Optimal Government Policies in Models with Heterogeneous Agents

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

April 7, 2015

Abstract

In this paper we develop a new approach for finding optimal government policies in economies with heterogeneous agents. Using the calculus of variations, we present three classes of equilibrium conditions from government’s and individual agent’s optimization problems: 1) the first order conditions: the government’s Lagrange-Euler equation and the individual agent’s Euler equation; 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 the resulting optimal distribution of agents in the steady state. This approach is applicable to a wide class of general equilibrium, Bewley type economies where the government looks for an optimal nonlinear, second-best fiscal or monetary policy. 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 Jim Costain, Mark Gertler, Max Gillman, Christian Hellwig, Boyan Jovanovic, Marek Kapitka, Dirk Krueger, Per Krusell, Michal Pakos, Josep Pijoan-Mas, Thomas Sargent, Jakub Steiner, Harald Uhlig, Gianluca Violante, Galyna Vereshchagina, and the participants at the CNB/CERGE-EI Macro Workshop 2004, SED 2004 conference, University of Stockholm, the Macroeconomic Seminar at the Federal Reserve Bank of Minneapolis, SED 2007 conference, Cardiff Business School, and the Stern School of Business. We are especially grateful to Michele Boldrin, Tim Kehoe and Stan Zin for their support and advice. Anton Tyutin, Jozef Zubrický and Tomáš Leško provided excellent research assistance. The financial support of the Czech Science Foundation project No. P402/12/G097 (DYME Dynamic Models in Economics) is gratefully acknowledged. All errors are our own.
1 Introduction

This paper provides a new approach for computing equilibria in which the stationary distribution of agents is a part of an optimal nonlinear, second-best government problem in a general equilibrium, Bewley type economy with heterogeneous agents. We formulate the optimal government policy problem as a calculus of variations problem where the government maximizes an objective functional subject to a system of operator 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 necessary conditions of the government functional problem given by the Euler-Lagrange equation (with transversality conditions) form altogether 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 standard projection method.

Our main contribution is the derived Euler-Lagrange equation for the government problem and the operator formulation of the individual agent’s Euler equation and of the endogenous stationary distribution. 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. The first and second order conditions, in the form of the Euler-Lagrange equation and a modified Legendre condition, respectively, represent the necessary and sufficient conditions for concavity and a unique maximum attained by the government policy function. Additionally, the first-order transversality conditions for the boundary agents allow for a qualitative analysis of the shape of the optimal government policy function. It should be emphasized that our approach does not use any additional restrictions or assumptions on the shape of the government policy function and 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. To our knowledge, this paper is the first one that provides a solution method for this kind of optimal government problem in an economy with heterogeneous agents in general equilibrium. The main innovation of our approach lies in its consideration of the general equilibrium effects arising from the endogeneity of the distribution function attained by the optimal policy function.

Our approach can be applied to a wide range of government optimal fiscal or monetary policy problems. We illustrate this general methodology on a steady state Ramsey (1927) problem, solving for the optimal tax schedule on the total income that maximizes aggregate welfare in a steady state of a standard neoclassical, dynamic general equilibrium model with heterogeneous agents and incomplete markets. We compare the steady state aggregate
levels, welfare, efficiency and distribution of resources associated with this 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 literature on optimal taxation has provided important insights due to restrictions imposed either on the functional form of the tax function (Diamond and Mirrlees (1971)) or on information available to the government (Mirrlees (1971) and Golosov, Kocherlakota, and Tsyvinski (2003)). In our example, we study the efficiency-equality tradeoff of the optimal tax function in a full information economy without any prior restriction on the shape of the tax function. The optimal tax schedule we find 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. The marginal tax rate is also a U-shaped function, but almost flat at the low incomes, reaching negative levels around the average income, and then rising to positive levels. The impact of the optimal tax schedule on aggregate levels and welfare is large. Compared to the simulated progressive tax schedule steady state, average welfare increases by 4.4% while relative to the flat-tax steady state, welfare goes up by 0.8%.\footnote{We compute transitions to the steady state associated with the optimal tax schedule and find that a majority of the population would benefit from a reform that replaces the progressive tax schedule by the optimal tax schedule. These results as well as a detailed efficiency and distributional analysis are described in the following sections.}

The efficiency and distributional effects of the optimal tax are the main mechanisms behind these large changes. Efficiency gains arise from the general equilibrium effects: a higher stock of capital increases productivity of labor and, therefore, the income of poor agents. For the steady state distributional effects, the U-shape optimal tax schedule provides incentives for agents to accumulate assets while preserving the equality measures in the economy, close to what a social planner with an access to the first-best, lump-sum transfers would do. The high tax rate at low income levels induces agents to save more for precautionary reasons while the even higher tax rate on high incomes provides resources for redistribution. Around the average total income levels, the optimal tax rate is lower than the one found for the flat-tax reform. For a comparison, the ad hoc flat tax reform also increases aggregate levels but does not take into account the distribution of agents, while the progressive tax schedule provides too much short-run insurance at the cost of the long-run average levels. Finally, the optimal tax schedule results in a very equal distribution of tax contributions across agents.

The shape of the optimal tax function reflects the important tradeoff between efficiency and equality in terms of income or wealth distribution stressed by Diamond and Saez (2011) or Golosov, Tsyvinski, and Troshkin (2011). In partial equilibrium models, Diamond (1998) shows that optimal marginal taxes are U-shaped if the distribution of skills is single-
peaked with a Pareto distribution above the peak. Saez (2001) generalizes these results for preferences with income effects. In a static model calibrated to empirical cross-sectional distribution of labor income and empirical tax rates, he finds that optimal marginal taxes are U-shaped and rise up to 50–70 percent. Mirrlees (1971), Kanbur and Tuomala (1994), and Mankiw et al. (2009) all note that the shape of the marginal tax schedule is very sensitive to the exogenous distribution of skill or income.  

Compared with the Mirrlees literature, in our approach it is the tax schedule that seeks to attain a steady state where the endogenous distribution of agents is optimal with respect to the aggregate welfare in the economy. For dynamic, general equilibrium models with heterogeneous agents and uninsurable idiosyncratic risk, the endogenous steady state distribution is the most important part of the resulting equilibrium. An important step towards quantitatively studying dynamic optimal taxation is made in Kapicka (2013) and Golosov, Tsyvinski, and Troshkin (2010). In a paper closest to our methodology, Golosov, Tsyvinski, and Werquin (2014) use variational approach for their analysis of optimal tax systems. In a partial equilibrium framework with exogenous distribution of agents, they compute Gateaux differentials of local tax perturbations and look for a globally optimal tax function that cannot be locally improved within a restricted class of tax functions. For many realistic parameters the optimal marginal tax rates are also U-shaped.

Such taxes are not allowed by conventional parameterizations in general equilibrium models, where the tax schedule is restricted to a specific, usually linear or monotone, functional form. Several important papers have analyzed the steady state implications (and transition paths) resulting from an ad hoc flat-tax reform or from a removal of double taxation of capital income. Heathcote, Storesletten, and Violante (2014) and Conesa and Krueger (2006) compute the optimal progressivity of the income tax code (the latter based on Gouveia and Strauss (1994)). Our paper shows that narrowing the analysis to monotone functions may be restrictive with respect to welfare maximization.

We limit our example 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 several reasons why we choose this setup. First, the tax on the total income enables us to study a 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. Second, to a large extent the current U.S. tax code does not distinguish between the sources of taxable income. The last reason for a simple tax on the total income is the complexity of the problem we solve. In this paper, we therefore do not address the following important issues related to optimal taxation: the issue of time-consistency, the issue of the optimal

\[\text{\footnotesize{In general, the information-constrained optimal marginal tax rates increase in the tail ratio of the skill distribution and in the desired degree of redistribution; they decrease when labor elasticity is high.}}\]
capital income tax rate in the steady state, and the distortionary effects of taxation on labor supply. Our government is fully and credibly committed, the tax schedule is constant over time.\textsuperscript{3} Due to the complexity of our work, we study the simplest utility maximization problem on the consumption-investment margin.

However, we would like to stress that our methodology can be applied to different aggregate welfare criteria and a wide variety of optimal government policy problems, including those with endogenous labor supply, separate taxation of labor and capital incomes as well as public goods, population growth, or a life-cycle earnings process as in Ventura (1999) or Weinzierl (2011). Finally, in 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 stationary Ramsey problem. Section 3 specifies the equilibrium as a system of functional equations and defines the operator stationary Ramsey problem. Section 4 formulates the Ramsey problem in the calculus of variations. The first-order necessary and the second-order sufficient conditions for the optimal government policy expressed in the form of a generalized Euler-Lagrange equations and Legendre condition, respectively, with related analytical results are described in Section 5. Section 6 illustrates our approach by an example of the optimal income tax schedule with numerical simulation presented in Section 7. Section 8 concludes. Appendices contain proofs, analytical results, and details on the simulated optimal tax schedule.

2 The Economy

The economy is populated by a continuum of infinitely lived agents on a unit interval. Each agent has preferences over consumption $c_t$ in period $t \geq 0$, given by a utility function

$$E \sum_{t=0}^{\infty} \beta^t U(c_t), \quad 0 < \beta < 1,$$

where $U : \mathbb{R}_+ \rightarrow \mathbb{R}$ is a twice continuously differentiable, strictly increasing and strictly concave function. We assume that the utility function satisfies the Inada conditions.

At all $t \geq 0$, each agent is identified by an endogenous state variable, the accumulated stock of capital, $k_t \in B = [k, \infty)$ with $k_0 = 0$, and by an exogenous labor productivity

\textsuperscript{3}For the time-consistency problem see Kydland and Prescott (1977) and Klein and Rios-Rull (2004). Aiyagari (1995) showed that for our class of models with incomplete insurance markets and borrowing constraints, the optimal tax rate on capital income is positive even in the long run (also Conesa et al. (2009)). For the optimal capital tax with two types of agents see Chamley (1986) and Judd (1985).
shock $z_t \in Z = \{z_1, z_2, \ldots, z_J\}$. The shock represents labor 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. We preserve the heterogeneity in the economy by assuming incomplete markets.

In each period, agents supply labor and accumulated capital stock to a representative firm with a production function $F(K_t, L_t)$, where $K_t \in B$ is the aggregate capital stock, $L_t \in \mathbb{R}_+$ is the aggregate effective labor. The production function is concave, twice continuously differentiable, increasing in both arguments, and displays constant returns to scale. Profit maximization implies the following factor prices

$$r_t = F_K(K_t, L_t) - \delta \quad \text{and} \quad w_t = F_L(K_t, L_t),$$

where $\delta \in (0, 1)$ is the depreciation rate of capital.

Finally, there is a government that finances its expenditures by taxing the agents in the economy. We assume the government is fully committed to a sequence of tax schedules $\{\pi_t\}_{t=0}^\infty$ to finance its expenditures. We assume that these expenditures equal a constant fraction of output, $g$, in each period, that they are not returned to the agents, and that the government cannot use the first best, lump-sum taxation.\(^4\) The tax schedule policy is applied to a broadly defined taxable activity of each agent, $x_t \in \mathbb{R}_+$. We will assume that $x_t = x(z_t, k_t)$ where $x : Z \times B \to \mathbb{R}_+$ and $x_z, x_k > 0$. Thus in each period, the policy schedule is a function $\pi_t : \mathbb{R}_+ \to \mathbb{R}$, so that an agent with a total income from labor and capital, $y_t \in \mathbb{R}_+$, $y_t = y(k_t, z_t) = r_t k_t + w_t z_t$, and a taxable activity $x_t = x(k_t, z_t)$ pays taxes $\pi_t(x_t)x_t$ and is left with an after-tax income $y_t - \pi_t(x_t)x_t$.\(^5\)

The economy’s state is characterized by the sequences of government policies $\{\pi_t\}_{t=0}^\infty$ and the distribution of agents over capital and shock in each period, $\{\lambda_t\}_{t=0}^\infty$. The latter is in each period a probability measure defined on subsets of the state space, describing the heterogeneity of agents over their individual state $(z, k) \in Z \times B$. Let $(B, \mathcal{B})$ and $(Z, \mathcal{Z})$ be measurable spaces, where $B$ denotes the Borel sets that are subsets of $B$ and $Z$ is the set of all subsets of $Z$.

\(^4\)Our analysis equally applies to the case when government finances any expenditures $\{G_t\}_{t=0}^\infty$ and the corresponding revenue-neutral reforms.

\(^5\)In Section 6, we illustrate this policy by a proportional taxation of the total income from labor and capital, i.e. when $x = y$. 

6
2.1 Recursive Formulation

We will analyze the economy in a stationary recursive competitive equilibrium in which the government policy schedule and the distribution of agents are time-invariant. Define the value function of each agent as

\[ v : Z \times B \rightarrow \mathbb{R} \]

and the savings function as

\[ h : Z \times B \rightarrow B. \]

Given the equilibrium prices and the time-invariant government policy \( \pi \), an agent \((z, k)\) solves the following dynamic programming problem

\[
v(z, k) = \max_{c, h} \left\{ u(c) + \beta \sum_{z'} v(z', h) Q(z, z') \right\},
\]

subject to a budget constraint

\[
c(z, k) + h(z, k) \leq y + k - \pi(x(z, k)) x(z, k),
\]

with taxable activity \( x(z, k) \), total income \( y = rk + wz \), and a borrowing constraint,

\[
h(z, k) \geq k.
\]

**Definition 1 (Stationary Recursive Competitive Equilibrium)** For a given share of government expenditures \( g \) and a time-invariant government policy schedule \( \pi \), a stationary recursive competitive equilibrium is a set of functions \((v, c, y, x, h)\), aggregate levels \((K, L)\), prices \((r, w)\), and a probability measure \( \lambda : Z \times B \rightarrow [0, 1] \), such that for given prices and government policies,

1. the policy functions solve each agent’s optimization problem (2);
2. firms maximize profit (1);
3. the probability measure is time invariant,

\[
\lambda(z', B') = \sum_z \int_{\{(z, k) \in Z \times B : h(z, k) \in B'\}} Q(z, z') \lambda(z, k) \, dk,
\]

for all \((z', B') \in Z \times B\);
4. the aggregate conditions hold,

\[
K = \sum_z \int h(z, k) \lambda(z, k) \, dk \quad \text{and} \quad L = \sum_z \int z \lambda(z, k) \, dk;
\]
5. the government budget constraint holds at equality,

\[
g = \sum_z \int \pi(x(z, k)) x(z, k) \lambda(z, k) \, dk / F(K, L);
\]
6. and the allocations are feasible,

\[
\sum_z \int \left[ c(z, k) + h(z, k) \right] \lambda(z, k) \, dk + gF(K, L) = F(K, L) + (1 - \delta)K.
\]
Agents have rational expectations, take the behavior of prices as given by a predetermined function that depends on aggregate variables in equation (1). The goal of our paper is to solve the following stationary Ramsey problem.

**Definition 2 (Stationary Ramsey Problem)** A solution to the Ramsey problem for a stationary economy with heterogeneous agents is a time-invariant government policy schedule \( \pi : \mathbb{R}_+ \to \mathbb{R} \) that maximizes social welfare in the steady state,

\[
\max_{\pi} \sum_z \int v(z, k; \pi) \lambda(z, k; \pi) \, dk,
\]

consistent with a given government consumption and with allocations satisfying the definition of the stationary recursive competitive equilibrium.

Our notation stresses that the value and distribution functions depend on the government policy, \( \pi \), i.e. \( v(z, k; \pi) \) and \( \lambda(z, k; \pi) \), respectively.

It is easy to show that the solution to the Stationary Ramsey Problem is equivalent to that of maximizing the average current period utility,

\[
\max_{\pi} \sum_z \int u(c(z, k; \pi)) \lambda(z, k; \pi) \, dk.
\]

In the following Sections, we will characterize the optimal government policy schedule using this latter specification.\(^6\)

Finally, because we do not impose any *ex ante* restrictions on the government policy function \( \pi \), the following condition on the individual savings function is required.

**Assumption 1 (Regularity Condition)** The government policy function \( \pi \) is such that for all \( z \in Z \), the individual savings function \( h : Z \times B \to B \) is a monotone function of \( k \) over the whole interval \([k(z), \infty)\).

A similar condition is generally required for the existence of a unique stationary recursive equilibrium in all models with heterogeneous agents (see Stokey, Lucas, and Prescott (1989)). It implies that the savings function does not display pathological features (for example, that wealthy agents save less than poor agents) and, therefore, that the stationary distribution has a unique ergodic set. We want to make explicit here that this is the only assumption imposed on the model and is completely innocuous.\(^7\)

\(^6\)Our analysis of optimal government policies can also be applied to other types of welfare functions.

\(^7\)The Regularity Condition guarantees that for all \( z \in Z \), the government policy function \( \pi \) is such that for given prices determined by \( K \), there exists an inverse function \( h^{-1} \) assigning a current value of capital \( k \) to savings \( h \) according to \( k = h^{-1}(z, h; K, \pi) \). As usual, the Regularity Condition is used in the law of motion for the stationary distribution in equation (11) below.
3 The Operator Stationary Ramsey Problem

Since the problem is to find an optimal, welfare maximizing time-invariant function \( \pi : \mathbb{R}_+ \to \mathbb{R} \), the Stationary Ramsey Problem can be transformed into an operator form and solved by the calculus of variations. The calculus of variations is much more suitable for solving our problem with a complicated set of functional constraints and complex boundary conditions than dynamic programming or optimal control methods.

In order to express the stationary recursive competitive equilibrium in this form, we define two operators: an operator on the Euler equation \( F \) and on the stationary distribution \( L \). For a given government policy schedule \( \pi : \mathbb{R}_+ \to \mathbb{R} \) the Euler equation operator \( F \) is defined on the savings function \( h : Z \times B \to B \), while the stationary distribution operator \( L \) is defined on the probability measure \( \lambda : Z \times B \to [0, 1] \) as well as on the savings function \( h \). We will assume that these functions are square integrable functions on some closed domain\(^8\): \( h, \lambda \in L^2(Z \times B) \) where \( L^2(Z \times B) \) is a Hilbert space with the inner product \( (u, v) = \int_{Z \times B} u(t)v(t)dt \). The operator \( F : C^1(Z \times B) \subset L^2(Z \times B) \to C^1(Z \times B) \subset L^2(Z \times B) \) is the mapping from a space of continuously differentiable functions into a space of continuously differentiable functions; and the operator \( L : C^1(Z \times B) \times C^1(Z \times B) \to C^1(Z \times B) \subset L^2(Z \times B) \).

In order to simplify our notation, from now on we will denote allocations in the next period with the superscript plus sign.

**Operator \( F \) on the Euler Equation** An individual agent’s allocations are characterized by the Euler equation from the optimization problem (2)-(4). For all \((z, k) \in Z \times B\),

\[
 u'(y - \pi(x)x + k - k^+) \geq \beta \sum_{x^+} u'(y^+ - \pi(x^+)x^+ + k^+ - k^{++}) \cdot \left\{ 1 + y_k^+ \left[ \pi(x^+) + \pi'(x^+)x^+ \right] x_k^+ \right\} Q(z, z^+),
\]

where \( y = rk + wz, x = x(z, k), y^+ = rk^+ + wz^+, x^+ = x(z^+, k^+), x_k^+ = x_k(z^+, k^+), \) and \( y_k^+ = r \frac{\partial k^+}{\partial k} \). Clearly, \( y - \pi(x)x \) is disposable income and \( \left\{ 1 + y_k^+ - [\pi(x^+) + \pi'(x^+)x^+]x_k^+ \right\} \) is the next-period ‘after-policy’ marginal return to capital where \( \pi' \) is the marginal government policy function (here the marginal tax schedule).

The solution of the Euler equation is a time invariant savings function \( h : Z \times B \to B \). In the stationary equilibrium, prices, as functions of the aggregate capital \( K \), are constant. When the government searches for the optimal policy schedule it needs to take into account the effect of its policy on equilibrium prices. For this reason it will be advantageous to introduce aggregate capital, \( K \), as an explicit variable in the specification of the equilibrium.

\(^8\) In more precise terms we actually assume that the functions are from the subspace \( W^{1,2}(Z \times B) \) which contains \( L^2(Z \times B) \)-functions which have weak derivatives of order one.
In order to make this effect more transparent, we will write the equilibrium prices as \( r(K) \) and \( w(K) \). As the savings function depends on the aggregate capital stock (through equilibrium prices) and on the government policy, \( \pi \), we denote \( k^+ = h(z, k; K, \pi) \) and \( k^{++} = h(z^+, k^+; K, \pi) = h(z^+, h(z, k; K, \pi); K, \pi) \). Since \( K = \sum_z k \lambda(z, k; \pi) dk \), the optimal savings function is affected through prices by the stationary distribution.

In the text below, we present only the case of the unconstrained agents (i.e. those for whom the Euler equation holds with equality and \( k^+(z, k; K, \pi) > k \)). The case of the borrowing constrained agents is presented in Appendix A.

The operator on Euler equation \( F \) is defined as
\[
F(h; \pi) = u'(c) - \beta \sum_{z^+} u'(c^+) \left\{ 1 + y_k^+ - \left[ \pi(x^+) + \pi'(x^+) x_k^+ \right] x_k^+ \right\} Q(z, z^+),
\]
where
\[
c = y(z, k; K) - \pi(x(z, k; K)) x(z, k; K) + k - h(z, k; K, \pi),
\]
\[
y(z, k; K) = r(K) k + w(K) z,
\]
\[
c^+ = y(z^+, h(z, k; K, \pi); K) - \pi(x(z^+, h(z, k; K, \pi); K)) x(z^+, h(z, k; K, \pi); K)
+ h(z, k; K, \pi) - h(z^+, h(z, k; K, \pi); K, \pi),
\]
and the terms \( y_k^+ = y_k(z^+, h(z, k; K, \pi); K) \) and \( x_k^+ = x_k(z^+, h(z, k; K, \pi); K) \) are the marginal effects of individual savings on income \( y \) and taxable activity \( x \) in the next period. The operator equation is simply \( F(h; \pi) = 0 \).

**Operator \( L \) on the Stationary Distribution**  
In the stationary recursive competitive equilibrium and under the Regularity Condition in Assumption 1, the operator \( L \) for equation (5) is
\[
L(\lambda, h; \pi) \equiv \lambda(z^+, k^+; K, \pi) - \sum_z \lambda[z, h^{-1}(z, k^+; K, \pi); K, \pi] Q(z, z^+),
\]
for all \((z^+, k^+) \in Z \times [h(k(z), z), \overline{h}(z)]\). The related operator equation is \( L(\lambda, h; \pi) = 0 \).

**Definition 3 (Operator Stationary Recursive Competitive Equilibrium)**  
*Given a share of government expenditures \( g \) and a time-invariant government policy schedule \( \pi \), an operator stationary recursive competitive equilibrium is a set of operators \( (F, L) \), prices \( (r, w) \), a savings function \( h \), a probability measure \( \lambda \), and aggregate levels \( (K, L) \), such that for given prices and government policies,*

1. *the policy function \( h \) is the solution to each agent’s optimization problem in the operator equation, \( F(h; \pi) = 0 \),*
2. *firms maximize profit (1),*
3. the time-invariant probability measure is the solution to the operator equation, 
\[ \mathcal{L}(h, \lambda; \pi) = 0, \]
4. the capital and labor markets clear, (6),
5. the government budget constraint (7) holds at equality,
6. and the allocations are feasible, (8).

We can now specify an operator version of the Stationary Ramsey Problem.

**Definition 4 (Operator Stationary Ramsey Problem)** A solution to the Operator Stationary Ramsey Problem in a stationary recursive competitive equilibrium in Definition 1 is a time-invariant government policy \( \pi \) that maximizes social welfare in the steady state,

\[
\arg\max_\pi \sum_z \int_{\underline{k}}^{\bar{k}} W(v(z, k; K, \pi)) \lambda(z, k; K, \pi) \, dk,
\]

subject to a system of operator equations \( \mathcal{F}(h; \pi) = 0 \) and \( \mathcal{L}(h, \lambda; \pi) = 0 \), consistent with equilibrium prices (1) and the market clearing conditions (6) in Definition 1, where \( v : Z \times B \rightarrow \mathbb{R} \) is the value function of individual agents, \( \lambda : Z \times B \rightarrow [0, 1] \) is the stationary distribution, \( W \) is a positive linear social aggregator function, \( \underline{k} \) is the exogenous lower bound and \( \bar{k} \) is the endogenous upper bound on individual savings, i.e. \( \bar{k} = h(z, \bar{k}; K, \pi) \).

4 Ramsey Problem as Calculus of Variations Problem

The first-order conditions for the solution \( \pi \) to the Operator Stationary Ramsey Problem in Definition 4 are best formulated in the calculus of variations.\(^9\) As it is standard, the government policy function \( \pi \) and its derivative \( \pi' \) are treated as two independent functions. Since we are looking for the government policy as a function of individual activity \( x \) rather than capital \( k \), we reformulate the problem with taxable activity \( x \) as the independent variable.

The social welfare function in equation (12) at new coordinates \( x \) is

\[
\sum_z \int_{\underline{x}}^{\bar{x}} W[u(c(z, k; K, \pi, \pi'))] \lambda(z, k; K, \pi, \pi') \, dk = \sum_z \int_{\underline{x}(z)}^{\bar{x}(z)} W[z, x; K, \pi(x), \pi'(x)] \, dx
\]

with

\[
W[z, x; K, \pi(x), \pi'(x)] \equiv W[u(c(z, k(z, x; K); K, \pi, \pi'))] \lambda[z, k(z, x; K); K, \pi, \pi'] k_x(z, x; K),
\]

\(^9\)Mirrlees (1976) also uses the calculus of variations to derive the first-order conditions for the optimal income tax schedule. However, while his problem is a static one with exogenously imposed distribution of unobserved abilities, we have a dynamic problem with an endogenous distribution of agents.
where \( k_x [z, x; K] = [x_k(z, k(z, x; K); K)]^{-1} \) is the inverse function to the marginal effect of individual savings on the taxable activity \( x_k \).

Before we proceed further, we have to clarify two important aspects of this dynamic optimal problem. First, observing that \( K \), which determines the equilibrium prices, is one of the arguments in the objective function above, we have to use the condition on the aggregate capital stock. Writing the condition properly using the fact that \( K \) also depends on the government policy \( \pi \) and its derivative \( \pi' \), we get at the new coordinates

\[
K = \sum_z \int_{\underline{k}}^{\overline{k}} k \lambda(z, k; K, \pi, \pi') \, dk = \sum_z \int_{\underline{\pi}(z)}^{\overline{\pi}(z)} K[z, x; K, \pi(x), \pi'(x)] \, dx,
\]

where

\[
K[z, x; K, \pi(x), \pi'(x)] \equiv k(z, x; K) \lambda[z, k(z, x; K); K, \pi, \pi'] k_x (z, x; K).
\]

Second, the bounds on taxable activity, \( \underline{x}(z) \) and \( \overline{x}(z) \), for each \( z \in Z \), are endogenous functions of a chosen government policy. The lower bound \( \underline{x}(z) = \underline{x}(z, k; K) \) depends on \( z \), on the exogenously given lower bound on capital \( \underline{k} \), and on the equilibrium prices determined by \( K \). The upper bound \( \overline{x}(z) = \overline{x}(z, \overline{k}; K) \) depends on \( z \), on the endogenous upper bound of capital \( \overline{k} \), and also on \( K \).

Finally, we also need to reformulate the side condition of the problem of the government budget constraint from equation (7), i.e.

\[
\sum_z \int_{\underline{k}}^{\overline{k}} [\pi(x(z, k; K))x(z, k; K) - gy(z, k; K)] \lambda(z, k; K, \pi, \pi') \, dk = \sum_z \int_{\underline{\pi}(z)}^{\overline{\pi}(z)} G[z, x; K, \pi(x), \pi'(x)] \, dx,
\]

where

\[
G[z, x; K, \pi(x), \pi'(x)] \equiv [\pi(x)x - gy(z, k(z, x; K); K, \pi, \pi')] \lambda[z, k(z, x; K); K, \pi, \pi'] k_x (z, x; K).
\]

**Definition 5 (Calculus of Variations Ramsey Problem)** The Ramsey problem in the calculus of variations is formulated as the following generalized isoperimetric problem,

\[
\max \sum_z \int_{\underline{\pi}(z)}^{\overline{\pi}(z)} W[z, x; K, \pi(x), \pi'(x)] \, dx,
\]

subject to

\[
\sum_z \int_{\underline{\pi}(z)}^{\overline{\pi}(z)} G[z, x; K, \pi(x), \pi'(x)] \, dx = 0,
\]

\(^{10}\text{Clearly, the maximal interval is } [\underline{x}(z), \overline{x}(z)] \text{ where } \underline{x}(z) \text{ is the lower bound of the lowest shock, } \underline{z}, \text{ and } \overline{x}(z) \text{ is the upper bound of the highest shock, } \overline{\pi}. \text{ So any taxable activity interval associated with a shock } z \in Z \text{ is a subinterval of the maximal interval, } [\underline{x}(z), \overline{x}(z)] \subset [\underline{x}(z), \overline{\pi}].\)
with the definition of the aggregate capital stock in equation (14), the individual policy function \( h \) given implicitly by the operator Euler equation \( F(h; \pi) = 0 \), the distribution function, \( \lambda \), given implicitly by the operator equation \( L(h, \lambda; \pi) = 0 \), the endogenously determined bounds of taxable activity, \( \underline{x}(z) \) and \( \overline{x}(z) \) for all values of \( z \in Z \), and the free values of the government policy at the extreme lower and upper bounds, \( \pi(\underline{x}(z)) \) and \( \pi(\overline{x}(z)) \).

Note that since \( \underline{x} \) is endogenous, the endpoint \( \pi(\overline{x}) \) is equality constrained.

5 Necessary and Sufficient Conditions for the Optimal Government Policy Function

In this Section we derive the first-order necessary and second order sufficient conditions for the optimal government policy function. In order to derive these conditions in the calculus of variations, we need to specify the derivatives of the functionals \( \mathcal{W} \) and \( \mathcal{G} \) with respect to marginal changes in government policy, \( \pi \) and \( \pi' \). For this purpose, we use the concept of generalized derivatives on mappings between two Banach spaces (B-spaces), the Fréchet derivatives. The Fréchet derivative is a generalization of the concept of a derivative on functional and operator spaces (see Luenberger (1969) or Ok (2007)).

**Definition 6 (Fréchet Derivative)** Given a nonlinear operator \( N(u) \) on function \( u \), the Fréchet differential \( \delta N(u; \delta h) = N_u \delta h \) is

\[
\lim_{||\delta h|| \to 0} \frac{||N(u + \delta h) - N(u) - N_u \delta h||}{||\delta h||} = 0,
\]

where \( N_u \) is the Fréchet derivative.

Define the Lagrange function \( L \) for the Calculus of Variations Ramsey Problem in Definition 5, for each \( z \in Z = \{z, \pi\} \), as

\[
L(z, x) = \begin{cases} 
0 & \text{for } x \in [\underline{x}(z), \overline{x}(z)), \\
\mathcal{W}(z, x) + \mu \mathcal{G}(z, x) & \text{for } x \in [\underline{x}(z), \pi(z)], \\
0 & \text{for } x \in (\pi(z), \overline{x}(z)].
\end{cases} \quad (17)
\]

\(^{11}\)The compliance of the Fréchet derivatives (also called the F-derivatives) with the derivations of the first order conditions in the calculus of variations is reflected by the fact that the F-differential is identical to the variation. Our derivations are more complicated than the standard Fréchet derivative because our functional equations are recursive. Practically, the Fréchet derivative can be obtained using a weaker concept of the Gateaux derivative \( N_u = \lim_{\varepsilon \to 0} N_u^{\varepsilon}(u + \varepsilon \delta h) \) when the obtained derivative is continuous.
Note that the social welfare function in (13) is the sum of integrands $W(z, x) = W[z, x; K, \pi(x), \pi'(x)]$ integrated on intervals $[\xi(z), \tau(z)]$ for each $z \in Z$. The same is true for integrands $L(z, x)$ in equation (17).

Interestingly, as we show in Theorem 1 below, the relevant Lagrange function that emerges from the solution to the maximization problem is one amended by a term which captures the effect of the distribution of capital on social welfare: $	ilde{L}(z, x) = L(z, x) + \Psi K(z, x)$, where $\Psi$ is the marginal effect of the aggregate capital on social welfare. In other words, the term $\Psi K(z, x)$ takes into account the effect of prices (determined by the aggregate capital) on social welfare of agents characterized by $(z, x)$.

**Theorem 1 (First Order Necessary Conditions)** Using the modified Lagrange function $	ilde{L}$ for the Calculus of Variations Ramsey Problem in Definition 5,

$$
\tilde{L}(z, x) = \begin{cases} 
0 & \text{for } x \in [\xi(z), \tau(z)), \\
W(z, x) + \mu G(z, x) + \Psi K(z, x) & \text{for } x \in [\xi(z), \tau(z)], \\
0 & \text{for } x \in (\tau(z), \tau(z)].
\end{cases}
$$

for each $z \in Z = \{ \tilde{z}, \bar{z} \}$, the first order necessary conditions for the Ramsey problem are

1. the Euler-Lagrange condition,

$$
\sum_z \left( \tilde{L}_x(z, x) - \frac{d}{dx} \tilde{L}_{x'}(z, x) \right) = 0; \quad (19)
$$

2. the transversality condition on the free boundary value, $\pi(\tau(\bar{z}))$, at the equality constrained endpoint, $\tau(\bar{z})$,

$$
\left[ \tilde{L}(\tau(\bar{z}), \pi(x')) - \left( \pi'(x) - \frac{k_\pi(\tau(\bar{z}), x)}{\omega_\pi(\tau(\bar{z}), x)} \right) \tilde{L}_{x'}(\tau(\bar{z}), x) \right]_{x=\tau(\bar{z})} = 0; \quad (20)
$$

3. the transversality condition on the free boundary value $\pi(\xi(\bar{z}))$ at $\xi(\bar{z})$

$$
\left[ \tilde{L}_{x'}(\xi(\bar{z}), x) \right]_{x=\xi(\bar{z})} = 0; \quad (21)
$$

4. and the condition on the Lagrange multiplier, $\mu$, at which (16) is satisfied.

The marginal effect $\Psi$ of the aggregate capital stock on social welfare is

$$
\Psi \equiv \frac{\delta L}{\delta K} = \Psi_K \sum_z \left\{ \int_{\tilde{z}(z)}^{\pi(z)} L_K(z, x) dx + L_K(z, \xi(z)) + L_K(z, \tau(z)) \right\}, \quad (22)
$$

\[12\]Note that $\Psi$ is the effect of variation in $K$ on the variation in $L$, i.e. $\delta L(\pi, \pi', K)/\delta K$. Klein, Krusell, and Rios-Rull (2008) derive a similar effect of government policy on aggregate capital in their generalized Euler equation.
where

$$\Psi_K^{-1} = 1 - \sum_z \left\{ \int_{\tau(z)}^{\pi(z)} K(z, x) dx + K(z, \pi(z)) + K(z, \nu(z)) \right\},$$

(23)

with the marginal effects at the lower and upper bounds defined for \( I = \{L, K\} \) as

$$I_K(z, \underline{\pi}(z)) \equiv \left[ (I(z, x) - \pi'(x)I_{\pi'}(z, x)) x_K \right]_{x=\underline{\pi}(z)},$$

$$I_K(z, \overline{\pi}(z)) \equiv \left[ (I(z, x) - \pi'(x)I_{\pi'}(z, x)) (x_K + x_k \omega_K) \right]_{x=\overline{\pi}(z)}.$$

Proof For the proof and more detailed specifications of all terms see the Appendix.

Inspecting the first order conditions in Theorem 1, we see that the condition (19) is a functional equation in the unknown government policy function, \( \pi \), with the side conditions (20)-(21) and the condition on the value of the Lagrange multiplier, \( \mu \). From the setup of the problem it is clear that the only free boundary values of the government policies are the values at the lower and upper bounds \( \underline{\pi}(z) \) and \( \overline{\pi}(z) \).

Theorem 2 (Second Order Sufficient Conditions) A tax schedule \( \pi \) satisfying the first-order conditions in Theorem 1 attains a strict maximum if and only if (i) the Lagrange function defined in equation (17) satisfies the second-order Legendre condition

$$\sum_{z \in \mathcal{Z}} L_{\pi', \pi'}(z, x) < 0 \quad \text{for all} \quad x \in \left[ \underline{\pi}(z), \overline{\pi}(z) \right],$$

and (ii) the interval \( [\underline{\pi}(z), \overline{\pi}(z)] \) contains no points conjugate to \( \overline{\pi}(z) \).

Proof For the proof see the Appendix.

The Euler-Lagrange equation in Theorem 1 and the modified Legendre condition in Theorem 2 represent necessary and sufficient conditions for concavity and a unique maximum of the Calculus of Variations Ramsey Problem.

5.1 The Effects of Government Policy on the Equilibrium

If we knew how agents’ saving policies \( h \) and simultaneously how the distribution \( \lambda \) depend on the government policy schedule, i.e. if we could solve at equilibrium prices for the optimal policy \( \pi \) which is a function of the distribution and prices which in turn are determined by \( h(\pi, \pi') \) which is itself a function of the optimal policy and prices, the

---

13 Additionally, the more detailed first-order conditions in the Appendix contain savings and distribution functions, \( h \) and \( \lambda \), and their derivatives with respect to \( \pi \), \( \pi' \), and \( K \).
task of the derivation of the first order conditions for this dynamic optimization would be straightforward. However, not only we have to solve for these functions simultaneously but also we are in a much more difficult situation since for any government policy schedule \( \pi \), agents’ saving policy and the distribution functions are known only implicitly as a solution to the two operator equations \((F(h; \pi) = 0 \text{ and } L(h, \lambda; \pi) = 0)\) and the aggregate equilibrium condition for stationary prices.

The next Lemma derives the effects of the government policy function \( \pi \) on the operator Euler equation by specifying five unknown “sensitivity” functions \( h_{\pi} : Z \times K \rightarrow \mathbb{R}^+ \), \( h_{\pi'} : Z \times K \rightarrow \mathbb{R}^+ \), \( h_{K} : Z \times K \rightarrow \mathbb{R}^+ \), \( h_{\pi', \pi} : Z \times K \rightarrow \mathbb{R}^+ \), and \( h_{\pi''} : Z \times K \rightarrow \mathbb{R}^+ \).

**Lemma 1 (The Effects of \( \pi, \pi', \text{ and } K \) on the Euler Equation)** The total \( F \)-derivatives of the operator Euler equation \( F \) with respect to the government policy function \( \pi \), to its derivative \( \pi' \), and to the aggregate capital stock \( K \), are,

\[
F_i = u''(c_i) - \beta \sum_{z^+} \left\{ u''(c^+)_i R^+ + u'(c^+) R^+_i \right\} Q(z, z^+) = 0, \tag{24}
\]

where the subscript \( F_i \) denotes derivatives with respect to \( i \in \{\pi, \pi', K\} \), and

\[
F_{ij} = u'''(c) c_i c_j + u''(c) c_{ij} - \beta \sum_{z^+} \left\{ [u'''(c^+) c^+_i c^+_j + u''(c^+) c^+_i R^+_j + u''(c^+) c^+_j R^+_i] + u'(c^+) R^{+}_{ij} \right\} Q(z, z^+) = 0. \tag{25}
\]

where the subscripts \( F_{ij} \) denote derivatives with respect to \( ij \in \{\pi'\pi', \pi'\pi\} \).

**Proof** For the proof and a full definition of terms see the Appendix.

Similarly, we derive the effects of the government policy on the shape of the distribution function \( \lambda \) by specifying functional equations which implicitly determine the unknown “sensitivity” functions \( \lambda_{\pi} : Z \times K \rightarrow \mathbb{R}^+ \), \( \lambda_{\pi'} : Z \times K \rightarrow \mathbb{R}^+ \), \( \lambda_K : Z \times K \rightarrow \mathbb{R}^+ \), \( \lambda_{\pi'\pi} : Z \times K \rightarrow \mathbb{R}^+ \), and \( \lambda_{\pi'\pi'} : Z \times K \rightarrow \mathbb{R}^+ \).

**Lemma 2 (The Effects of \( \pi, \pi', \text{ and } K \) on the Stationary Distribution Function)** The total \( F \)-derivative of the operator stationary distribution function \( L \) with respect to the government policy function \( \pi \), to its derivative \( \pi' \), and to the aggregate capital stock \( K \) are,

\[
L_i = \lambda_i(z^+, k^+) - \sum_z \left\{ \lambda_k \left[ z, h^{-1}(z, k^+) \right] h^{-1}_{i}(z, k^+) + \lambda_i \left[ z, h^{-1}(z, k^+) \right] \right\} Q(z, z^+) = 0, \tag{26}
\]

14 The proper specification of the derivative of the savings and distribution functions with respect to the government policy function again requires the generalized concept of the Fréchet derivative.
where the subscript $L_i$ denotes derivatives with respect to $i \in \{\pi, \pi', K\}$, and

\[ L_{ij} = \lambda_{ij}(z^+, k^+) - \sum_z \left\{ \lambda_{ij} \left[ z, h^{-1}(z, k^+) \right] + \lambda_{ik} \left[ z, h^{-1}(z, k^+) \right] h_j^{-1}(z, k^+) \\
+ \left[ \lambda_{jk} \left[ z, h^{-1}(z, k^+) \right] + \lambda_{kk} \left[ z, h^{-1}(z, k^+) \right] h_j^{-1}(z, k^+) \right] h_i^{-1}(z, k^+) \\
+ \lambda_{k} \left[ z, h^{-1}(z, k^+) \right] h_{ij}^{-1}(z, k^+) \right\} Q(z, z^+) = 0. \tag{27} \]

where the subscripts $L_{ij}$ denote derivatives with respect to $ij \in \{\pi', \pi', \pi'\}$.

**Proof** For the proof and the definition of terms see the Appendix.

In this way we obtain functional equations (10), (24)-(25), (11), and (26)-(27) in the unknown functions $h$, $h_\pi$, $h_{\pi'}$, $h_K$, $h_{\pi'\pi}$, $h_{\pi'\pi'}$, $\lambda$, $\lambda_\pi$, $\lambda_{\pi'}$, $\lambda_K$, $\lambda_{\pi'\pi'}$, and $\lambda_{\pi'\pi}$, respectively.

Finally, by adding the first-order conditions from Theorem 1, the problem of finding the optimal government policy $\pi$ is a system of thirteen functional equations in thirteen unknown functions with two side conditions and one condition on the Lagrange multiplier.

### 5.2 Analytical Results

Since the first-order conditions in Theorem 1 specify the explicit transversality conditions on agents with the lowest and the highest taxable activity, they can be used for qualitative results on the behavior of the policy schedule at these boundary points.

Despite the complexity of the problem we are able to derive several analytical results. It will be useful to compare the optimal government policy and its effects to those of a flat-tax economy in which the government consumes the same fraction $g$ of the total output. Note that the term\(^{15}\) $gy(z, x)$ in equation (16) captures the fraction of government expenditures “related” to an agent $(z, x)$. This fraction of government expenditures $gy(z, x)$ is identical to the amount of taxes hypothetically paid by the agent $(z, x)$ under a flat tax $g$ on total income. We will further use this concept as a useful benchmark in the following analysis.

First, we relate the tax payment of the “poorest” agent under the optimal tax policy to the tax payment in the economy with a flat tax $g$. This poorest agent is constrained at the lowest value of capital $k$ and is hit by the lowest productivity shock $\tilde{z}$.

**Proposition 1 (Optimal Tax at Lowest Taxable Activity)** The transversality condition in equation (21) in Theorem 1 implies that for an agent with the lowest taxable activity level $z(\tilde{z})$, the difference between optimal taxes paid $\pi(z(\tilde{z}))\tilde{z}(\tilde{z})$ and taxes paid under a flat tax regime, $gy(z(\tilde{z}), \tilde{z}(\tilde{z}))$, is proportional to the sum of the agent’s utility and the marginal contribution of his savings to the aggregate welfare,

\[ \pi(z(\tilde{z}))\tilde{z}(\tilde{z}) - gy(z(\tilde{z}), \tilde{z}(\tilde{z})) = -\frac{W[u(c(z, z(\tilde{z}))]}{\mu} + \Psi_k, \]

\(^{15}\)To shorten the notation we write $gy(z, x) = gy(z, k(z, x))$. 

17
where \( c(z; x(z)) = y(z; x(z)) - \pi(x(z))z - k \) is consumption, \( y(z; x(z)) = rk + wz \) is before-tax income, \( \mu < 0 \) is the shadow price of government expenditures, and \( \Psi \) is the marginal effect of aggregate capital on aggregate welfare defined in Theorem 1.

**Proof** See the Appendix.

The following Corollary states the conditions for which the amount of taxes paid by the poorest agent under the optimal tax policy is larger than that under a flat income tax, i.e. \( \pi(x(z))z > gy(z; x(z)) \).

**Corollary 1** If the savings contribution of the poorest agent to aggregate welfare is non-negative, \( \Psi k \geq 0 \), then at the lowest taxable activity level \( x(z) \) the amount of taxes paid under the optimal tax policy \( \pi(x(z))z \) is larger than that under a flat income tax, \( gy(z; x(z)) \).

The poorest agent pays more taxes than in the flat-tax regime if one of these three conditions is satisfied: \( \Psi = 0 \), or if \( k = 0 \), or if simultaneously \( \Psi > 0 \) and \( k > 0 \). The Corollary illustrates the incentives to the agent to accumulate more assets.\(^\text{16}\) We discuss these incentives in the following Section.

In a similar way, the Proposition below specifies the optimal tax payment of the “richest” agent relative to his tax burden under a flat tax equal to \( g \). The richest agent’s current and future savings are at the endogenous upper bound \( \tilde{k}(z) = \bar{k} \).

**Proposition 2 (Optimal Tax at Highest Taxable Activity)** Assuming \( \pi'(x(z)) \neq \left[ \frac{\delta x}{\delta x}(x(z)) \right]^{-1} \), the transversality condition in equation (20) in Theorem 1 implies that for an agent with the highest taxable activity \( x(z) \), the difference between optimal taxes paid \( \pi(x(z))z \) and the taxes paid under the flat tax regime \( gy(z; x(z)) \), is proportional to the sum of the agent’s utility and the marginal contribution of his savings to aggregate welfare,

\[
\pi(x(z))z - gy(z; x(z)) = -\frac{W[u(r(z, x(z)))] + \Psi\tilde{k}}{\mu},
\]

where \( r(z, x(z)) = y(z, x(z)) - \pi(x(z))z - \bar{k}(z) \) is consumption, \( y(z, x(z)) = rk + wz \) is before-tax income, \( \mu < 0 \) is the shadow price of government expenditures, and \( \Psi \) is the marginal effect of aggregate capital stock on aggregate welfare defined in Theorem 1.

**Proof** See the Appendix.

Notice that as \( \omega = \frac{k}{1 - \rho_k} < 0, \left[ \frac{\delta x}{\delta x}(x(z)) \right]^{-1} = \frac{1}{x_{x}(x(z))\omega_k(z, \bar{k})} < 0 \) implies that the assumption on the slope of the tax schedule at \( x(z) \) is satisfied whenever it is non-decreasing, i.e. \( \pi'(x(z)) \geq 0 \), and that \( \Psi \geq 0 \).

\(^\text{16}\)If a substantial debt can be accumulated, \( \bar{k} < 0 \), the ability of the government to tax capital stock becomes limited. Finally, it is easy to show that the equilibrium might not exist if \( \Psi < 0 \).
The following Corollary states the conditions for which the amount of taxes paid by
the richest agent under the optimal tax schedule is larger than that under a flat tax, i.e.
\( \pi(\overline{x}(z))\overline{x}(z) > g\overline{y}(z, \overline{x}(z)). \)

**Corollary 2** If the marginal contribution of the aggregate capital stock to aggregate welfare
is nonnegative, \( \Psi \geq 0 \), then at the highest taxable activity level the optimal tax contribution
\( \pi(\overline{x}(z))\overline{x}(z) \) is larger than that under a flat tax regime \( g\overline{y}(z, \overline{x}(z)) \),
\[
\pi(\overline{x}(z))\overline{x}(z) - g\overline{y}(z, \overline{x}(z)) = -\frac{W(u(\overline{x}(z, \overline{x}(z))))}{\mu} + \frac{\Psi\overline{K}}{W(u(\overline{x}(z, \overline{x}(z)))) + 1} > 0.
\]
Additionally,
\[
\frac{\pi(\overline{x}(z))\overline{x}(z) - g\overline{y}(z, \overline{x}(z))}{\pi(\overline{x}(z))\overline{x}(z) - g\overline{y}(z, \overline{x}(z))} \geq \frac{W(u(\overline{x}(z, \overline{x}(z))))}{W(u(\overline{x}(z, \overline{x}(z)))) + 1} \left( \frac{\Psi\overline{K}}{W(u(\overline{x}(z, \overline{x}(z)))) + 1} \right)
\]
implies that the amount of taxes paid by the richest agent is greater than that of the poorest
agent, i.e.
\( \pi(\overline{x}(z))\overline{x}(z) > \pi(\overline{x}(z))\overline{x}(z). \)

These Corollaries show that the socially optimal tax payments of the ‘boundary’ agents
depend only on these agents’ individual characteristics and the two aggregate shadow
prices: the shadow price of government spending, \( \mu \), and the shadow price of the aggregate
capital, \( \Psi \), both expressed in terms of social welfare.

Together, Corollaries 1 and 2 reveal a lot about the shape of the optimal government
policy schedule. If \( k = 0 \), then both ends of the tax schedule are above the flat-tax level
and the implied tax schedule is a “U-shape” function. Its trough lies under the flat-tax
level to clear the government budget constraint proportional to \( g \). This is also the case if
\( \Psi = 0 \) and if both \( \Psi > 0 \) and \( k > 0 \). Finally, if \( \Psi > 0 \) and \( k < 0 \), the optimal tax schedule
could either be “U-shaped” or an increasing function of income.

The marginal effect \( \Psi \) of the aggregate capital stock on social welfare reflects the
importance of the optimal wealth distribution in analyzing tax policies. The shape of the
distribution of the individual state variable is behind the U-shape marginal tax rates in
Diamond (1998) and Saez (2001) as well as in the dynamic models of Kapicka (2013) and
Golosov, Tsyvinski, and Werquin (2014). These qualitative assessments are confirmed by
the numerical results in our example in the following Section.

6 An Example: The Optimal Income Tax Schedule

In this section we demonstrate our method by finding the optimal government policy \( \pi \)
defined as a tax schedule on total income from capital and labor. Therefore, the taxable
activity is
\[ x(z, k; K) = y(z, k; K) = rk + wz, \]
and the individual budget constraint is
\[ c + k^+ \leq (1 - \pi(x))x + k. \]

We assume that the borrowing constraint is \( k = 0 \) and that there are only two shocks, \( Z = \{ z, \bar{z} \} \). The total tax revenues are equal to a fraction \( g \) of the total output.

Thus the Euler equation (9) for a \((z, k)\)-agent’s optimal savings function \( k^+(z, k) \) for all \((z, k) \in Z \times B\) is now
\[ u'(c) \geq \beta \sum_{z^+} u'(c^+) \left[ (1 - \pi(x^+) - \pi'(x^+)x^+) r + 1 \right] Q(z, z^+), \]
where \( c^+ = (1 - \pi(x^+))x^++k^++k^{++}, x^+ = rk^++wz^+, \) and \( k^{++} = k^+(z, k, z^+) \).
Note that for this specification \( x_k = y_k = r \) and \( k_x = 1/x_k = 1/r \).

For this concrete example we first analyze the conditions for the existence of the stationary recursive competitive equilibrium.

### 6.1 Existence and Uniqueness of Stationary Recursive Competitive Equilibrium

Because the tax schedule is an arbitrary function, we must ensure that the first order approach is valid.\(^{17}\) 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). For each agent \((z, k) \in B \times Z\), denote the after-tax gross income as
\[ \psi(z, k) \equiv (1 - \pi(x(z, k))) x(z, k) + k. \]

Using \( \psi(z, k) \), rewrite the Euler equation as
\[ u'(\psi(z, k) - k^+(z, k)) = \beta \sum_{z^+} u'(\psi(k^+(z, k), z^+) - k^+(k^+(z, k), z^+))\psi_1(k^+(z, k), z^+)Q(z, z^+), \]
where
\[ \psi_1(k^+(z, k), z^+) = (1 - \pi(x(k^+(z, k), z^+)) - \pi'(x(k^+(z, k), z^+)) x(k^+(z, k), z^+)) r + 1 \]
is the marginal after-tax return of investment. In the following Theorem we establish the validity of the first order approach and the existence of the competitive equilibrium.

**Theorem 3** *For a given tax schedule \( \pi : \mathbb{R}_+ \rightarrow \mathbb{R} \), if for each \((z, k) \in B \times Z\)*

\(^{17}\)Again, we analyze the case of borrowing constrained agents in the Appendix.
1. $\psi_1(z, k) > 0$, and
2. $\psi$ is quasi-concave,
then the solution to each agent’s maximization problem and the stationary recursive competitive equilibrium exist.

**Proof** See the Appendix.

The following corollary characterizes the set of admissible tax schedules that satisfy the conditions of Theorem 3. We have earlier introduced the notation for the endogenous upper bound on capital for any agent, $\bar{k}$. Let $\{\underline{w}, \bar{w}\}$ and $\{\underline{r}, \bar{r}\}$ denote some arbitrary, non-binding lower and upper bounds for equilibrium wage and interest rate, respectively. Finally, $\varepsilon_\pi^x(x) \equiv \frac{\pi'(x)}{\pi(x)} x$ is the elasticity of the tax rate to the taxable income.

**Corollary 3 (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 $\pi \in C^2(\mathbb{R}_+)$ belongs to the set of admissible tax schedules $\Upsilon$,

$$\Upsilon = \left\{ \pi \in C^2(\mathbb{R}_+) : \pi(x) \left(1 + \varepsilon_\pi^x(x)\right) < 1 + \frac{1}{r} \right\}$$

for all $x \in [\underline{r}k + \underline{w}z, \bar{r}\bar{k} + \bar{w}\bar{z}]$, then it satisfies the conditions of Theorem 3.

The above statement follows directly from the fact that $\psi_1(z, k) > 0$ and that $\psi$ is quasi-concave. The corollary implies that the marginal tax rate must be smaller than $1 + 1/r$. This upper bound is not likely to bind for a very wide range of tax schedules.\(^{18}\)

Application of Theorem 1 and Theorem 2 imply necessary and sufficient conditions for the unique maximum attained by the tax schedule.

### 6.2 The Shape of the Optimal Income Tax Schedule

To obtain the first order conditions for the optimal income schedule we insert the terms from our example on total-income tax into the general conditions in Theorem 1 and Lemma 1 and Lemma 2 (see the Appendix C). Then, adapting Propositions 1 and 2 and the related Corollaries 1 and 2 to our example economy we obtain the following analytical results.

**Corollary 4 (Optimal Tax Rates at Lowest and Highest Income)** If $\pi$ is the optimal tax schedule on the total income, then

---

\(^{18}\)As an example, for a realistic equilibrium interest rate $r = 0.05$, the upper bound on the marginal tax rate is equal $1 + 1/r \approx 21$. Therefore, for a high level of tax rate, say $\pi(x) = 0.5$, the elasticity of the tax schedule at that level of income would have to be $\varepsilon_\pi^x = 41$ in order to violate the admissibility condition. When numerically solving for the optimal tax schedule in the next Section we do not impose any of these exogenous bounds but we check the admissibility of the optimal tax schedule ex post.
1. the optimal tax rate at the lowest income level \( \bar{\pi}(z) \) is strictly greater than a flat tax \( g \) at that income level,

\[
\pi(\bar{\pi}(z)) > g;
\]

2. and, provided that \( \Psi \geq 0 \), the optimal tax rate at the highest income level \( \bar{\pi}(\overline{z}) \) is strictly greater than the optimal tax rate at the lowest income level,

\[
\pi(\bar{\pi}(\overline{z})) > \pi(\bar{\pi}(z)) > g.
\]

**Proof** The results follow directly from Corollaries 1 and 2.

Thus both the poorest and richest agents face higher taxe rates than in a corresponding flat-tax economy. The government budget constraint and the continuity of the optimal tax function imply that there must exist a measure of agents at intermediate income levels who face lower tax rates than the flat tax rate. Thus the optimal tax schedule must be a “U-shape” function.

We interpret these results from the point of view of a social planner who could use the first best, lump sum taxes. The social planner would set such a tax schedule on each pair of individual states \((z, k)\) to arrive at a stationary equilibrium in which all agents would accumulate the same amount of capital \( K \) and consume the same amount of goods. In other words, the distribution would consist of a mass of agents at two points, \((z, K)\) and \((\overline{z}, K)\). A U-shape transfer (tax) system is the way to induce agents to arrive at this optimal outcome.

Although our government is constrained from using the first best policy, it strives to accomplish the same outcome. By imposing higher than average taxes on the poorest and richest agents, the government tries to provide incentives for agents to save towards the average capital stock. The poorest agents are motivated by high but decreasing tax schedule, \( \frac{d(\pi(x,k))}{dx} \bar{\pi}(z) < 0 \). Thus, according to the Euler equation, their expected return on capital increases. On the other hand, the richest agents are discouraged from further savings by the increasing tax schedule at high levels of capital. Finally, agents with savings around the average capital stock are motivated by lower than average taxes to keep their savings at the current level. The U-shape tax function provides the incentives for agents to eventually arrive and stay at the individual capital stock around \( k = K \). Similar incentives for agents in dynamic models of optimal taxation are present in Kapicka (2013), Golosov, Tsyvinski, and Troshkin (2010), or Golosov, Tsyvinski, and Werquin (2014). We illustrate the optimal policy by numerical simulations in the following Section.
7 Numerical Solution

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. In order to evaluate the welfare implications of these tax reforms, we conduct a transition analysis.

The solution to the problem of finding the optimal policy in terms of the Operator Steady State Ramsey Problem in our example leads, as in the general case, to solving the functional system of the first-order conditions given by Theorem 1 together with the functional equations specifying the stationary equilibrium (expressed by \( h \) and \( \lambda \)) and the sensitivity functions given by the F-derivatives. We obtain a system of functional equations with two side conditions and one condition on the Lagrange multiplier. We solve this complex functional equations problem by the least squares projection method. Its application to our problem, all definitions of the functional equations, and the approximation of the optimal tax schedule can be found in Appendix D.

7.1 Parameterization

As empirical studies find a significant degree of persistence in the idiosyncratic shocks to labor productivity (see Storesletten et al. (2004)), 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 by a first-order autoregression of the form

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

with \( \epsilon \sim N(0, \sigma^2_\epsilon) \). They find that \( \rho = 0.53 \) and \( \sigma^2_\epsilon = 0.063 \). 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.\(^{20}\) We set the discount factor at \( \beta = 0.95 \). The rest of the parameters is standard, \( \alpha = 0.36 \), \( \delta = 0.1 \), and the preference parameter \( \sigma = 1 \).

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

\(^{19}\)The least squares projection method is an efficient and well-behaved method for functional equations problems. For a more detailed explanation of the use of the projection methods to stationary equilibria in economies with a continuum of heterogenous agents see Bohacek and Kejak (2002).

\(^{20}\)Similar parameterization is in Storesletten, Telmer, and Yaron (2004) 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 \( Q(z_L, z_L) = 0.981 \), \( Q(z_H, z_H) = 0.926 \).
7.2 The U.S. Progressive Tax Schedule

We model the progressive tax schedule as Ventura (1999) model of a flat-tax reform with heterogeneous agents. \(^{21}\) An agent’s budget constraint is

\[ c + k' \leq rk + zw + k - T, \]

where \( T \) represents the amount of taxes paid by an agent according to the progressive tax schedule applied to the total taxable income, \( I = rk + \max\{0, zw - I^*\} \), with a labor-income tax deductible amount \( I^* \geq 0 \). There are \( M \) tax 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 the 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 a flat rate \( \tau_k = 0.25 \).

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

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

The government budget constraint is cleared by finding an equilibrium value of the tax exemption level \( I^* \). Aggregate statistics of the steady state are shown in the left column of Table 1.

7.3 A Flat-Tax Reform

The flat-tax reform consists of replacing the progressive tax schedule with a single flat tax \( \tau \) 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 \( \tau = 0.254 \).

Relative to the progressive tax schedule steady state, the flat-tax reform in the middle column of Table 1 increases the steady state levels by similar 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. \(^{22}\)

\(^{21}\) Compared to Ventura (1999), 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.

\(^{22}\) 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%.

24
Table 1: Steady State Results.

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Progressive</th>
<th>Flat Tax</th>
<th>Optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital</td>
<td>2.54</td>
<td>3.29</td>
<td>3.80</td>
<td></td>
</tr>
<tr>
<td>Output</td>
<td>1.40</td>
<td>1.54</td>
<td>1.62</td>
<td></td>
</tr>
<tr>
<td>Consumption</td>
<td>0.86</td>
<td>0.90</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Median wealth</td>
<td>2.45</td>
<td>2.55</td>
<td>3.60</td>
<td></td>
</tr>
<tr>
<td>Median/mean wealth ratio</td>
<td>0.97</td>
<td>0.77</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>Interest rate (%)</td>
<td>9.87</td>
<td>6.77</td>
<td>5.33</td>
<td></td>
</tr>
<tr>
<td>Wage</td>
<td>0.894</td>
<td>0.983</td>
<td>1.035</td>
<td></td>
</tr>
<tr>
<td>Gini total income</td>
<td>0.22</td>
<td>0.31</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Gini total after tax income</td>
<td>0.21</td>
<td>0.32</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Constrained agents (%)</td>
<td>1.16</td>
<td>1.88</td>
<td>1.42</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Constrained agents is a fraction of agents whose wealth equals the exogenous lower bound on capital.

7.4 The Optimal Tax Schedule

We use our methodology described in the previous Sections to solve for the optimal tax schedule that maximizes average steady state welfare. The right column in Table 1 summarizes the optimal tax schedule steady state. The impact of the optimal tax schedule is very large. Steady state average welfare increases by 4.4%. Aggregate capital stock rises by 49%, output by 15.8%, and consumption by 5.8%. 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%.

Figure 1 shows the optimal tax schedule and the optimal marginal tax rate function. The average tax rate is a U-shaped function taxing the lowest total income at 45%, decreasing to a minimum of 19% and rising to 62% at the highest level of total income. Although the whole shape of the tax function is important for savings incentives and the resulting equilibrium allocations, a majority of agents face the decreasing or the flat part of the tax schedule. The marginal tax rate is also U-shaped, almost flat and close to zero for low incomes, falling to negative levels around the average total income and then rising at high income levels. Note that the maximal marginal tax rate is 2.5 and that the optimal tax schedule easily satisfies the admissibility condition from Corollary 3 (the equilibrium
The Optimal Tax Schedule

The Optimal Marginal Tax Rate

Figure 1: The optimal tax schedule and the optimal marginal tax rate.

interest rate implies an upper bound equal to 19.7). Optimal marginal tax function with a similar shape is also found in dynamic models of Kapicka (2013) or Golosov, Tsyvinski, and Troshkin (2010) and in static models by Diamond (1998) and Saez (2001).\textsuperscript{23}

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. Conesa and Krueger (2006), also in a general equilibrium framework but with added life cycle features, searched for the optimal progressivity of a tax schedule, limiting their class

\textsuperscript{23}In the original contribution of Mirrlees (1971), the welfare maximizing tax schedule is close to a linear non-decreasing function with the marginal tax rate between zero and one, and zero at both ends of the distribution due to labor incentives related to the distribution of skills and consumption-leisure preferences. Building on Mirrlees (1971) and Mirrlees (1976) seminal work, Kocherlakota (2005), Golosov, Kocherlakota, and Tsyvinski (2003) or Albanesi and Sleet (2006) study optimal social planner policies with asymmetric information. In these models, informational frictions are necessary for characterization of optimal policies.
of tax schedules to monotone functions as in Gouveia and Strauss (1994). In this class of functions, the optimal tax schedule is basically a flat tax with a fixed deduction, delivering a welfare gain of 1.7% compared to the existing progressive tax in the United States. The class of monotone functions seems rather restrictive for the optimal tax schedule. Our class of admissible functions includes all progressive tax schedules but these were found significantly inferior with respect to the welfare criterion.

7.5 The Tradeoff Between Efficiency and Equality

Apart from the general equilibrium effects, the huge welfare impact of the optimal tax schedule arises from the distributional effects. The stationary distributions of agents over assets in the three steady states are shown in Figure 2. Although both the flat and the optimal tax schedules increase the aggregate levels, the difference between them is that the ad hoc 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 0.77. 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 steady state 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 behind the large increase in the aggregate levels are the incentive effects provided by the shape of the optimal tax schedule. The U-shaped function in the top panel of Figure 1 effectively motivates the agents to concentrate 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 provides tax revenues for redistribution. In the middle of the total income levels, the tax rate is lower than that found in the flat-tax reform. The optimal tax schedule preserves the median/mean wealth ratio of the progressive tax schedule by increasing the median by 47% and the mean by 49%. The support of the invariant distribution becomes wider but inequality measures do

\[24\text{The simple search method cannot be used for computation of a stationary competitive equilibrium which is the limit of the optimal dynamic tax schedule. We show in Bohacek and Kejak (2004) that under some parameterization the first order conditions of the general dynamic Ramsey problem can simplify to the first order conditions of the steady state Ramsey problem analyzed in this paper.}\]
Figure 2: Stationary distribution of agents over assets in the optimal, flat, and progressive tax schedule steady states.

not increase as much as in the flat-tax reform.\textsuperscript{25}

To further analyze the tradeoff between efficiency and distribution, we adopt the approach in Domeij and Heathcote (2004) who distinguish the efficiency gain from distributional gains. The efficiency gain for an individual agent is the percentage of the original consumption that would allow the agent to consume the same fraction of the aggregate consumption after the reform as in the original steady state. The distributional gain is the difference between the usual individual welfare gain and the efficiency gain.\textsuperscript{26}

The top panel of Table 2 displays the average efficiency and distributional gains in the

\textsuperscript{25}Table 1 also shows the fraction of agents constrained in their borrowing: only 1.16\% of agents is constrained in the progressive tax schedule steady state. The flat-tax schedule increases this number to 1.88\%, while the optimal tax steady state it is 1.42 (Domeij and Heathcote (2004) obtained similar results).

\textsuperscript{26}The individual welfare gain is the percentage of the original consumption level that would make an agent as well off as in the optimal tax steady state. In the case of logarithmic utility, the gain is the same for all agents (see Domeij and Heathcote (2004) for a simple proof and other details).
optimal steady state relative to the other two steady states. It is apparent that the steady state associated with the optimal tax function is welfare and efficiency superior to the other two steady states: both average welfare and efficiency measures are positive, and greater for the comparison with the progressive steady state. As Figure 2 shows, the optimal tax schedule obtains an average distributional loss relative to the progressive tax (−0.57%) but a gain relative to the flat-tax steady state (2.86%).

In Appendix D we present a detailed analysis of individual welfare and distributional gains. In terms of the income and after-tax income distribution, the progressive tax helps the bottom quintile while the flat tax helps the top quintile. The optimal tax equalizes tax contribution shares across all quintiles. Both the flat-tax and progressive-tax steady states put more relative burden on the higher income quintiles. The large welfare gains for poor agents are mostly due to the general equilibrium effects, namely the higher wage in the optimal-tax steady state.

### 7.6 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. In Domeij and Heathcote (2004), the expected discounted present value of welfare losses during transition are so large that they overturn the steady state welfare improvement.\(^{27}\)

\(^{27}\)This is similar to Garcia-Mila, Marce, 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...
The bottom part of Table 2 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.\textsuperscript{28} 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 the reform makes both the mean and the median agents in the progressive-tax economy better off. The welfare gains are positive but smaller than in the steady-state comparison (3.44\% and 3.86\%, respectively, measured in terms of per-period consumption transfers as a percentage of the initial steady state average consumption). In total, 73\% of the population is better off from the reform.

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 1.81\% and 1.97\%, respectively). The poor and also the wealthy agents, who now face much higher taxes, are worse off during the transition. Political support is not sufficient, equal only to 33\% of the population. We do not know whether an optimal transition would be welfare improving from this steady state.

As usual, this transition exercise shows that a tax reform is not Pareto improving for all agents. However, the efficiency gains from the optimal tax reform applied to the progressive-tax economy are so large that they not only provide resources for potential compensations but also obtain political support from the majority of agents despite the transitional costs (see Appendix D for more details).\textsuperscript{29}

\section{Conclusions}

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

\begin{footnotesize}
\begin{itemize}
\item 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.
\item Of course, it is not the optimal transition to the optimal tax schedule steady state. An optimal reform would implement a time-specific optimal tax schedule at each period of the transition.
\item Conesa and Krueger (2006) also find that a majority of the population would benefit from their tax reform. However, in their case the poor and rich benefit, while it is the middle class (38\%) who would be against the reform.
\end{itemize}
\end{footnotesize}
high aggregate levels and welfare improvements can be attained.

To our knowledge, this paper is the first one that provides a solution method for such unconstrained optimal government policies in an economy with heterogeneous agents and general equilibrium. 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. As in Diamond (1998), Saez (2001), or Golosov, Tsyvinski, and Troshkin (2010), the optimal tax schedule is U-shaped, it increases all aggregate levels by providing the right incentives but not at the cost of increased inequality. Welfare gains in the steady state are large, positive for both the mean and the median agent. Moreover, these gains persist also in a transition following an unanticipated optimal tax reform of the progressive tax schedule steady state.

The approach 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 endogenous labor supply and realistic life-cycle income profiles as in Golosov, Tsyvinski, and Troshkin (2009) and Farhi and Werning (2012). We would also like to explore different (Rawlsian) welfare functions. 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 and Conesa, Kitao, and Krueger (2009)). Finally, we plan to use this methodology to analyze optimal dynamic taxation.

References


Kapicka, M. (2013). Efficient allocations in dynamic private information economies with


