat{\tau}, \widehat{\mathbf{k}'}, \widehat{\mathbf{\Lambda}})}{\partial \tau} + \frac{\partial \mathcal{L}(\widehat{\tau}, \widehat{\mathbf{k}'}, \widehat{\mathbf{\Lambda}})}{\partial k'} \widehat{\mathcal{D}k'}_j + \widehat{\mathcal{D}\lambda'}_i - \widehat{\mathcal{D}\lambda'}_j \, Q(z_j, z'_i) \right), \tag{35}$$

to use the Chebyshev polynomials on the proper domain. It is straightforward to show that $\xi(k) = 2(k - k_L)/(k_H - k_L) - 1$.

where the vector of parameters $\mathbf{a} \equiv (a^1, a^2, \ldots, a^J, a^{J+1}, \ldots, a^{J+2})$ is of a size $S = J \times (M + N + O + P) + Q$, $\widehat{\mathbf{k}}' \equiv (\widehat{k}'_1, \widehat{k}'_2, \ldots, \widehat{k}'_J)$, $\widehat{\mathbf{\Lambda}} \equiv (\widehat{\lambda}_1, \widehat{\lambda}_2, \ldots, \widehat{\lambda}_J)$, and $\frac{\partial \mathcal{W}(\widehat{\tau}, \widehat{\mathbf{k}}', \widehat{\mathbf{\Lambda}})}{\partial \tau} \equiv \frac{\partial \mathcal{W}(\widehat{\tau}(k;\mathbf{a}), \widehat{\mathbf{k}}'(k;\mathbf{a}), \widehat{\mathbf{\Lambda}}(k;\mathbf{a}))}{\partial \tau}$ etc. Further, $\mathcal{F}_i(\widehat{\tau}, \widehat{\mathbf{k}}') = \mathcal{F}_i(\widehat{\tau}, \widehat{k}'_1(\widehat{k}'_i), \ldots, \widehat{k'_J}(\widehat{k}'_i))$ for any $i = 1, \ldots, J$ is obtained from equation (17) where we substitute $\widehat{k}'_i(k;\mathbf{a})$ for $k'(k,z)$ and $\widehat{k}'_j(\widehat{k}'_i(k;\mathbf{a});\mathbf{a})$ for $k'(k'(k,z),z')$ as well as z_i for z and z'_j for z'. We can similarly get formulas for $\mathcal{L}$ by using equation (18) respectively.[11] To obtain the residual equation (23) we use (41)-(42) and substitute $\widehat{\mathcal{D}k'}_i(k;\mathbf{a})$ for $\widehat{\mathcal{D}k'}_i$. Similarly, we find the residual equations (31) and (35) from equations (20),(38)-(40), and (24),(44)-(46), respectively.

The least squares projection method looks for parameters $\mathbf{a}$ that minimize the sum of weighted residuals,

$$\sum_{i=1}^{J} \int_{k_L}^{k_H} \left([R_i^F(k;\mathbf{a})]^2 + [R_i^L(k;\mathbf{a})]^2 + [R_i^{FT}(k;\mathbf{a})]^2 + [R_i^{LT}(k;\mathbf{a})]^2 \right) w(k) dk + \int_{k_L}^{k_H} [R^W(k;\mathbf{a})]^2 \, w(k) \, dk, \tag{36}$$

with the weighting function given by $w(k) \equiv \left(1 - \left(2\frac{k-k_L}{k_H-k_L}\right)^2\right)^{-1/2}$ and $i = 1, \ldots, J$. After approximating the integrals by the Gauss-Chebyshev quadrature, we obtain a minimization problem

$$\min_{\mathbf{a} \in \mathbb{R}^S} \sum_{i=1}^{J} \sum_{k} \left([R_i^F(k;\mathbf{a})]^2 + [R_i^L(k;\mathbf{a})]^2 + [R_i^{FT}(k;\mathbf{a})]^2 + [R_i^{LT}(k;\mathbf{a})]^2 \right) + \sum_{k} [R^W(k;\mathbf{a})]^2,$$

with k's being the zeros of the polynomial ϕ of a degree greater than the biggest degree of approximation $\max\{M, N, O, P, Q\}$.

Since the least squares projection method sets up an optimization problem we can use standard methods of numerical optimization, e.g. the Gauss-Newton or the Levenberg-Marquardt methods. Again, the discussion of these methods is not the aim of our paper. However, we found that these traditional methods did not work in our high-dimensional problem mainly due to possible multiple local solutions. We tried several other methods (simulated annealing or genetic algorithm with quantization, for example) and finally succeeded with a genetic algorithm with multiple populations and local search.

[11] The hat above the integration $\widehat{\int}$ means that we need to approximate the integration on the interval $[k_L, k_H]$ in equation (19).

6 Results

In this Section we solve for the optimal tax schedule and compare the associated steady state allocations to those resulting from the existing progressive tax schedule in the U.S. economy and from the usual flat-tax reform.

6.1 Parameterization

Given the complexity of our steady state Ramsey problem, we cannot model the earnings process so well as Ventura (1999). In this model, each agent supplies labor inelastically and the uninsurable idiosyncratic shock to labor productivity follows a two-state, first order Markov chain. We use the results of Heaton and Lucas (1996) who, using the PSID labor market data, estimate the household annual labor income process between 1969 and 1984 by a first-order autoregression of the form

$$\log(\eta_t) = \bar{\eta} + \rho \log(\eta_{t-1}) + \epsilon_t,$$

with $\epsilon \sim N(0, \sigma_\epsilon^2)$. They find that $\rho = 0.53$ and $\sigma_\epsilon^2 = 0.063$. Tauchen and Hussey (1991) approximation procedure for a two-state Markov chain implies $z_L = 0.665$, $z_H = 1.335$ and $Q(z_L, z_L) = Q(z_H, z_H) = 0.74$. These values imply an aggregate effective labor supply equal to one with agents evenly split over the two shocks.[12] The rest of the parameters is taken from Prescott (1986), in particular $\alpha = 0.36$, $\delta = 0.1$, and the preference parameters $\sigma = 1$ and $\beta = 0.95$.

Finally, for all steady states we consider a Ramsey problem in which government is required to raise a predetermined amount of tax revenues equal to 20% of the total output, i.e. $g = 0.2$.

6.2 The U.S. Progressive Tax Schedule

We model the progressive tax schedule as Ventura (1999), the closest model analyzing a flat-tax reform in an economy with heterogeneous agents.[13] An agent's budget constraint

[12] Similar parameterization is in Storesletten, Telmer, and Yaron (1999) with $z_L = 0.73$, $z_H = 1.27$ and $Q(z_L, z_L) = Q(z_H, z_H) = 0.82$. Diaz-Jimenez, Quadrini, and Rios-Rull (1997) use $z_L = 0.5$, $z_H = 3.0$ and $Q(z_L, z_L) = 0.9811, Q(z_H, z_H) = 0.9261$.

[13] Compared to his model, our agents are infinitely lived, so we omit the life-cycle variables, accidental bequests, government transfers, and social security tax and benefits. Except for capital depreciation, we do not consider tax deductions.

can be written as

$$c + k' \leq rk + zw + k - T,$$

where T represents the amount of tax paid by the agent according to the progressive tax schedule. The amount of tax is determined according to which tax bracket the total taxable income, $I = rk + \max\{0, zw - I^*\}$, falls in, with a labor-income tax deductible amount $I^* \geq 0$. There are M brackets with associated tax rates, $\tau_m, m = 1, \ldots, M$, defined on intervals between brackets' bounds $I_0, \ldots, I_{M-1}$. For $M = 5$ the tax rates are $\tau_m \in \{0.15, 0.28, 0.31, 0.36, 0.396\}$ and tax brackets, expressed as a multiple of the average income, $I_{m-1} \in \{0, 0.85, 2.06, 3.24, 5.79\}$. In addition, capital income, rk, is taxed at flat rate $\tau_k = 0.25$.

For income $I \in (I_{m-1} - I_m]$, the total tax is then

$$T = \tau_1(I_1 - I_0) + \tau_2(I_2 - I_1) + \ldots + \tau_m(I - I_{m-1}) + \tau_k rk.$$

The government budget constraint is cleared by finding an equilibrium value of the tax exemption level I^*.

6.3 A Flat-Tax Reform

The flat-tax reform consists of replacing the progressive tax schedule with a single flat tax τ on the total income from labor and capital. The budget constraint of each agent becomes

$$c + k' \leq (1 - \tau)(rk + zw) + k.$$

Note that the flat tax reform, like in Ventura (1999), does not eliminate taxation of capital income. We find that the equilibrium flat tax rate is $\tau = 0.254$.

6.4 The Optimal Tax Schedule

Finally, we use our methodology described in the previous Sections to solve for the optimal tax schedule that maximizes average steady state welfare while raising tax revenues amounting to 20% of GDP. Figure 2 shows the numerical solution for the optimal tax schedule. It is a U-shaped function taxing the lowest total income at a 45%, decreasing to a minimum of 19% and rising to 62% at the highest level of total income.

Table 1 compares the aggregate levels of the three steady states. The left column describes the steady state values of the U.S. progressive tax schedule, the middle column those of the flat-tax reform, and the right column is for the optimal tax schedule steady state.

Our simulation of the flat-tax reform increases the steady state levels by magnitudes found in the literature: capital stock increases by 30%, output by 10.8%, consumption by 4.6%, and welfare by 3.9%. As in Ventura (1999), the flat-tax reform increases inequality: Gini income coefficients rise from 0.22 to 0.31 before tax and from 0.21 to 0.32 after tax.[14]

The impact of the optimal tax schedule is even greater. Capital stock increases by 49%, output by 15.8%, consumption by 5.8% and welfare by 4.8%. Inequality increases too but not as much as in the flat-tax reform: Gini income coefficients are 0.28 before and 0.27 after tax, respectively. General equilibrium effects cause the interest rate to drop by almost one half and the wage to increase due to a higher productivity of labor used in production with such a high capital stock. Compared to the flat-tax steady state, capital stock increases by 15%, output by 4.5%, consumption by 1.1%, and welfare by 0.8%.

The huge increases in these levels arise from the distributional effects of the optimal tax schedule, shown in Figure 3. Although both the flat and the optimal tax schedules increase the aggregate levels, the difference between them is that the flat tax schedule does not take into account the distribution of agents. The flat tax reform helps more the agents with high incomes: the mean wealth increases much more than the median so that the median/mean ratio falls to 0.86. In the flat-tax steady state the aggregate levels increase but from "the optimal distribution" point of view the mass of agents moves too much to the left while wealthy agents emerge at the right tail of the distribution. The progressive tax schedule has the lowest inequality measures because the high taxes on rich agents narrows the distribution towards the mean. However, the low tax rates on low incomes do not provide incentives for the poor households to save and move to higher income levels. In other words, it provides too much short-run insurance at the cost of the long-run average levels.

This is exactly what the optimal income tax schedule improves. The main mechanism

[14]Elimination of capital income tax in Lucas (1990) increases capital stock by 30-34% and consumption by 6.7%. A flat-rate reform with heterogeneous agents in Ventura (1999) increases the total capital stock by one third, output by 15%. Average labor hours do not change but aggregate labor in efficiency units increases because more productive agents work more. Without a well calibrated life-cycle earnings process we are not able to match well the inequality coefficients, especially that of wealth.

behind the large growth in the aggregate levels is the incentive effect of the optimal tax schedule. The U-shaped function in Figure 2 effectively concentrates the agents around the mean, something what a social planner with an access to lump-sum transfers would do. The high tax rate at low income levels provides incentives for these agents to save more and move to higher income levels. On the other hand, the even higher tax rate on high income discourages further savings by the wealthiest agents. In the middle of the total income levels, the tax rate is lower than that found for the flat-tax reform. The majority of agents is concentrated there. The optimal tax schedule preserves the median/mean wealth ratio of the progressive tax schedule by increasing the median by 42% and the mean by 49%. The support of the invariant distribution becomes wider but inequality measures do not increase as much as in the flat-tax reform.

The optimal tax function τ is strictly positive and very nonlinear. Both results are very different from Mirrlees (1971), who shows that in a static model with fixed distribution of skills, a welfare maximizing tax schedule is close to a linear, non-decreasing function. The highest marginal tax rate is for low incomes and falls thereafter, a negative income tax is supported. However, his results are sensitive to the fixed distribution of skills and consumption-leisure preferences. We want to emphasize that our stationary distribution is endogenous and there are no restrictions on the optimal tax schedule to be positive or to be of any particular shape.

The welfare gain in the optimal tax schedule steady state is large, almost 5% of the average consumption given to each agent in the progressive tax steady state in perpetuity. Table 2 shows the distribution of resources for quintiles of the wealth distribution. Because of the high tax rate on incomes in the bottom quintile, agents in the optimal tax schedule steady state consume 6.5% less than those of the progressive tax schedule. From the second quintile of the optimal tax schedule steady state, agents consume on average more than in the other two steady states. Dividing these levels by the average consumption in each steady state, we can calculate average quintile consumption relative to the steady state average. In the optimal tax schedule, the bottom quintile consumes 73% of the average consumption, in the flat tax it is 77%, in the progressive it is 82%. This shows that the incentive effects overweight the insurance (i.e., the redistribution) aspect of both the progressive and the flat tax schedules. All other agents consume much more than the average, those of the top quintile by 12%. This is similar to the flat-tax steady state but the magnitudes are higher.

The distribution of capital reveals that the incentives contained in the optimal-tax schedule move the distribution to higher capital levels. The poorest quintile owns on average 17% more assets than in the progressive steady state. This increase is even larger for the other quintiles (40% on the top). Again, the flat-rate steady state leads to lower level of savings by the bottom two quintiles. These levels are reflected in the shares of the total capital stock. For all steady states the bottom quintile owns only around 5% of the total stock while the top quintile around one third (43% in the flat-tax steady state).

The investment-to-income ratios reveal the agents in the bottom quintile of the optimal schedule invest much more than similar agents in the other two steady states. Agents in the optimal tax schedule steady state invest 30% of their income, more than those in the flat-tax (27%) and progressive (22%) steady states. The investment is also more evenly distributed over the quintiles. Note also that the flat-rate tax schedule favors capital accumulation by the top quintile.

The income and after-tax income distribution show the differences between the three tax schedules. The progressive tax helps the bottom quintile while the flat tax helps the top quintile. The U-shape of the optimal tax provides the right incentives at the cost of lowest after-tax income for the poor agents. Finally, the optimal tax actually equalizes tax contribution share of total tax revenues across the quintiles. Both the flat-tax and progressive-tax steady states put more relative burden on the higher income quintiles.

Figure 4 shows the sensitivity functions $\mathcal{D}k'$ and $\mathcal{D}\lambda$. The top panel shows the effect of a change in the optimal tax schedule on the savings decision of agents. For the low shock it is close to zero, for the high shock it is negative and monotonically decreasing. The bottom panel displays the same effects on the probability density function of the stationary distribution λ, again for each shock. We know from the stationarity condition of the distribution that the integral of these functions must be zero. Our numerical solution is only very close to zero due to approximation errors and the complexity of the problem we face.

6.5 Transition to the Optimal Tax Schedule Steady State

Pure welfare steady-state comparisons could be misleading because tax changes imply substantial redistribution in the short run. Changes in the mix between capital and labor income taxes redistribute the tax burden across households. Domeij and Heathcote

(2004) find that capital tax cuts imply large welfare gains in the representative agent economy but not in economies with heterogeneous agents. In their model, the expected discounted present value of welfare losses during transition are so large that they overturn the steady state welfare improvement. The short-run cost in the form of higher labor taxes is too heavy a price to pay for all except for the wealth-richest households.[15]

Table 3 shows the results for our tax reform experiment. It compares the expected present discounted value from an unanticipated optimal tax reform of the progressive and flat-tax steady state. In each case the optimal tax schedule is imposed on the stationary distribution of the initial steady state. We guess a sufficiently large number of convergence periods and iterate on paths of equilibrium interest rates and wages to clear markets in each period of the transition, returning possible excess tax revenues to all agents in each period. The convergence is relatively fast lasting around thirty periods.

Contrary to Domeij and Heathcote (2004), we find that both the mean and the median agents in the progressive tax schedule economy are better off. The welfare gains are smaller but positive (3.8% and 4.4%, respectively, measured as per period consumption transfers as a percentage of the initial steady state average consumption). The top panel in Figure 5 shows the expected present discounted values in the progressive-rate steady state and at the moment of the unanticipated reform to the optimal tax schedule. The reform is not Pareto improving as the poorest 27% of all households are worse off (they are hit by the high taxes the optimal schedule imposes on low income levels).

On the other hand, a transition from the flat-tax steady state would not be supported by the mean nor by the median agent (they loose 1% and 1.4%, respectively). The poor and the wealthy, for whom the tax increases dramatically, are worse off during the transition. The bottom panel in Figure 5 shows the expected present discounted values of the flat-rate steady state and of the transition to the optimal tax schedule. Political support is not sufficient, equal only to 38% of the population.

This transition exercise shows that a simple tax reform (an immediate switch to a new tax schedule) is not usually Pareto improving for all agents. However, the gains

[15]This is similar to Garcia-Mila, Marcet, and Ventura (1995) and Auerbach and Kotlikoff (1987) who find that reducing capital income taxation shifts the tax burden away from households who receive a large fraction of their income from capital and towards those who receive a disproportionate fraction from labor. Transition costs in Lucas (1990) reduce the welfare gains from zero capital tax reform to 0.75-1.25 percent of average consumption in the initial steady state.

from the optimal tax reform of the existing progressive tax schedule are so large that they are supported by majority of agents despite their transitional costs.

7 Conclusions

Quah (2003) shows that average levels are of the first order importance for economic growth and welfare, much more important than within-country inequality. Government policies focusing on aggregate levels, including obviously optimal fiscal policy and taxation, are essential. However, it is the distribution of agents that delivers these aggregate levels. This paper clearly shows that it is crucial to think of policies that target the distribution of agents. Only in this way the high aggregate levels and welfare improvements can be achieved.

To our knowledge, this paper is the first one that provides the solution method for such optimal government policies in steady state equilibria in heterogeneous agent economies. We think of these policies as optimal because they take into account their effects on the distribution of agents. As an example, we find the optimal tax schedule for a steady state Ramsey problem in an economy with heterogeneous agents. The optimal tax schedule is U-shaped, it increases all aggregate levels by providing the right incentives for the agents to accumulate high aggregate levels but not at the cost of increased inequality. Welfare gain in the steady state is large, almost 5%, and it is positive for both mean and median agent in a transition following an unanticipated optimal tax reform of the progressive tax schedule steady state.

The methodology developed in this paper can be applied to any optimal government policy. Within the field of optimal taxation, in our future research we plan to study the optimal tax schedule with elastic labor supply. An endogenous labor-leisure decision might significantly affect the shape of the optimal tax schedule, the aggregate labor supply and the distribution of labor hours. Another topic that has received a lot of attention is the optimal capital taxation in models with heterogeneous agents (see Aiyagari (1995) for the initial contribution). Finally, we plan to use this methodology to analyze optimal dynamic taxation.

Appendix

Proof of Theorem 1

We need to show that the first order approach to each agent's maximization problem is valid. First, agents maximize over a quasi-convex set: $\Psi = \{x \in B : 0 \leq x \leq \psi(k,z) \text{ for all } (k,z) \in B \times Z\}$. If the function ψ is increasing and quasi-concave, then the set Ψ is quasi-convex. Further, we need to satisfy Assumptions 18.1 in Stokey, Lucas, and Prescott (1989), particularly that (i) $\beta \in (0,1)$; (ii) utility function $u : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously differentiable, strictly increasing and strictly concave function; (iii) for some $\bar{k}(z) > 0, \psi(k,z) - k$ is strictly positive on $[0, \bar{k}(z))$ and strictly negative for $k > \bar{k}(z)$, where the value $\bar{k}$, the maximum sustainable capital stock out of after-tax income for any agent, is defined as $\bar{k} = \max\{\bar{k}(z_1), \ldots, \bar{k}(z_J)\}$; and, (iv) given the tax-schedule function the right-hand side of the Euler equation is strictly positive

$$\beta \sum_{z'} u'(\psi(k'(k,z),z') - k'(k'(k,z),z'))\psi_1(k'(k,z),z')\, Q(z,z') > 0, \tag{37}$$

where

$$\psi_1(k'(k,z),z')Q(z,z') = (1 - \tau(y(k'(k,z),z')) - \tau'(y(k'(k,z),z'))\, y(k'(k,z),z'))\; r + 1.$$

It can be easily checked that the assumptions (i)-(iii) are satisfied from our previous assumptions and the model. The assumption (iv) follows directly from the fact that ψ is increasing in k, i.e. $\psi_1 > 0$.

The other assumptions needed for proving the existence of a stationary recursive competitive equilibrium (see Assumption 18.2 in Stokey, Lucas, and Prescott (1989)) are satisfied : (i) the equilibrium marginal return on capital for any $k \in B$ is finite (in our case the interest rate r); and (ii) that $\lim_{c\to 0} u'(c) = \infty$.

Then to prove the Schauder's Theorem, let $C(B,Z)$ be the set of continuous bounded functions $h : B \times Z \to B$ and define a subset $F = \{h \in C(B,Z)\}$ where the function h satisfies $0 \leq h(k,z) \leq \psi(k,z)$, all $(k,z) \in B \times Z$, and h and $\psi - h$ are nondecreasing. Note that $B \times Z$ is a bounded subset of $\mathbb{R}^2$ and that the family of functions F is nonempty, closed, bounded, and convex. Define an operator T on F

$$\begin{aligned} u'(\psi(k,z) - (Th)(k,z)) \;&=\; \beta \sum_{z'} u'(\psi((Th)(k,z),z') - h[(Th)(k,z),z']) \\ &\quad \cdot [(1 - \tau(y(\cdot)) - \tau'(y(\cdot))y(\cdot))\, r + 1]\, Q(z,z'), \end{aligned}$$

where $y(\cdot) = y((Th)(k,z),z')$.

Then it is easy to prove that T is well defined, continuous and that $T : F \to F$. From the conditions on function h and finite return on capital, it follows that F is an equicontinuous family. That the operator T has a fixed point in F follows from the Schauder's Theorem (see e.g. Theorem 17.4 in Stokey, Lucas, and Prescott (1989)).

The existence of the stationary recursive competitive equilibrium is standard from the monotonicity, Feller and mixing property of Q and the non-decreasing policy functions (see Chapter 12 in Stokey, Lucas, and Prescott (1989)). $\square$

Proof of Proposition 1

First, construct a Lagrangian for the Ramsey problem as

$$L = \sum_z \int \mathcal{W}(\tau, k', \lambda) + \mu_F \left(\sum_{z'} \mathcal{F}(\tau, k') \right) + \mu_\lambda \left(\sum_z \mathcal{L}(\tau, k', \lambda) \right),$$

where μ_F and μ_λ are multipliers related to the Euler equation operator and the stationary distribution operator, respectively. The first order conditions are the derivatives of L with respect to τ, μ_F and μ_λ,

$$\frac{d}{d\tau} L(\tau, k', \lambda) \equiv \left(\frac{d}{d\varepsilon} L(\tau, k', \lambda) \right)_{\varepsilon=0} = 0,$$

and similarly,

$$\begin{aligned} \frac{d}{d\mu_F} L(\tau, k', \lambda) = \mathcal{F}(\tau, k') = 0, \\ \frac{d}{d\mu_\lambda} L(\tau, k', \lambda) = \mathcal{L}(\tau, k', \lambda) = 0. \end{aligned}$$

Further,

$$\frac{d}{d\tau} L(\tau, k', \lambda) = \sum_z \int \frac{d}{d\tau} \mathcal{W}(\tau, k', \lambda) dk + \mu_F \sum_{z'} \frac{d}{d\tau} \mathcal{F}(\tau, k') + \mu_\lambda \sum_z \frac{d}{d\tau} \mathcal{L}(\tau, k', \lambda).$$

The stationary recursive competitive equilibrium responds to a tax policy change by changing the saving policy and the stationary distribution function. These effects on the policy k' and the distribution function λ are captured by 'sensitivity' functions $\mathcal{D}k'$ and $\mathcal{D}\lambda$, respectively. Then,

$$\begin{aligned} \frac{d}{d\tau} \mathcal{W}(\tau, k', \lambda) &= \frac{\partial \mathcal{W}(\tau, k', \lambda)}{\partial \tau} + \frac{\partial \mathcal{W}(\tau, k', \lambda)}{\partial k'} \mathcal{D}k' + \frac{\partial \mathcal{W}(\tau, k', \lambda)}{\partial \lambda} \mathcal{D}\lambda, \\ \frac{d}{d\tau} \mathcal{F}(\tau, k') &= \frac{\partial \mathcal{F}(\tau, k')}{\partial \tau} + \left[\frac{\partial \mathcal{F}(\tau, k')}{\partial k'} + \frac{\partial \mathcal{F}(\tau, k')}{\partial (k'(k'))} \frac{\partial k'(k')}{\partial k'} \right] \mathcal{D}k' + \frac{\partial \mathcal{F}(\tau, k')}{\partial (k'(k'))} \mathcal{D}k'(k'), \\ \frac{d}{d\tau} \mathcal{L}(\tau, k', \lambda) &= \frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial \tau} + \frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial k'} \mathcal{D}k' + \frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial \lambda} \mathcal{D}\lambda. \end{aligned}$$

Since the Euler and the stationary distribution operators $(\mathcal{F}, \mathcal{L})$ are specified in an implicit form, the 'total' derivatives with respect to the tax function have to be equal to zero, i.e.,

$$\begin{aligned}
\sum_{z'} \frac{\partial \mathcal{F}(\tau, k')}{\partial \tau} + \left[\frac{\partial \mathcal{F}(\tau, k')}{\partial k'} + \frac{\partial \mathcal{F}(\tau, k')}{\partial (k'(k'))} \frac{\partial k'(k')}{\partial k'} \right] \mathcal{D}k' + \frac{\partial \mathcal{F}(\tau, k')}{\partial (k'(k'))} \mathcal{D}k'(k') &= 0 \\
\sum_{z} \frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial \tau} + \frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial k'} \mathcal{D}k' + \frac{\partial \mathcal{L}(\tau, k', \lambda)}{\partial \lambda} \mathcal{D}\lambda &= 0.
\end{aligned}$$

Thus, the first order condition (38) simplifies to

$$\sum_{z} \int \left(\frac{\partial \mathcal{W}(\tau, k', \lambda)}{\partial \tau} + \frac{\partial \mathcal{W}(\tau, k', \lambda)}{\partial k'} \mathcal{D}k' + \frac{\partial \mathcal{W}(\tau, k', \lambda)}{\partial \lambda} \mathcal{D}\lambda \right) dk = 0.$$

Therefore, the first order conditions are given by (20)-(22) and (23)-(24). These conditions form a system of five functional equations in the five unknown functions which was to be proved. □

Lemma 1

Lemma 1 *The first order necessary conditions in Proposition 1 for the Ramsey problem in Definition 4 form the following system of operator equations:*

$$\begin{aligned}
\frac{\partial \mathcal{W}(k', \lambda, \tau)}{\partial k'} &= -u'(c[k'(k,z), z; \tau])\, \lambda(k,z), && (38) \\
\frac{\partial \mathcal{W}(k', \lambda, \tau)}{\partial \lambda} &= u(c[k'(k,z), z; \tau]), && (39) \\
\frac{\partial \mathcal{W}(k', \lambda, \tau)}{\partial \tau} &= u'(c[k'(k,z), z; \tau]) \frac{\partial c[k'(k,z), z; \tau]}{\partial \tau}\, \lambda(k,z), && (40) \\
\frac{\partial \mathcal{F}(k', \tau)}{\partial k'} &= -u''(c[k'(k,z), z; \tau]) \\
&- \beta u''(c[k'(k'(k,z), z'), z'; \tau] \frac{\partial c[k'(k'(k,z), z'), z'; \tau]}{\partial k'} \\
&\cdot R[k'(k,z), z'; \tau] Q(z,z') + \beta u'(c[k'(k'(k,z), z'), z'; \tau]) \\
&\cdot \left(2 \frac{\partial \tau[y(k'(k,z), z')]}{\partial y} + \frac{\partial^2 \tau[y(k'(k,z), z')]}{\partial y^2} y(k'(k,z), z') \right) \\
&\cdot \frac{\partial y(k'(k,z), z')}{\partial k}\, rQ(z,z'), && (41) \\
\frac{\partial \mathcal{F}(k', \tau)}{\partial \tau} &= u''(c[k'(k,z), z; \tau]) \frac{\partial c[k'(k,z), z; \tau]}{\partial \tau} \\
&- \beta u''(c[k'(k'(k,z), z'), z'; \tau] \frac{\partial c[k'(k'(k,z), z'), z'; \tau]}{\partial \tau} \\
&\cdot R[k'(k,z), z'; \tau] Q(z,z')
\end{aligned}$$

$$
\begin{aligned}
&+\quad \beta u'(c[k'(k'(k,z),z'),z';\tau])\frac{dR[k'(k,z),z';\tau]}{d\tau}Q(z,z'), &(42)\\
\frac{\partial \mathcal{F}(k',\tau)}{\partial k''} &= -\beta u''(c[k'(k,z),z;\tau])R[k'(k,z),z';\tau]Q(z,z'), &(43)\\
\frac{\partial \mathcal{L}(k',\lambda,\tau)}{\partial k'} &= \frac{\partial \lambda[(k')^{-1}(k,z),z)]}{\partial k}\left\{\frac{\partial k'[(k')^{-1}(k,z),z)]}{\partial k}\right\}^{-1}Q(z,z'), &(44)\\
\frac{\partial \mathcal{L}(k',\lambda,\tau)}{\partial \lambda} &= 0, &(45)\\
\frac{\partial \mathcal{L}(k',\lambda,\tau)}{\partial \tau} &= \mathcal{D}\lambda(k,z)-\mathcal{D}\lambda[(k')^{-1}(k,z),z)]\,Q(z,z'), &(46)
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{\partial c[k'(k,z),z';\tau]}{\partial \tau} &= \left(1-\tau[y(k,z)]-\frac{\partial \tau[y(k,z)]}{\partial y}y(k,z)\right)\frac{\partial y(k,z)}{\partial \tau}-y(k,z),\\
R[k'(k,z),z';\tau] &= \left(1-\tau(y[k'(k,z),z'])-\frac{\partial \tau[y(k'(k,z),z')]}{\partial y}\cdot\right.\\
&\cdot\quad y(k'(k,z),z'))\,r+1,\\
\frac{\partial c[k'(k'(k,z),z'),z';\tau]}{\partial k'} &= -y(k'(k,z),z')+1-\frac{\partial k'(k'(k,z),z')}{\partial k}\\
&+\quad (R[k'(k,z),z';\tau]-1)\frac{\partial y(k'(k,z),z')}{\partial k},\\
\frac{dR[k'(k,z),z';\tau]}{d\tau} &= \left[-1-\left(2\frac{\partial \tau[y(k'(k,z),z')]}{\partial y}+\frac{\partial^2 \tau[y(k'(k,z),z')]}{\partial y^2}\cdot\right.\right.\\
&\cdot\quad \left.y(k'(k,z),z'))\,\frac{\partial y(k'(k,z),z')}{\partial \tau}\right]r+(R[k'(k,z),z';\tau]-1)\frac{\partial r}{\partial \tau},\\
\frac{\partial y(k,z)}{\partial \tau} &= \frac{\partial r}{\partial \tau}k+\frac{\partial w}{\partial \tau}z,\\
\frac{\partial y(k'(k,z),z')}{\partial \tau} &\equiv \frac{\partial r}{\partial \tau}k'(k,z)+\frac{\partial w}{\partial \tau}z',\\
\frac{\partial r}{\partial \tau} &= F_{KK}(K,L)\,K_\tau,\\
\frac{\partial w}{\partial \tau} &= F_{LK}(K,L)\,K_\tau,\\
K_\tau &= \sum_z\int k\,\mathcal{D}\lambda(k,z)\,dk.
\end{aligned}
$$

Note that in equation (44) the current capital is determined backwards from the next period capital k' and the current shock z, i.e. $k=(k')^{-1}(k',z)$. It means that both terms on the right hand side of the equation $\partial\lambda/\partial k=\partial\lambda(k,z)/\partial k$ and $\partial k'/\partial k$ are evaluated at $k=(k')^{-1}(k',z)$. □

References

Aiyagari, R. S. (1995). Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy 103*, 1158–1175.

Albanesi, S. and C. Sleet (2003). Dynamic optimal taxation with private information. *Mimeo, University of Iowa*.

Auerbach, A. J. and L. J. Kotlikoff (1987). *Dynamic Fiscal Policy.*

Bohacek, R. and M. Kejak (2002). Projection methods for economies with heterogeneous agents. *CERGE-EI Working Paper*.

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica 54*(3), 607–622.

Diaz-Jimenez, J., V. Quadrini, and J. V. Rios-Rull (1997, Spring). Dimensions of inequality: Facts on the u.s. distribution of earnings, income, and wealth. *Federal Reserve Bank of Minneapolis Quarterly Review*.

Domeij, D. and J. Heathcote (2004). On the distributional effects of reducing capital taxes. *International Economic Review 45*(2), 523–554.

Garcia-Mila, T., A. Marcet, and E. Ventura (1995). Supply side interventions and redistributions. *Working Paper UPF 115.*

Golosov, M., N. R. Kocherlakota, and O. Tsyvinski (2003). Optimal indirect and capital taxation. *Review of Economic Studies 70*, 569–587.

Heaton, J. and D. Lucas (1996). Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy 104*(3), 443–487.

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics 28*, 59–83.

Judd, K. L. (1992). Projection methods for solving aggregate growth models. *Journal of Economic Theory 58*, 410–452.

Judd, K. L. (1998). *Numerical Methods in Economics.* Cambridge, MA: MIT Press.

Kamien, M. I. and N. L. Schwartz (1991). *Dynamic Optimization: The Calculus Of Variations and Optimal Control in Economics and Management* (2nd ed.). New York: North-Holland.

Klein, P. and J.-V. Rios-Rull (2004). Time-consistent optimal fiscal policy. *International Economic Review 44*(4), 1217–1245.

Kocherlakota, N. R. (2003). Zero expected wealth taxes: A mirrless approach to dynamic optimal taxation.

Kydland, F. and E. C. Prescott (1977). Rules rather than discretion: the inconsistency of optimal plans. *Journal of Political Economy 84*, 473–491.

Lucas, Jr., R. E. (1990). Supply-side economics: An analytical review. *Oxford Economic Papers* (42), 293–316.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*, 175–208.

Mirrlees, J. (1976). Optimal tax theory: A synthesis. *Journal of Public Economics 6*, 327–358.

Prescott, E. C. (1986). Theory ahead of business cycle measurement. *Carnegie-Rochester Conference Series on Public Policy 25*, 11–44.

Quah, D. (2003). One third of the world's growth and inequality. In T. Eicher and S. Turnovsky (Eds.), *Growth and Inequality: Issues and Policy Implications*, Cambridge. MIT Press.

Ramsey, F. P. (1927). A contribution to the theory of taxation. *The Economic Journal 37*, 47–61.

Stokey, N. L., R. E. Lucas, Jr., and E. C. Prescott (1989). *Recursive Methods in Economic Dynamics*. Cambridge: Harvard University Press.

Storesletten, K., C. Telmer, and A. Yaron (1999). Asset pricing with idiosyncratic risk and overlapping generations. *Mimeo, Carnegie Mellon University*.

Tauchen, G. and R. Hussey (1991). Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica 59*, 371–396.

Ventura, G. (1999). Flat tax reform: A quantitative exploration. *Journal of Economic Dynamics and Control 23*, 1425–1458.

Steady State Results			
	Tax Schedule		
Average	Progressive	Flat Rate	Optimal
Capital	2.54	3.29	3.80
Output	1.39	1.54	1.62
Consumption	0.86	0.90	0.91
Median Wealth	2.57	2.85	3.63
Median/Mean Wealth Ratio	1.012	0.866	0.955
Interest Rate (%)	9.83	6.79	5.31
Wage	0.895	0.983	1.035
Gini Total Income	0.22	0.31	0.27
Gini Total After Tax Income	0.21	0.32	0.28
Welfare	0.861	0.895	0.902
Welfare Gain (%)	4.91	0.89	—

Notes: Wealth in terms of accumulated capital stock. Welfare measured as consumption level corresponding to the average utility. Welfare gain measured as percentage of the average consumption each agent would have to receive in the progressive and the flat-rate tax economy so that the average welfare equals that in the optimal tax economy.

Table 1: Steady State Results.

Steady State Distribution of Resources					
	Quintile				
Tax Schedule	1st	2nd	3rd	4th	5th
	Average Consumption Level				
Optimal	0.6673	0.8272	0.9171	1.0076	1.1489
Flat	0.6990	0.8189	0.8904	0.9685	1.1213
Progressive	0.7120	0.8113	0.8642	0.9237	1.0148
	Average Asset Level				
Optimal	0.9166	2.4051	3.6545	4.9934	7.0381
Flat	0.6336	1.7708	2.8488	4.1933	6.9921
Progressive	0.7705	1.8195	2.5666	3.2509	4.2725
	Average Investment/Income Ratio				
Optimal	0.1222	0.2148	0.2935	0.3692	0.4643
Flat	0.0869	0.1630	0.2363	0.3197	0.4624
Progressive	0.1058	0.1774	0.2252	0.2592	0.2964
	Average Income Level				
Optimal	1.0141	1.1520	1.2353	1.3221	1.4620
Flat	0.9644	1.1009	1.1881	1.2883	1.4884
Progressive	0.9158	1.0565	1.1410	1.2337	1.3739
	Average After-Tax Income Level				
Optimal	0.6996	0.8341	0.9142	0.9964	1.1240
Flat	0.7194	0.8213	0.8863	0.9611	1.1103
Progressive	0.7319	0.8168	0.8644	0.9184	0.9948
	Average Tax Contribution Share				
Optimal	0.1944	0.1957	0.1995	0.2014	0.2090
Flat	0.1588	0.1824	0.1982	0.2133	0.2473
Progressive	0.1311	0.1702	0.1988	0.2258	0.2741

Table 2: Distribution of Resources in Steady States.

Transition to the Optimal Tax Schedule Steady State		
	Transition From	
	Progressive	Flat Rate
	Average Agent	
Welfare in Steady State	0.861	0.895
Welfare in Transition	0.895	0.885
Welfare Gain (%)	3.84	–1.08
	Median Agent	
Welfare in Steady State	0.863	0.872
Welfare in Transition	0.901	0.860
Welfare Gain (%)	4.40	–1.40
	Political Support	
% of Population	73.1	37.9

Notes: Welfare in the steady state defined as in the note to Table 1. Welfare in transition measured as consumption level corresponding to the average expected present discounted value from transition. Welfare gain is a percentage of the average consumption each agent would have to receive in each period of transition so that the average welfare from transition equals in expected present discounted value that of the initial steady state.

Table 3: Transition to the Optimal Tax Schedule Steady State.

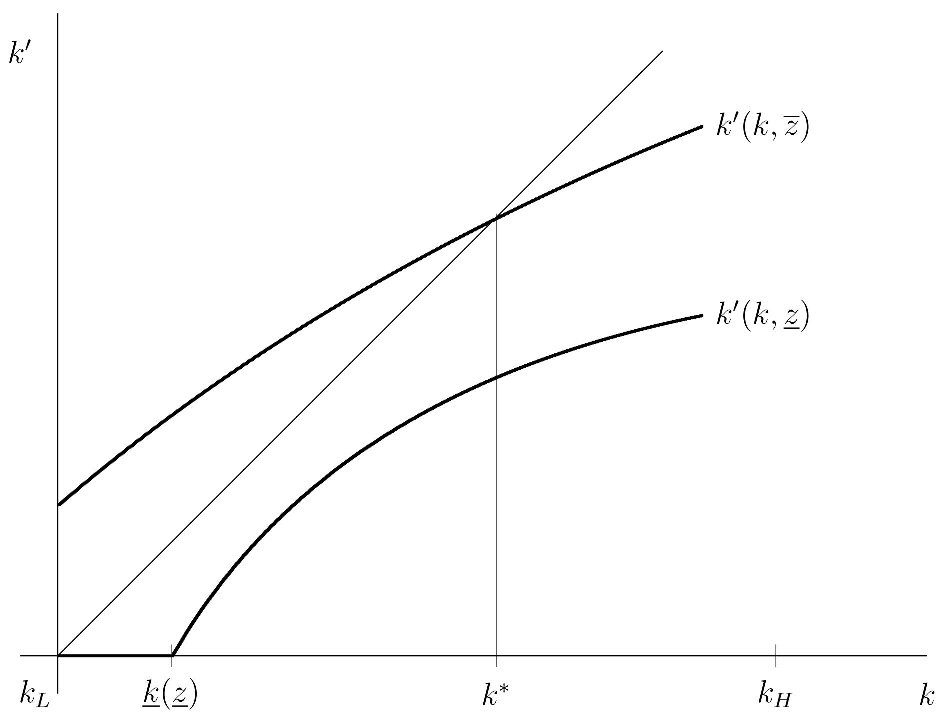

Figure 1: Policy functions for the next period capital stock. An example with two productivity shocks $\overline{z} > \underline{z}$. There is an exogenous lower bound k_L and an endogenous upper bound $k^* < k_H$. The stationary distribution has a unique ergodic set $E = [k_L, k^*]$. Agents with shock $\underline{z}$ and capital stock $k < \underline{k}(\underline{z})$ are borrowing constrained.

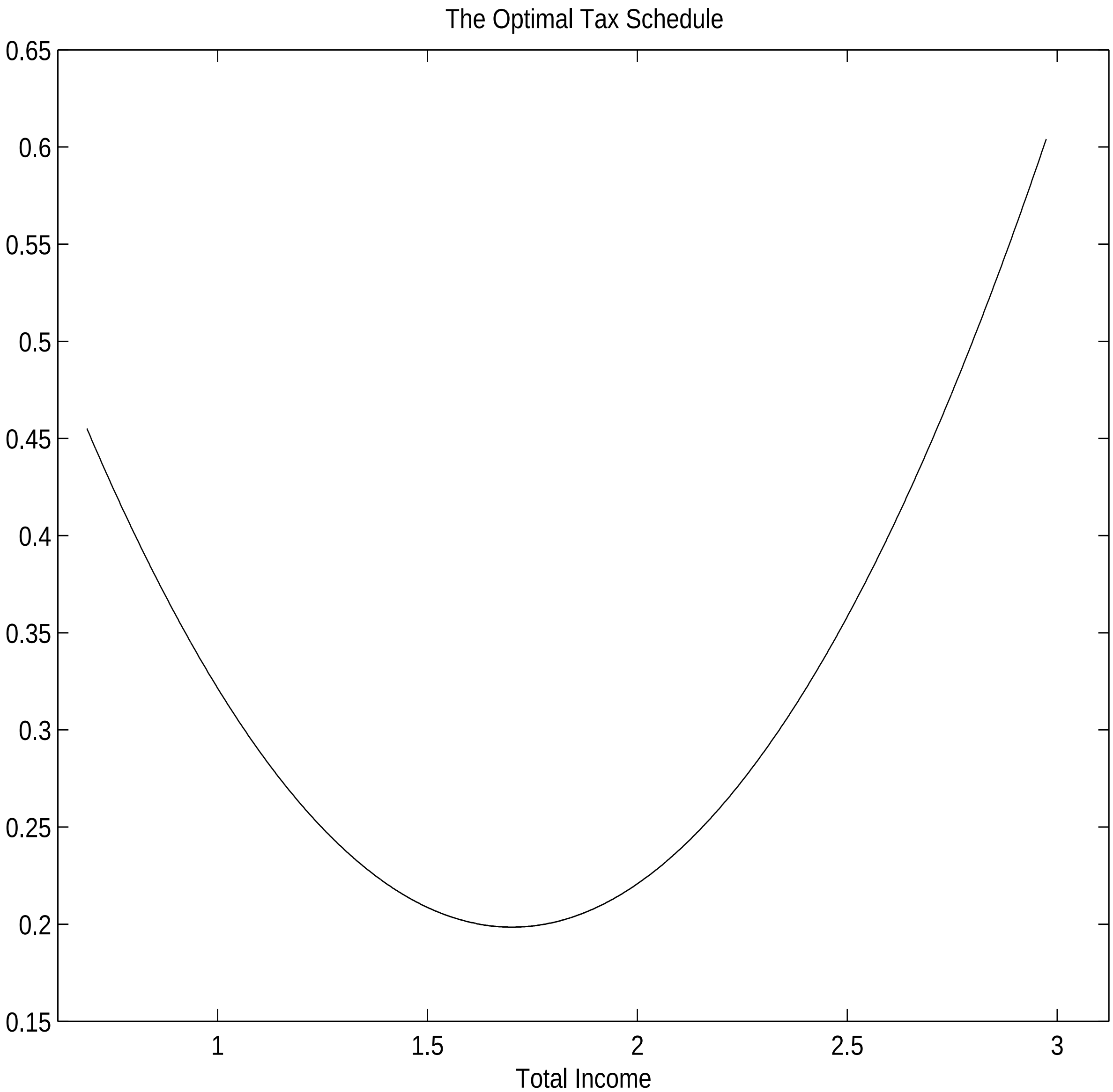

Figure 2: The Optimal Tax Schedule.

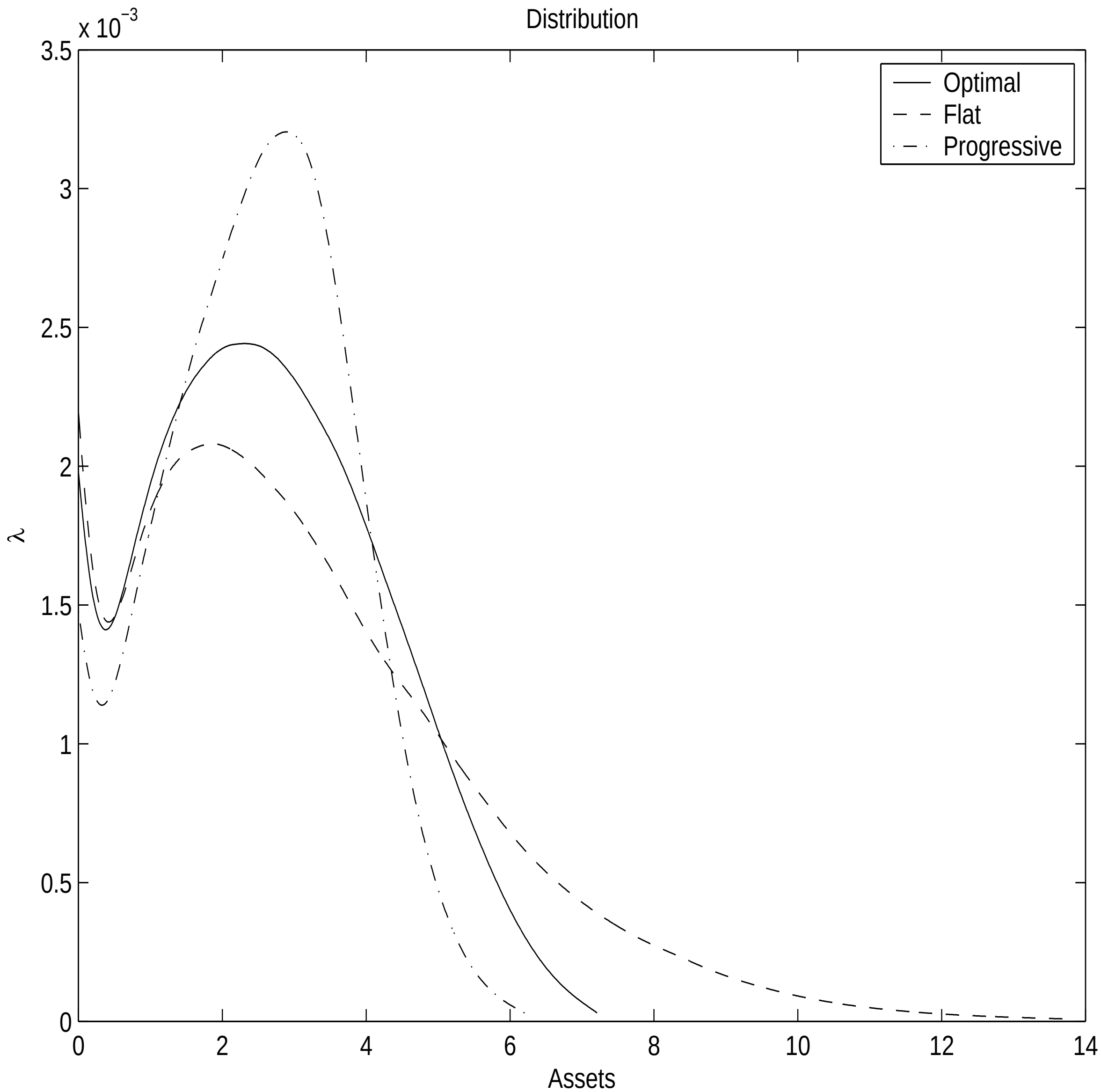

Figure 3: Stationary Distribution of Agents Over Assets.

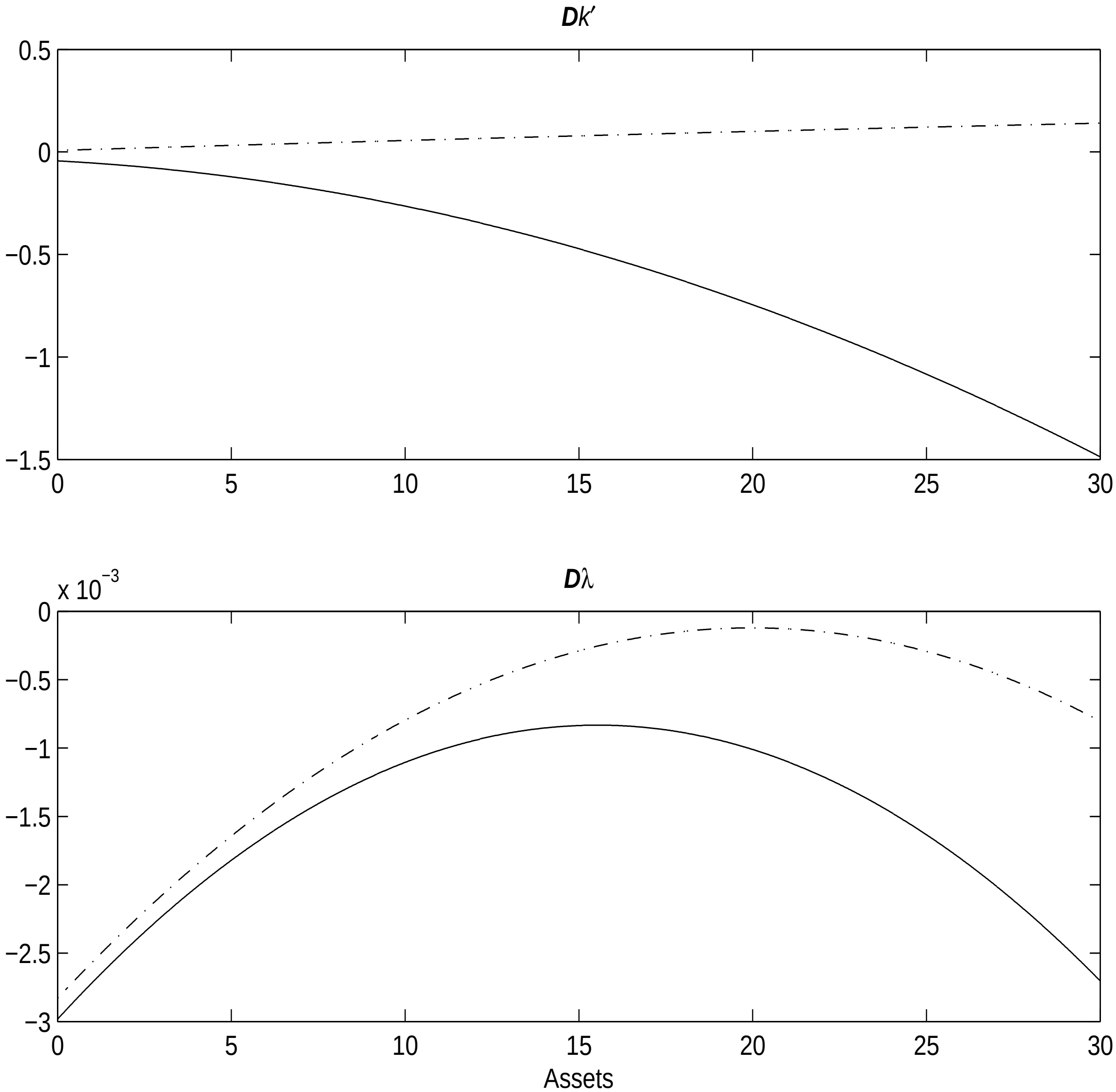

Figure 4: Optimal Policy Functions for High (-) and Low (-.-) Productivity Shocks.

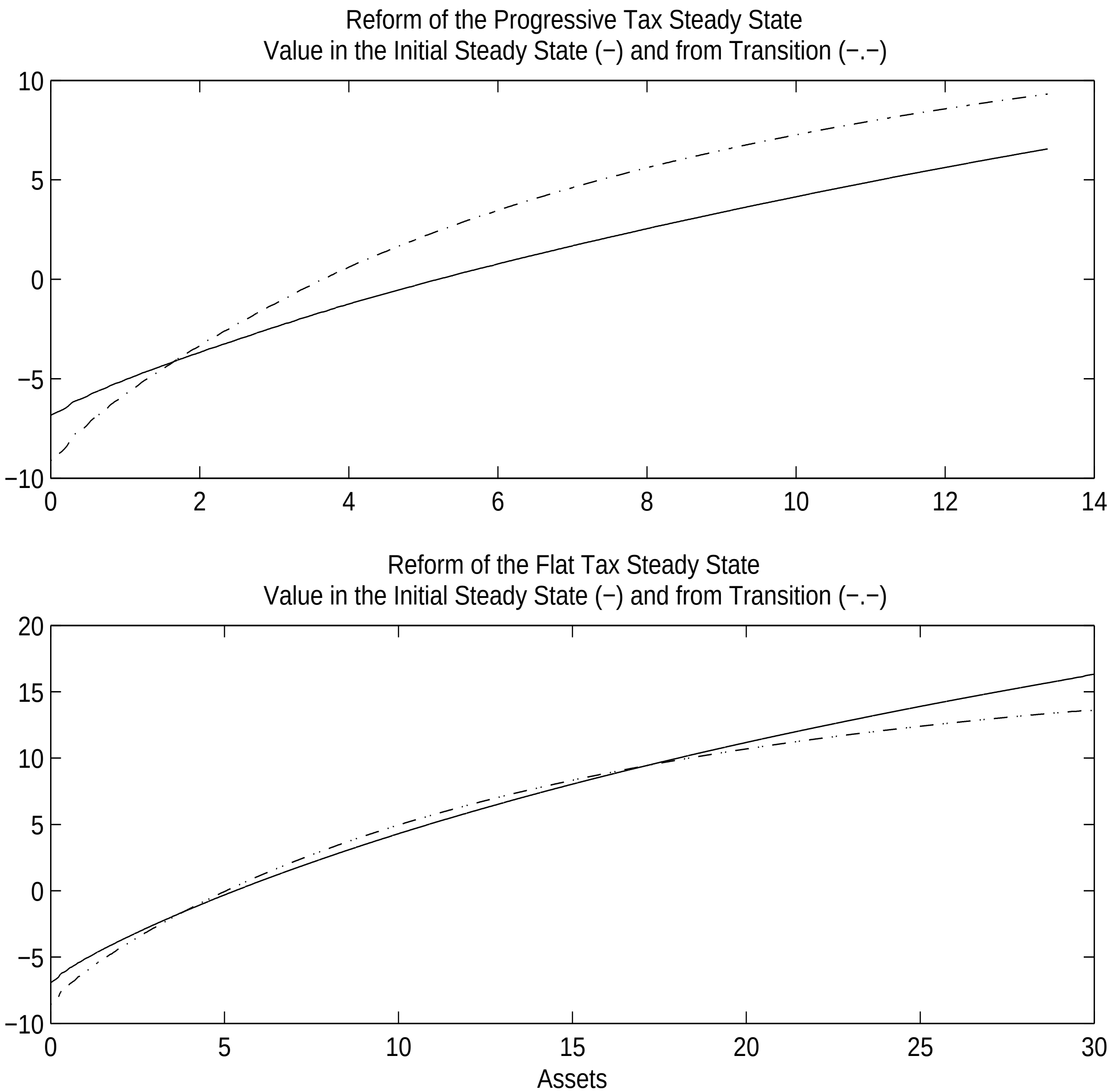

Figure 5: Transition from the Progressive and the Flat Tax Steady State to the Optimal Tax Schedule Steady State.