ADEMU
A Dynamic Economic and Monetary Union

ADEMU WORKING PAPER SERIES

Optimal Government Policies in Models with Heterogeneous Agents

Radim Bohácek[†]

Michal Kejak[Ť]

April 2018

WP 2018/130

www.ademu-project.eu/publications/working-papers

Abstract

In this paper we develop a new approach for funding 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 e_ect of the government policy on individual allocations, the resulting optimal distribution of agents in the steady state and, therefore, equilibrium prices. We illustrate the methodology on a Ramsey problem with heterogeneous agents, finding the optimal limiting tax on total income.

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

[†] CERGE-EI,Email: radim.bohacek@cerge-ei.cz
[Ť] CERGE-EI,Email: michal.kejak@cerge-ei.cz

Horizon 2020
The Framework Programme
for Research and Innovation

Acknowledgments

First version: September 2002. For helpful comments we thank the editor and the referees, Jim Costain, Mark Gertler, Max Gillman, Christian Hellwig, Boyan Jovanovic, Marek Kapicka, 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 Zubricky and Tomas Lesko provided excellent research assistance. The authors would like to acknowledge the support of the ADEMU project, "A Dynamic Economic and Monetary Union, funded by the European Union Horizon 2020 Program under grant agreement N 649396 (ADEMU)". The financial support of the Czech Science Foundation project No. P402/12/G097 (DYME Dynamic Models In Economics) is also gratefully acknowledged. All errors are our own.

The ADEMU Working Paper Series is being supported by the European Commission Horizon 2020 European Union funding for Research & Innovation, grant agreement No 649396.

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 condition for 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 an Euler-Lagrange equation (with transversality conditions) form 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.

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. There are two restrictions we impose on the solution: the government cannot use taxes that are state-contingent (to preserve incomplete markets and heterogeneity in the economy) and, because of the variational approach, the tax function must belong to the class of continuously differentiable functions. We do not impose additional assumptions on the shape of the government policy function. The optimal policy is 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.

We formulate the government problem as a modified Golden Rule. That is, we solve for an optimal limiting government policy under an assumption that the economy converges to a steady state. The optimal limiting government policy is a long-run optimal outcome that takes into account intertemporal discounting and the convergence to the steady state. In a related paper, Davila, Hong, Krusell, and Rios-Rull (2012) consider a social planner that attains a constrained optimum by directly manipulating the savings decision of each agent. They derive a functional first-order necessary condition with an added pecuniary externality arising from general equilibrium effects and use the variational approach for its characterization. Compared with Davila et al. (2012), the contribution of our paper is in the formulation of the Euler-Lagrange equations and the joint consideration of general equilibrium and distributional effects of the optimal limiting tax policy function.

We illustrate this methodology on a Ramsey problem, solving for the optimal limiting tax schedule on total income that maximizes average welfare in a steady state of a standard neoclassical, dynamic general equilibrium model with heterogeneous agents and incomplete markets calibrated as in Davila et al. (2012). For this calibration with a realistic wage and wealth inequality, we compare steady state allocations corresponding to the optimal limiting tax schedule with those related to a progressive tax schedule approximated by Heathcote et al. (2016) and to a flat-tax reform.

The optimal limiting average tax schedule is a U-shaped function. The marginal tax rates are also U-shaped, balancing a trade-off between the general equilibrium and the distributional effect. The former effect arises from providing incentives to accumulate a higher stock of aggregate capital that increases productivity of labor and, therefore, the income of poor agents (the average price effect in Davila et al. (2012)). The latter effect redistributes resources across agents in the stationary equilibrium. The marginal tax rates at low incomes induce agents to save more in order to escape relative poverty and secure better insurance against idiosyncratic risk, while the high tax rates on wealthy agents provide resources for short-run redistribution.

Without agent-specific lump-sum transfers the optimal limiting tax schedule cannot attain the constrained optimum in Davila et al. (2012). The optimal limiting tax function only slightly increases the aggregate capital stock but significantly reduces inequality. We compute transitions to the optimal limiting tax steady state and find that only when a high initial capital stock can be consumed during a transition, the tax reform improves welfare of the majority of the population. Following Farhi and Werning (2007), Farhi et al. (2012), and Krueger and Ludwig (2018), we analyze the effects of the optimal limiting tax schedule when the government puts different weights on current redistribution compared to long-term equilibrium effects. Finally, we perform several sensitivity tests and discuss important differences and effects from an alternative parameterization based on Aiyagari (1994). The U-shape of the optimal limiting average tax schedule is obtained in most of the simulated economies.

Our work contributes to recent advances in the literature on dynamic optimal taxation. Throughout the paper we compare our methodology and results to Davila et al. (2012) and show the equivalence of our Euler-Lagrange equation to the first-order condition of their constrained efficiency problem. Several papers have also built on this seminar work: Park (2017) adds an important dimension by introducing human capital. The first-order condition with respect to human capital investment also has an extra term whose sign is opposite to that of savings. Because the planner can improve welfare only by altering equilibrium prices, qualitative results of Davila et al. (2012) do not change. However, endogenous human capital decreases inequality relative to the economy with exogenous labor. Park (2014) derives positive capital income taxes from the pecuniary externality of the aggregate capital stock in a limited commitment economy. Evans (2017) studies the role of capital taxation in a model with uninsurable investment risk. The optimal capital tax rate balances the pecuniary effect of increased savings with redistribution to agents whose investment failed. Krueger and Ludwig (2018) analyze capital taxes in a Ramsey economy where the pecuniary externality cancels out the precautionary savings effect.

Influential studies in the literature analyze the steady state implications of a flat-tax or a capital income tax reform (Lucas (1990), Ventura (1999), and Conesa et al. (2009)), or restrict the tax schedule to a specific functional form: Heathcote et al. (2016) and Conesa and Krueger (2006) compute gains from the optimal *progressivity* of the income tax code. Heathcote et al. (2017) allow the degree of tax progressivity to vary with age. Bakis et al. (2015) apply the same parametric form to compute the optimal tax policy for a dynastic economy. Useful insights have been obtained by imposing restrictions on information available to the government in Golosov et al. (2003), Golosov et al. (2011), Kapicka (2013), or Heathcote and Tsujiyama (2017). Our paper shows that narrowing the analysis to monotone functions may be rather restrictive and that the shape of the optimal tax schedule is sensitive to parameterization and the resulting stationary distribution (see Mirrlees (1971), Saez (2001), Mankiw et al. (2009), or Diamond and Saez (2011)). Compared with the Mirrleesian literature, it is the tax schedule that attains a steady state where the endogenous distribution is optimal with respect to average welfare.

Finally, our paper adds to the new literature on quantitative methods. In a partial equilibrium framework, Golosov, Tsyvinski, and Werquin (2014) also use variational approach to 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. Perturbation methods have been recently used to analyze the optimal government responses to aggregate shocks. Bhandari et al. (2017a) study public debt in an economy where taxes and transfers are chosen optimally subject to heterogeneous agents' borrowing constraints and the distribution of debt ownership. In Bhandari et al. (2017b), the Ramsey planner optimally sets nominal interest rates, transfers and proportional labor taxes in response to aggregate shocks in a New Keynesian model where agents are heterogeneous with respect to co-movements of aggregate variables and measures of inequality.

In our example, we apply our methodology to an optimal limiting tax schedule on total income from labor and capital. There are several reasons why we choose this setup. First, the tax on total income preserves incomplete markets 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 total income is the complexity of the optimization problem. We discuss how our methodology can be extended to address important issues with respect to endogenous labor supply, taxation of capital income, borrowing constraints, government debt, or more detailed life-cycle features.

The paper is organized as follows. The following section defines the stationary Ramsey problem in a competitive equilibrium. Section 3 formulates the limiting Ramsey problem in the calculus of variations. The necessary and sufficient conditions in terms of a generalized Euler-Lagrange equations and Legendre condition are developed in Section 4. Sections 5 and 6 present an example with the optimal limiting income tax schedule. Section 7 concludes. Appendices contain proofs, additional results, and a sensitivity analysis.

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

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

where $U : \mathbb{R}_+ \to \mathbb{R}$ 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, $k_t \in B = [\underline{k}, \overline{k}]$, and by a discrete, exogenous labor productivity shock $z_t \in Z = \{\underline{z}, \ldots, \overline{z}\}$. We assume that there is a borrowing constraint that prevents the individual savings from being negative. The lower bound could be motivated by solvency constraints or by an explicit borrowing constraint. As is standard, the upper bound $\overline{k}$ is set very high and verified not to be binding in equilibrium. 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). The labor productivity shock is independent across agents and we preserve the heterogeneity in the economy by assuming incomplete markets: Agents do not have access to state-contingent contracts but can only accumulate the risk-free capital stock.

In each period, agents inelastically 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

$$r_t = F_K(K_t, L_t) - \delta \quad \text{and} \quad w_t = F_L(K_t, L_t), \tag{2}$$

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

Finally, there is a government that finances its expenditures by taxation. In order to preserve incomplete markets and, therefore, heterogeneity in the economy, we impose that the government cannot use state-contingent taxes. We assume that the government cannot issue debt and is fully committed to a sequence of tax functions $\{\pi_t\}_{t=0}^{\infty}$ to finance its expenditures equal to a fraction g of total output net of depreciation, not returned back to the agents.[1] The tax schedule is applied to a broadly defined taxable activity of each agent, $x_t \in \mathbb{R}_+$. We assume $x_t = x(z_t, k_t)$ where $x : Z \times B \to \mathbb{R}_+$ and $x_z, x_k > 0$. In each period, the policy schedule is a twice continuously differentiable 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 $\pi_t(x_t)$ and is left with an after-tax income $y_t - \pi_t(x_t)$.[2] An individual budget constraint in each period is then

$$c_t + k_{t+1} \leq r_t k_t + w_t z_t - \pi_t(x_t) + k_t.$$

The economy's aggregate state is characterized by the sequences of government policies $\{\pi_t\}_{t=0}^{\infty}$ and the distribution of agents over capital and productivity 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 $\mathcal{B}$ denotes the Borel sets that are subsets of B and $\mathcal{Z}$ is the set of all subsets of Z. Agents have rational expectations and take prices as given by equation (2). In order to determine prices, agents also have to know the evolution of the distribution function from an initial distribution λ_0, for each sequence of government policies $\{\pi_t\}_{t=0}^{\infty}$.

The objective of the government is to choose a sequence of $\{\pi_t\}_{t=0}^{\infty}$ to maximize

$$\sum_{t=0}^{\infty} \beta^t \sum_z \int u(c_t(z, k)) \lambda_t(z, k) dk, \tag{3}$$

[1] Specifying government expenditures net of depreciation simplifies the derivation of analytical properties that can be related to agents' incomes. In numerical simulations we use the usual formulation of expenditures as a fraction of total output. Our methodology equally applies to the case when government finances any level of expenditures $\{G_t\}_{t=0}^{\infty}$ and the corresponding revenue-neutral reforms.

[2] In Section 5, we compute an economy with a tax on total income, i.e. when $x = y$.

subject to agents' optimal allocations in each period, equilibrium prices determined by the aggregate capital stock and labor,

$$K_t = \sum_z \int k\lambda_t(z,k)dk, \quad \text{and} \quad L_t = \int \sum_z z\lambda_t(z,k)dk, \tag{4}$$

the government budget constraint,

$$g\left(F(K_t, L_t) - \delta K\right) = \sum_z \int \pi_t(x_t(z,k))\, \lambda_t(z,k)\, dk; \tag{5}$$

and a law of motion of the distribution,

$$\lambda_{t+1}(z', B') = \sum_z \int_{\{(z,k)\in Z\times B:\, h_t(z,k)\in B'\}} Q(z,z')\, \lambda_t(z,k)\, dk, \tag{6}$$

given an initial distribution λ_0.

The government problem assigns equal weights to all agents. This utilitarian approach is chosen for two main reasons. We prefer to start the economy from an initial distribution λ_0 where all agents have identical wealth and labor productivity. Second, assigning equal weights to all agents allows us to treat identical agents identically and derive properties for the long-run equilibrium associated with the optimal limiting government policy function.[3]

Definition 1 (Stationary Ramsey Problem) *A solution to the Stationary Ramsey Problem is a time-invariant limiting government tax policy function $\pi : \mathbb{R}_+ \to \mathbb{R}$ such that $\pi = \lim_{t\to\infty} \pi_t$ maximizes the government problem in (3)-(6).*

Note that our analysis is not a pure steady state utility maximization. The optimal limiting government policy is a long run optimal outcome that takes into account intertemporal discounting and the convergence to the steady state. In other words, we study a steady state under a modified Golden Rule. That is, we study a steady state of an economy for which the optimal limiting government policy implies a convergence to that steady state.

2.1 Recursive Formulation

The limiting optimal policy π needs to take into account its effects on equilibrium prices and agents' decisions.[4] Define the value function of each agent as $v : Z \times B \to \mathbb{R}$ and the savings function as $h : Z \times B \to B$. Given π and equilibrium prices $r(K)$ and $w(K)$, an

[3] See Davila et al. (2012) for a similar discussion. The initial distribution is arbitrary.

[4] With exogenous labor supply, the aggregate labor converges deterministically to a constant due to the law of large numbers. In the following exposition, we normalize the aggregate labor supply and write the equilibrium prices as functions of the aggregate capital K only.

agent (z,k) solves the following dynamic programming problem

$$v(z,k) = \max_{c,h} \Big\{ u(c(z,k)) + \beta \sum_{z'} v(z', h(z,k))\, Q(z,z') \Big\}, \tag{7}$$

subject to a budget constraint

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

with a taxable activity $x(z,k)$, total income $y(z,k) = r(K)k + w(K)z$, and a borrowing constraint, $h(z,k) \geq \underline{k}$.

Definition 2 (Recursive Competitive Equilibrium) *For a given share of government expenditures g and the government policy π on a taxable activity x, a recursive competitive equilibrium is a set of functions (v, c, h), aggregate levels (K, L), prices (r, w), and a probability measure $\lambda : Z \times B \to [0,1]$, such that for given prices and government policies,*

1. *the policy functions solve each agent's optimization problem (7);*
2. *firms maximize profit (2);*
3. *the probability measure evolves according to a law of motion,*

$$\lambda'(z', B') = \sum_{z} \int_{\{(z,k)\in Z\times B:\, h(z,k)\in B'\}} Q(z,z')\, \lambda(z,k)\, dk, \qquad \textit{for all } (z', B') \in \mathcal{Z}\times\mathcal{B}; \tag{8}$$

4. *the aggregation conditions hold, $K = \sum_z \int k\lambda(z,k)dk$, and $L = \sum_z \int z\lambda(z,k)dk$;*
5. *the government budget constraint, $g(F(K,L) - \delta K) = \sum_z \int \pi(x(z,k))\lambda(z,k)dk$.*

In the recursive formulation, the optimal limiting government policy maximizes

$$W(\lambda) = \max_{\pi} \int \sum_{z} u(c(z,k))dk + \beta W(\lambda'), \tag{9}$$

subject to allocations satisfying the conditions in Definition 2.[5]

The steady state of the economy corresponding to the limiting optimal government policy π is characterized by a time-invariant distribution λ. That the optimal limiting government policy π allows for a convergence to the steady state requires a regularity condition on its properties. Denote the interval of individual savings at which an agent with a productivity shock z is borrowing constrained as $[\underline{k}, \underline{k}(z)]$. For future reference also denote $\overline{k}(z)$ as the highest savings by an agent with a current productivity shock z.

Assumption 1 (Regularity Condition) *The government policy function π is such that for each $z \in Z$, the individual savings function $h : Z \times B \to B$ is a strictly increasing function for $k > \underline{k}(z)$ and is constant $h(k,z) = \underline{k}$ for $k \in [\underline{k}, \underline{k}(z)]$.*

[5]Note that if the allocations satisfy the definition of the recursive competitive equilibrium then they are also feasible.

A similar condition is required for the existence of a unique stationary recursive equilibrium in all models with heterogeneous agents (see Stokey et al. (1989)). It implies that the savings function does not display pathological features (for example, that wealthy agents save less than poor agents) so that the stationary distribution has a unique ergodic set. We want to make explicit here that this assumption is completely innocuous.[6]

3 Solution to the Stationary Ramsey Problem as a Calculus of Variations Problem

Since the problem is to find an optimal limiting, welfare maximizing continuous function $\pi \in \mathcal{C}^2(\mathbb{R}_+, \mathbb{R})$, we transform the Stationary Ramsey Problem into an operator form and solve it by the calculus of variations.[7] The calculus of variations is much more suitable for solving a problem with complicated 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: on the Euler equation and on the stationary distribution. For a given government policy function π, the Euler equation operator $\mathcal{F}$ is defined on the savings function $h : Z \times B \to B$. The distribution operator $\mathcal{L}$ is defined on the probability measure $\lambda : Z \times B \to [0, 1]$ and on the savings function h. We 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 $\mathcal{F} : C^1(Z \times B) \subset L^2(Z \times B) \to C^1(Z \times B) \subset L^2(Z \times B)$ is a mapping from a space of continuously differentiable functions into a space of continuously differentiable functions; and the operator $\mathcal{L} : C^1(Z \times B) \times C^1(Z \times B) \to C^1(Z \times B) \subset L^2(Z \times B)$. All functions in the calculus of variations depend on the government policy π and its derivative π_x.

[6]The Regularity Condition guarantees that for all $z \in Z$, the government policy function π 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 $k = h^{-1}(z, h)$. The Regularity Condition is used only in the law of motion for the distribution λ in the operator for the distribution function in equation (11). Davila et al. (2012) make a similar assumption (an increasing savings function h, in Appendix). In their case the savings policy is not distorted by a government policy function.

[7]Mirrlees (1976), Davila et al. (2012) or Golosov et al. (2014) use the same approach.

[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.

Operator $\mathcal{F}$ on the Euler Equation An individual agent's optimization problem in (7) is characterized by the Euler equation with an operator[9]

$$\mathcal{F}(h) \equiv u_c(c) - \beta \sum_{z'} u_c(c') \left[1 + y'_k - \pi_x(x')x'_k\right] Q(z, z'), \tag{10}$$

where in the next period $c' = y(z', h(z,k)) - \pi(x(z', h(z,k))) + h(z,k) - h(z', h(z,k))$, $y' = r(K')h(z,k) + w(K')z'$, and $x' = x(z', h(z,k))$. The term $[1 + y'_k - \pi_x(x')x'_k]$ is the after-tax marginal return to capital when $\pi_x(x')$ is the next-period marginal government policy. Finally, $y'_k = y_k(z', h(z,k)) = r(K')$ is the marginal effect of individual savings on total income and $x'_k = x_k(z', h(z,k))$ is the marginal effect of individual savings on the taxable activity x' next period.[10] The operator equation is $\mathcal{F}(h) = 0$.

Operator $\mathcal{L}$ on the Distribution Function Under the Regularity Condition, the operator $\mathcal{L}$ for the distribution function in equation (8) is

$$\mathcal{L}(\lambda, \lambda', h) \equiv \lambda'(z', k') - \sum_{z} Q(z, z') \frac{\lambda\left[z, h^{-1}(z, k')\right]}{\frac{d}{dk} h\left(z, h^{-1}(z, k')\right)}, \tag{11}$$

for all $(z', k') \in Z \times [h(z, \underline{k}(z)), h(z, \overline{k}(z))]$. The operator equation is $\mathcal{L}(\lambda, \lambda', h) = 0$.

A variational approach to the Ramsey problem is based on a sequential formulation of the government optimization problem. For descriptive purposes we present the recursive formulation.[11] Only two adjacent periods from the infinite time series are relevant for the government maximization problem. For a given tax policy π' and a savings function k'' in the next period, and for a given distribution λ, the government chooses π to maximize

$$\sum_{z} \int u\big[k(1 + r(K)) + zw(K) - \pi(x) - k'\big]\, \lambda(z,k) dk$$
$$+\beta \sum_{z'} \int u\big[k'(1 + r(K')) + z'w(K') - \pi'(x') - k''\big]\, \lambda'(z', k') dk',$$

with the updating operator for the next-period distribution

$$\lambda'(z', k') = \sum_{z} Q(z, z') \frac{\lambda\left[z, h^{-1}(z, k')\right]}{\frac{d}{dk} h\left(z, h^{-1}(z, k')\right)}.$$

The term $h^{-1}(z, k')$ denotes the value at k' of the inverse of h for each $z \in Z$. We use a change of variable $k' = h(z,k)$ to express $dk' = \frac{d}{dk} h(z,k) dk$. Merging the sums over z and

[9] In the text below, we present only the case of the unconstrained agents (for whom the Euler equation holds with equality and $h(z,k) > \underline{k}$). The case of borrowing constrained agents is in Appendix A.

[10] The Euler equation for an individual agent is standard. When the taxable activity equals total income, $\mathcal{F}(h) \equiv u_c(c) - \beta \sum_{z'} u_c(c') \left[1 + r(K') - \pi_x(x') r(K')\right] Q(z, z')$.

[11] See Appendix for derivation and a similar discussion in Davila et al. (2012).

the integration with respect to k, the optimal limiting tax function π maximizes

$$\sum_z \int \Big\{ u[k(1+r(K)) + zw(K) - \pi(x) - h(z,k)] \\ + \beta \sum_{z'} Q(z,z') u\big[h(z,k)(1+r(K')) + z'w(K') - \pi'(x') - k''\big] \Big\} \lambda(z,k) dk,$$

where the aggregate capital stock in the next period equals

$$K' = \sum_{z'} \int k' \sum_z Q(z,z') \frac{\lambda\left[z, h^{-1}(z,k')\right]}{\frac{d}{dk} h\left(z, h^{-1}(z,k')\right)} dk' = \sum_z \int h(z,k)\lambda(z,k) dk.$$

The Ramsey problem in the calculus of variations is then

$$\max_\pi \sum_z \int_{\underline{k}(z)}^{\overline{k}(z)} \Big[\mathcal{W}(z,k;\pi,\pi_x) + \beta \sum_{z'} Q(z,z') \mathcal{W}'(z,z',k;\pi,\pi_x) \Big] \lambda(z,k) dk,$$

where

$$\begin{aligned} \mathcal{W}(z,k;\pi,\pi_x) &\equiv u[k(1+r(K)) + zw(K) - \pi(x(z,k)) - h(z,k)], \\ \mathcal{W}'(z,z',k;\pi,\pi_x) &\equiv u[h(z,k)(1+r(K')) + z'w(K') - \pi'(x(z',h(z,k))) - k'']. \end{aligned}$$

We make explicit the dependence of operators $\mathcal{W}$ and $\mathcal{W}'$ on the tax function π and also on its derivative π_x. Solving for the optimal tax function, we need to move from the coordinates k to the taxable activity $x = x(z,k)$.[12] Therefore,

$$\max_\pi \sum_z \int_{\underline{x}(z)}^{\overline{x}(z)} \Big[\mathcal{W}(z,x;\pi,\pi_x) + \beta \sum_{z'} Q(z,z') \mathcal{W}'(z,z',x;\pi,\pi_x) \Big] d\lambda(x), \tag{12}$$

where

$$\begin{aligned} \mathcal{W}(z,x;\pi,\pi_x) &\equiv u[k(z,x)(1+r(K)) + zw(K) - \pi(x(z,x)) - h(z,k(z,x))], \\ \mathcal{W}'(z,z',x;\pi,\pi_x) &\equiv u[h(z,k(z,x))(1+r(K')) + z'w(K') - \pi'(x(z',h(z,k(z,x)))) - k''], \\ d\lambda(x) &\equiv \lambda(z,k(z,x))k_x(z,x). \end{aligned}$$

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 $\underline{x}(z) = \underline{x}(z,\underline{k})$ depends on z, on the exogenously given lower bound on capital $\underline{k}$, and on the equilibrium prices. Similar arguments apply to the upper bound $\overline{x}(z) = \overline{x}(z,\overline{k})$.[13]

[12] The taxable activity x is now the independent variable and $k = k(z,x)$ is its function. We want to stress again the dependence of the operator $\mathcal{W}$ and $\mathcal{W}'$ on the aggregate capital stock K and K' through general equilibrium effects, although it is not written as one of its arguments.

[13] Clearly, the maximal interval is $[\underline{x}(\underline{z}), \overline{x}(\overline{z})]$ where $\underline{x}(\underline{z})$ is the lower bound of the lowest shock, $\underline{z}$, and $\overline{x}(\overline{z})$ is the upper bound of the highest shock, $\overline{z}$. So any taxable activity interval associated with a shock $z \in Z$ is a subinterval of the maximal interval, $[\underline{x}(z), \overline{x}(z)] \subset [\underline{x}(\underline{z}), \overline{x}(\overline{z})]$.

The aggregate capital stock in the next period is defined as

$$K' \equiv \sum_z \int_{\underline{x}(z)}^{\overline{x}(z)} h(z, k(z,x))\, d\lambda(x), \tag{13}$$

and the side conditions for the government budget constraint is

$$\sum_z \int_{\underline{x}(z)}^{\overline{x}(z)} \mathcal{G}[z, x; \pi, \pi_x]\, d\lambda(x) = 0, \quad \text{where} \;\; \mathcal{G}[z, x; \pi, \pi_x] \equiv \pi(x) - gy(z, k(z,x)). \tag{14}$$

Definition 3 (Calculus of Variations Ramsey Problem) *The Ramsey Problem in the calculus of variations is a generalized isoperimetric maximization problem (12), subject to the government budget constraint (14), with the individual policy function h given implicitly by the operator Euler equation $\mathcal{F}(h) = 0$, the law of motion for the distribution function, λ, given implicitly by the operator equation $\mathcal{L}(\lambda, \lambda', h) = 0$, the aggregate capital stock (13), the endogenous 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.*

Note that since the upper bounds $\overline{k}(z)$ are endogenous, the endpoints $\overline{x}(z)$ are equality constrained. This might be also true for the lower bounds $\underline{k}(z)$ and their endpoints $\underline{x}(z)$.

4 Necessary and Sufficient Conditions for the Stationary Government Policy Function

In this Section we derive the first-order necessary and second order sufficient conditions for the optimal limiting 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, π and π_x. 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)).[14]

Definition 4 (Fréchet Derivative) *Given a nonlinear operator $\mathcal{N}(u)$ on function u, the Fréchet differential $\delta\mathcal{N}(u; \delta h) = \mathcal{N}_u \delta h$ is*

$$\lim_{\|\delta h\| \to 0} \frac{\|\mathcal{N}(u + \delta h) - \mathcal{N}(u) - \mathcal{N}_u \delta h\|}{\|\delta h\|} = 0,$$

[14]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 $\mathcal{N}_u = \lim_{\varepsilon \to 0} \mathcal{N}(u + \varepsilon \delta h)/\varepsilon$ when the obtained derivative is continuous.

where $\mathcal{N}_u$ is the Fréchet derivative.

Define the Lagrange function $\mathbf{L}$ for the Calculus of Variations Ramsey Problem in Definition 3, for each $z \in Z$, as

$$\mathbf{L}(z,x) = \begin{cases} 0 & \text{for } x \in [\underline{x}(\underline{z}), \underline{x}(z)), \\ \mathcal{W}(z,x) + \mu\mathcal{G}(z,x) & \text{for } x \in [\underline{x}(z), \overline{x}(z)], \\ 0 & \text{for } x \in (\overline{x}(z), \overline{x}(\overline{z})]. \end{cases} \tag{15}$$

Note that the social welfare function is the sum of integrands $\mathcal{W}(z,x) = \mathcal{W}[z,x;\pi(x),\pi'(x)]$ integrated on intervals $[\underline{x}(z), \overline{x}(z)]$ for each $z \in Z$. The same is true for integrands $\mathbf{L}(z,x)$.

Theorem 1 (First Order Necessary Conditions) *Using a modified Lagrange function $\tilde{\mathbf{L}}$ for the Calculus of Variations Ramsey Problem in Definition 3,*

$$\tilde{\mathbf{L}}(z,x) = \begin{cases} 0 & \text{for } x \in [\underline{x}(\underline{z}), \underline{x}(z)), \\ \left[\mathbf{L}(z,x) + \beta \sum_{z'} Q(z,z')\mathbf{L}'(z,z',x)\right]\lambda(z,x)k_x(z,x) & \text{for } x \in [\underline{x}(z), \overline{x}(z)], \\ 0 & \text{for } x \in (\overline{x}(z), \overline{x}(\overline{z})], \end{cases}$$

for each $z \in Z$, the first order necessary conditions for the Ramsey problem are

1. *the Euler-Lagrange equation,*

$$\sum_z \left(\tilde{\mathbf{L}}_\pi(z,x) - \frac{d}{dx}\tilde{\mathbf{L}}_{\pi_x}(z,x) \right) = 0; \tag{16}$$

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

$$\left[\tilde{\mathbf{L}}(\underline{z},x) - \left(\pi_x(x) - \frac{k_x(\underline{z},x)}{\omega_\pi(\underline{z},x)} \right) \tilde{\mathbf{L}}_{\pi_x}(\underline{z},x) \right]_{x=\underline{x}(\underline{z})} = 0; \tag{17}$$

3. *the transversality condition on the free boundary value, $\pi(\overline{x}(\overline{z}))$, at the equality constrained endpoint, $\overline{x}(\overline{z})$,*

$$\left[\tilde{\mathbf{L}}(\overline{z},x) - \left(\pi_x(x) - \frac{k_x(\overline{z},x)}{\omega_\pi(\overline{z},x)} \right) \tilde{\mathbf{L}}_{\pi_x}(\overline{z},x) \right]_{x=\overline{x}(\overline{z})} = 0; \tag{18}$$

4. *and the condition on the Lagrange multiplier, μ, at which (14) is satisfied.*

Proof At the endogenous upper bound the endpoint is equality constrained. If the extreme lower bound is exogenous, then the condition 3 simplifies to $\tilde{\mathbf{L}}_{\pi_x}(\underline{z},x)|_{x=\underline{x}(\underline{z})} = 0$. For the proof and more detailed specifications of all terms see the Appendix.

The total variation with respect to the tax schedule is equal to the sum of the total variation of utilities and the total variation of the budget constraint weighted by the shadow price μ in two consecutive periods. The total variation with respect to the optimal limiting tax policy schedule, i.e. $\pi = \lim_{t\to\infty} \pi_t$, is equal to zero. Denoting $\Delta \equiv \frac{\delta}{\delta\pi} - \frac{d}{dx}\frac{\delta}{\delta\pi_x}$ as the operator for the total variation, the Euler-Lagrange equation is simply $\sum_z \Delta\tilde{\mathbf{L}} = 0$.

The Lagrange-Euler conditions in Theorem 1 contain the tradeoff between the tax level, π, and its curvature captured by the marginal tax rate, π_x. At the optimum, the marginal effect of a change in the level of the tax schedule π on social welfare, $\tilde{\mathbf{L}}_\pi$, has to be equal to the marginal effect of an implied change in the tax schedule curvature expressed by the derivative of the implied change, $\frac{d}{dx}\tilde{\mathbf{L}}_{\pi_x}$, due to the changing marginal tax rate, π_x.[15]

Lemma 1 *The Euler-Lagrange equation (16) can be written as*

$$\begin{aligned}
&\sum_z \Bigg\{ \Big[-u_c(c) + \beta \sum_{z'} Q(z,z') u_c(c')[1 + r - \pi_x(x')x'_k] \Big] h^*_\pi \qquad (19)\\
&-u_c(c) + \mu \\
&+\beta \sum_{z'} Q(z,z') \Big(\int \psi(z,z',x) \big[u_c(c') - \mu g\big] h^*_\pi \lambda^*(z,k(z,x))k_x(z,x)dx + \mu\xi(z,z',x)h^*_\pi \Big) \\
&-\frac{d}{dx}\Big[\beta \sum_{z'} Q(z,z') \Big(\int \psi(z,z',x)\big[u_c(c') - \mu g\big] h^*_{\pi_x} \lambda^*(z,k(z,x))k_x(z,x)dx + \mu\xi(z,z',x)h^*_{\pi_x} \Big) \Big] \\
&-\varepsilon^{\lambda^* k_x}_x(z,x) \Bigg\} \lambda^*(z,k(z,x))k_x(z,x)dx \leq 0,
\end{aligned}$$

where

$$\begin{aligned}
\psi(z,z',x) &\equiv h^*(z,k(z,x))r_K(K^*) + z'w_K(K^*), \\
\xi(z,z',x) &\equiv \pi_x(x(z',h^*(z,k(z,x))))x_k(z',h^*(z,k(z,x))) - gr(K^*), \\
\varepsilon^{\lambda^* k_x}_x(z,x) &\equiv \frac{\frac{d}{dx}\left(\lambda^*(z,k(z,x))k_x(z,x)\right)}{\lambda^*(z,k(z,x))k_x(z,x)} = \varepsilon^{\lambda^*}_x(z,x) + \varepsilon^{k_x}_x(z,x),
\end{aligned}$$

Proof The Euler-Lagrange equation results from a substitution of terms defined in Appendix B into Theorem 1.

The government constructs its optimal limiting tax schedule by balancing the *distributional* effect on the *individual* savings function, h^*, and the *general equilibrium* effect on the *aggregate* capital stock. The first line in the Euler-Lagrange equation (19) represents an agent's intertemporal first-order condition that takes into account the tax schedule and is zero except for the borrowing-constrained agents.[16] In the second line the optimal limiting tax schedule directly affects the disposable income and, therefore, the marginal utility

[15] If we restrict the tax policy to be only a flat tax and the marginal tax rate is constant, the first-order conditions degenerate to the standard optimization problem $\tilde{\mathbf{L}}_\pi(z,x) = 0$.

[16] The Euler-Lagrange equation is written with inequality as it includes agents at the lower bound. If written with equality, the transversality condition at the lower bound applies to the case of borrowing constrained agents.

of consumption in the current period, $-u_c(c)$, counterbalanced by an opposite effect of the Lagrange multiplier, μ, on the government budget constraint. The indirect effects of agents' savings, h^*_π, are aggregated to changes in the next-period aggregate capital stock weighted by marginal utility, $\beta \sum_{z'} Q(z,z') \int \psi u_c(c') h^*_\pi \lambda^* k_x dx$. In a similar way the shape of the tax schedule influences the equilibrium through the effect of the marginal tax rate π_x on individual savings function, $h^*_{\pi_x}$, expressed by $\beta \sum_{z'} Q(z,z') \int \psi u_c(c') h^*_{\pi_x} \lambda^* k_x dx$. Note that both aggregate effects arise through the sensitivity functions h^*_π and $h^*_{\pi_x}$. Finally, the last line of the Euler-Lagrange equation is the semi-elasticity of the transformed distribution function with respect to the taxed activity, $\varepsilon_x^{\lambda^* k_x}$. It can be decomposed into a sum of the semi-elasticity of the distribution function and the semi-elasticity of the related capital k_x with respect to x.

The pecuniary externality effect operates through changes in equilibrium prices and can be written as in Davila et al. (2012),

$$\psi(z,z',x) = r_K(K^*)K^* \left(\frac{h^*(z,k(z,x)))}{K^*} - \frac{z'}{L} \right).$$

For a labor intensive income the term in the brackets is negative and $\psi > 0$ (since $r_K < 0$).

Compared to Davila et al. (2012), the government budget constraint causes additional pecuniary externality effects. First, there are aggregate capital effects on the government budget balance through the change of the tax rate and the marginal tax rate, $-g\beta \sum_{z'} Q(z,z') \int \psi u_c(c') h^*_\pi \lambda^* k_x dx$, and $-g\beta \sum_{z'} Q(z,z') \int \psi u_c(c') h^*_{\pi_x} \lambda^* k_x dx$. Second, individual savings decisions impact the government budget constraint through the change in the tax schedule, $\xi(z,z',x) = [\pi_x(x')x_k(x') - gr(K^*)]h^*_\pi$. The first term captures the change in tax contributions and the second the change in government spending due to the change in the agent's capital income. The total effect is $\beta \sum_z \sum_{z'} Q(z,z') \mu \xi(z,z',x) h^*_\pi \lambda^*(z,k(z,x)) k_x(z,x) dx$, weighted by $\lambda^*(z,k(z,x))k_x(z,x)dx$ across all agents with (z,x). A similar decomposition applies to marginal tax changes $h^*_{\pi_x}$.

Finally, to relate our results to Davila et al. (2012), we formulate the Euler-Lagrange equation for their constrained efficiency problem in which the social planner manipulates savings of each agent. As each agent receives back a lump-sum transfer equal to changes in investment, there is neither the government budget constraint (Lagrange multiplier μ) nor the direct tax effect on disposable income, $-u_c(c)$. Since the tax is not a function of x, the terms with $\pi(x)$ as well as d/dx will disappear from the equation. As the social planner chooses savings directly, $h^*_\pi = 1$. The first-order condition is then

$$\begin{aligned} &-u_c(c) + \beta \sum_{z'} Q(z,z') u_c(c')[1+r] \\ &\qquad + \beta \sum_z \sum_{z'} Q(z,z') \int \psi(z,z',x) u_c(c') \lambda^*(z,k(z,x)) k_x(z,x) dx \leq 0, \end{aligned}$$

identical to the first-order condition in Davila et al. (2012) with the last term equal to the extra term Δ.[17] In our formulation, the function ψ is weighted by agents' marginal utilities and the whole term also enters an individual agent's Euler equation as a number.

Theorem 2 (Second Order Sufficient Conditions) *A tax schedule π satisfying the first-order conditions in Theorem 1 attains a strict maximum if and only if (i) the Lagrange function defined in equation (15) satisfies the second-order Legendre condition*

$$\sum_{z\in Z} \tilde{\mathbf{L}}_{\pi_x\pi_x}(z,x) < 0 \quad \textit{for all } x \in [\underline{x}(\underline{z}), \overline{x}(\overline{z})]\,,$$

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

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 Stationary Ramsey Problem. In the following Sections we illustrate these effects numerically.

4.1 The Effects of Government Policy on the Equilibrium

If we knew how agents' saving policies h and *simultaneously* how the distribution λ depend on the government policy schedule, i.e. if we could solve at equilibrium prices for the optimal policy which is a function of the distribution and prices which in turn are determined by $h(\cdot)$ which is itself a function of the optimal policy and prices, the 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, agents' saving policy and the distribution functions are known only implicitly as a solution to the two operator equations ($\mathcal{F}(h) = 0$ and $\mathcal{L}(\lambda, \lambda', h) = 0$) and the aggregate conditions for equilibrium prices.

The next Lemma derives the effects of the government policy function π on the operator Euler equation by specifying four unknown "sensitivity" functions $h_\pi : Z \times B \longrightarrow \mathbb{R}$, $h_{\pi_x} : Z \times B \longrightarrow \mathbb{R}$, $h_{\pi_x\pi} : Z \times B \longrightarrow \mathbb{R}$, and $h_{\pi_x\pi_x} : Z \times B \longrightarrow \mathbb{R}$. Denote the next-period after-tax marginal return to capital as $R' = 1 + y'_K - \pi_x(x')x'_K$.

Lemma 2 (The Effects of π and π_x on the Euler Equation) *The total F-derivatives of the operator Euler equation $\mathcal{F}$ with respect to the government policy function π*

[17] Note that the summation over current shocks z disappears as the planner chooses allocations contingent on each shock. The Euler-Lagrange equation for a flat-tax schedule is in the Appendix.

and to its derivative π_x *are*

$$\mathcal{F}_i = u_{cc}(c)c_i - \beta \sum_{z'} Q(z,z')\big\{u_{cc}(c')c_i'R' + u_c(c')R_i'\big\} = 0, \tag{20}$$

where the subscript denotes derivatives with respect to $i \in \{\pi, \pi_x\}$*, and*

$$\begin{aligned} \mathcal{F}_{ij} &= u_{ccc}(c)c_ic_j + u_{cc}(c)c_{ij} - \\ \beta \sum_{z'} Q(z,z')&\Big\{ \big[u_{ccc}(c')c_i'c_j' + u_{cc}(c')c_{ij}'\big]R' + u_{cc}(c')\big[c_i'R_j' + c_j'R_i'\big] + u'(c')R_{ij}'\Big\} = 0. \end{aligned} \tag{21}$$

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

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

We obtain five functional equations (10), and (20)-(21) in unknown functions h, h_π, h_{π_x}, $h_{\pi_x\pi_x}$, and $h_{\pi_x\pi}$, respectively. Finally, by adding the first-order conditions from Theorem 1, and the functional equation for the distribution function, λ, the problem of finding the optimal government policy π is a system of seven functional equations in seven unknown functions with two side conditions and one condition on the Lagrange multiplier.

5 An Example: The Optimal Income Tax Schedule

In this section we demonstrate our method by finding the optimal limiting government policy π defined as a tax on total income from capital (net of depreciation) and labor. The taxable activity is

$$x(z,k) = y(z,k) = r(K)k + w(K)z,$$

and the individual budget constraint is

$$c(z,k) + h(z,k) \leq x(z,k) - \pi(x(z,k)) + k.$$

As before, there is a borrowing constraint $\underline{k} = 0$ and the total tax revenues are equal to a fraction g of the total output. In the abbreviated notation, the Euler equation for a (z,k)-agent's optimal savings function h is now

$$u'(c) \geq \beta \sum_{z'} u'(c')\left[(1 + r(K') - \pi(x')r(K')\right] Q(z,z'),$$

where $c' = x' - \pi(x') + h - h'$, $x' = r(K')h + w(K')z'$, and $h' = h(z', h(z,k))$. Note that for this specification $x_k = y_k = r(K)$ and $k_x = 1/x_k = 1/r(K)$.

5.1 Admissible Tax Functions

Because the tax schedule is an arbitrary continuous function, we must ensure that the first-order approach is valid and that the stationary recursive equilibrium exists.[18] In order to characterize the admissible tax functions and to prove the Schauder Theorem for economies with distortions, we follow the notation in Stokey et al. (1989). For each agent $(z, k) \in B \times Z$, denote the after-tax gross income as $\varphi(z, k) \equiv x(z, k) - \pi(x(z, k)) + k$, and rewrite the Euler equation as

$$u'(\varphi(z, k) - h(z, k)) = \beta \sum_{z'} u'(\varphi(z', h(z, k)) - h(z', h(z, k)))\varphi_k(z', h(z, k))Q(z, z'),$$

where $\varphi_k(z', h(z, k)) = 1 + r(K') - \pi_x(x(z', h(z, k)))r(K')$ is the marginal after-tax return of investment.

Theorem 3 *For a given tax schedule $\pi : \mathbb{R}_+ \to \mathbb{R}$, if for each $(z, k) \in B \times Z$, $\varphi_k(z, k) > 0$, and φ 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 functions.

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

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

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

The above statement follows directly from the fact that $\varphi_k(z, k) > 0$ and that φ is quasi-concave. The corollary implies that the marginal tax rate must be smaller than $1+1/r(K)$. This upper bound is not likely to bind for a very wide range of tax schedules.[19] Application of Theorem 1 and Theorem 2 then implies necessary and sufficient conditions for the unique maximum attained by the tax schedule.

[18] Again, we analyze the case of borrowing constrained agents in the Appendix.

[19] For the targeted equilibrium interest rate in the progressive-tax steady state $r = 0.04$, the upper bound on the marginal tax rate is equal to 26. When numerically solving for the optimal tax schedule in the next Section we do not impose any bounds on the marginal tax rate but we check the admissibility of the optimal tax schedule ex post.

Table 1: Parameters of the Benchmark Economy

$\beta = 0.9309$	$\sigma = 2.0$	α= 0.36	$\delta = 0.08$	g=0.189
Earnings Process (Davila et al. (2012)):				
$z \in Z = \{1.000, 5.290, 46.550\}$		$Q(z, z') = \begin{bmatrix} 0.992 & 0.008 & 0.000 \\ 0.009 & 0.980 & 0.011 \\ 0.000 & 0.083 & 0.917 \end{bmatrix}$		

6 Numerical Solution

In this Section we solve for the optimal limiting tax schedule and compare the associated steady state allocations to those resulting from the progressive tax schedule in the U.S. economy and from the usual flat-tax reform. In order to evaluate welfare implications, we conduct a transition analysis.

We use the same calibration as Davila et al. (2012). The three-state, first order Markov process of the uninsurable idiosyncratic shock to labor productivity is based on Diaz-Jimenez et al. (2003) with large and asymmetric wage risk necessary for a realistic dispersion of both earnings and wealth. We set the discount factor so that in the progressive-tax steady state the capital-output ratio equals 3 and the equilibrium interest rate is 4 percent. Other parameters in Table 1 are standard, $\alpha = 0.36$, $\delta = 0.08$, and the risk aversion parameter $\sigma = 2$ (intertemporal elasticity of substitution of 1/2). The stationary distribution of productivity shocks is $\{0.498, 0.443, 0.059\}$.

We model the progressive tax schedule as Heathcote, Storesletten, and Violante (2016) by specifying a tax on total income as a function

$$\pi^{PT}(y(z, k)) = y(z, k) - \chi y(z, k)^{1-\tau},$$

where τ is a parameter of the rate of progressivity and χ is a level parameter that clears the government budget constraint. The tax function is progressive if $0 < \tau < 1$ with strictly increasing marginal tax rates.[20] For the U.S. economy, Heathcote et al. (2016) estimate from the 2000-2006 PSID data that $\tau = 0.161$. In each period the government is required to collect tax revenues equal to 18.9% of the total output. As depreciation of capital is deducted from taxable income, in the flat-tax rate steady state $\tau = g/(1 - \delta K/Y)$.

The optimal limiting tax policy is a solution to the system of functional equations defined in Theorem 1, the functional equation for the stationary distribution λ, the two side conditions, and the Lagrange multiplier condition. We solve this functional equations problem by the least squares projection method described in Appendix C. Definitions of functional equations are in Appendix D, together with additional results and a sensitivity analysis for different parameters of risk aversion.[21] In Appendix E we also compute an

[20] Examples of monotone tax functions are Conesa and Krueger (2006) and Gouveia and Strauss (1994).
[21] The least squares projection method is an efficient and well-behaved method for functional equations

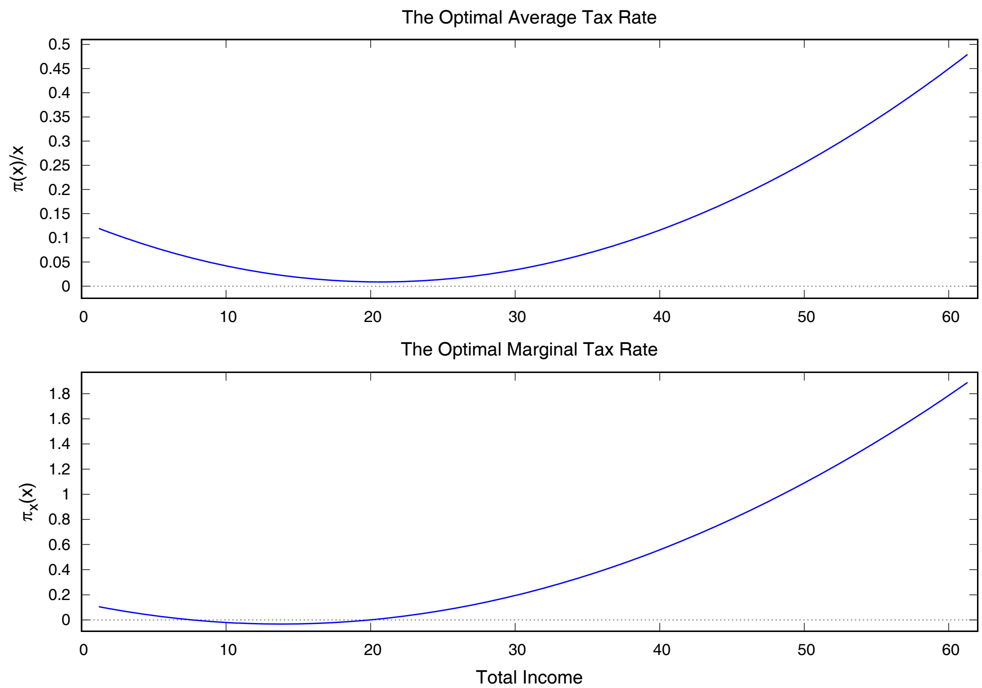

Figure 1: The optimal average tax rate and the optimal marginal tax rate.

alternative calibration for a low wealth dispersion economy based on Aiyagari (1994).

6.1 Steady State Results

The optimal limiting average tax schedule is U-shaped. In Figure 1, the average tax rate at the lowest total income is 14%, decreases to a minimum of 2%, and rises to 48% at the highest level of total income. The marginal tax rates are also U-shaped: close to zero for low incomes, negative at medium incomes, and positive at higher income levels.[22]

Table 2 shows steady state outcomes of the progressive, flat, and optimal limiting tax schedules. Under the progressive tax, the persistent earnings process generates a large fraction of poor agents and a substantial right tail of the wealth distribution with a Gini coefficient of wealth above 0.8, exactly as in Davila et al. (2012). The flat-tax reform increases the aggregate capital stock by 31% and output by 10%, while only slightly increasing inequality. The optimal limiting income tax increases the aggregate capital stock by only 3 percent. The coefficient of variation of wealth falls substantially and the Gini coefficient of wealth inequality becomes 0.698. While in the progressive- and flat-tax steady

problems. For a detailed explanation of applying projection methods to stationary equilibria in economies with a continuum of heterogeneous agents see Bohacek and Kejak (2002).

[22]The optimal limiting tax schedule easily satisfies the admissibility condition in Corollary 1. Similar marginal tax rates are found in dynamic models of Kapicka (2013) or Golosov et al. (2011) and in static models of Diamond (1998) and Saez (2001). Building on Mirrlees (1971) and Mirrlees (1976) seminal work, Kocherlakota (2005), Golosov et al. (2003), or Albanesi and Sleet (2006) study optimal social planner policies with asymmetric information needed for characterization of optimal policies.

Table 2: Steady State Results

	Progressive	Flat	Optimal
τ	0.161	0.265	—
Aggregate assets	3.000	3.934	3.093
Output	1.000	1.102	1.011
Capital-output ratio	3.000	3.568	3.059
Interest rate (%)	4.000	2.092	3.771
Coeff. of variation of wealth	2.125	2.313	1.501
Gini wealth	0.812	0.838	0.698
Perc. of wealth of the top 5%	0.433	0.475	0.276

states the top five percent of richest agents hold more than 40 percent of all assets, the optimal limiting tax schedule lowers this share to 27.6 percent.

There are two important effects the optimal limiting tax schedule takes into account. The first is the general equilibrium price effect derived in Davila et al. (2012). Poor agents in the benchmark economy have labor intensive incomes and an increased aggregate capital stock improves their welfare through higher equilibrium wages: the extra term defined in Davila et al. (2012) is positive in all our simulations, suggesting an overaccumulation of capital in a competitive equilibrium relative to the constrained efficient allocation by a social planner.

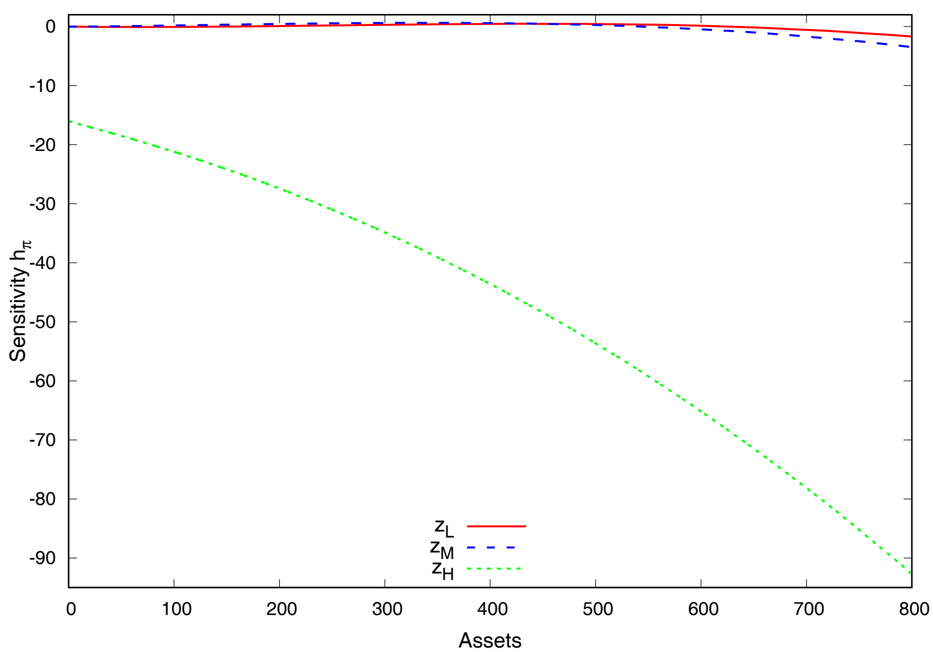

Figure 2: Sensitivity of savings to changes in the tax schedule h_π.

The second, distributional effect is related to agents' insurance and mobility within the stationary distribution. The large and persistent dispersion of earnings makes redistribution important. While in Davila et al. (2012) the wealthy agents are only induced to save

Table 3: Steady State Distribution of Assets and Tax Contributions

	Assets			Tax Contributions		
	Progress.	Flat	Optimal	Progress.	Flat	Optimal
1st Quintile	0.1	0.1	0.2	-3.4	5.8	2.6
2nd Quintile	0.2	0.2	0.3	-3.4	5.9	2.9
3rd Quintile	1.7	1.5	8.1	4.0	14.4	3.3
4th Quintile	9.9	7.0	18.8	9.6	16.4	4.3
5th Quintile	88.3	91.9	72.7	89.9	57.4	89.5
Top 10%	66.4	71.6	48.5	81.7	54.3	87.0
Top 5%	43.3	47.5	27.6	64.5	41.1	76.3
Top 1%	12.1	13.5	6.6	14.5	8.8	17.4

Note: Each entry is the percentage share of assets owned or taxes paid by each group.

more, under the optimal limiting tax schedule they are taxed in order to provide transfers to agents with high marginal utility. Figure 2 shows the sensitivity functions of savings to changes in the optimal tax schedule, h_π, for three labor productivity shocks $z_L < z_M < z_H$. The response of savings by agents with the low and medium productivity shocks is negligible across the ergodic set of the wealth distribution as all perturbations of the tax schedule are consumed by the agents. Only agents with the highest labor productivity alter their savings: a perturbation of the convex increasing tax schedule reduces savings, $h_\pi(k, \underline{z}) < 0$.

The substantial right tail of the income distribution contributes a large fraction of the total tax revenues. Table 3 shows that the optimal tax schedule puts most of the tax burden on the top decile of the wealth distribution. The top decile pays almost 90 percent and the top five percent pays more than 3/4 of the total tax revenues, respectively, much more than in the progressive-tax steady state. For a comparison, the flat-tax reform dramatically increases the aggregate capital stock without taking into account the distribution of agents. Under the flat tax, the share of the tax burden of the top decile is only 54.3% of all tax contributions compared to above eighty percent in the other steady states.

In general, the low and initially decreasing marginal tax rates of the limiting optimal tax policy insure poor agents and improve efficiency by lowering distortions in the economy. The marginal tax rates motivate the savings of poor agents towards the desired aggregate capital level while the increasing rates on high incomes deliver revenues for redistribution. The optimal limiting tax schedule lowers the share of assets held by the top quintile to 72.7%; the top decile owns less than one half of total assets. This redistribution is reflected in the low wealth inequality in the optimal limiting tax steady state.

6.2 Welfare Gains from the Optimal Limiting Tax Schedule

The top part of Table 4 shows that the optimal tax schedule delivers large average welfare gains in the steady state, 1.7% and 6.1% with respect to the progressive and the flat

Table 4: Welfare Gains from the Optimal Tax Schedule (in %)

	Progressive	Flat
Steady State	1.735	6.129
Transition		
Average Welfare	-0.767	14.884
Aggregate Component	-0.580	5.670
Distributional Component	-0.188	8.720
Political Support	46.471	94.129

tax steady state, respectively.[23] Because a pure steady state comparison involves different stationary distributions and ignores transition costs, the rest of Table 4 shows welfare gains in terms of expected present discounted values from an unanticipated reform in which the optimal limiting tax schedule is imposed on the initial progressive- or flat-tax steady state.[24] We adopt the approach of Domeij and Heathcote (2004) and decompose the average welfare gain into an aggregate and a distributional component. The former denotes a hypothetical expected present value of per-period consumption if a household consumes each period the same fraction of aggregate consumption as in the pre-reform steady state. The latter is the difference between the average welfare gain and the aggregate component.[25]

Imposing the optimal limiting tax schedule on the progressive-tax steady state leads to short-run costs in both distributional and aggregate components. The progressive tax steady state provides better short-time insurance to the poorest agents as well as requires more savings. On the other hand, in the flat-tax steady state the aggregate capital stock is very high and can be deaccumulated during the transition. This additional consumption as well as increased insurance delivers a 14.8% welfare gain per period and a majority political support. Detailed results in Appendix D show that all but the high productivity agents support the tax reform.

Obviously, we do not compute an optimal transition process that also adjust the shape of the optimal tax policy in each period of the transition. Also, efficiency gains in the terminal steady state are large and compensation schemes could be designed to alleviate welfare losses from the transition in order to obtain political support for the reform.[26]

[23]Steady state average welfare is defined as the expected discounted value of being born into a stationary equilibrium, expressed in consumption units per period.

[24]The average welfare gain from the optimal tax reform is defined as a constant percentage increase in consumption of each household in the pre-reform steady state that delivers the same expected utility as when the optimal limiting tax schedule is implemented. We guess a sufficiently large number of convergence periods and iterate on paths of equilibrium prices and tax levels to clear markets and the government budget constraint in each period of the transition.

[25]In a representative agent economy, the distributional component is zero. If the aggregate component is positive, the reform is Pareto-improving if it leaves the distribution of consumption unchanged. In Domeij and Heathcote (2004), the distributional component is negative and outweighs the aggregate component as the capital income tax reform shifts taxation to labor income.

[26]See Gottardi et al. (2011) for an analysis of issuing debt against subsequent efficiency gains.

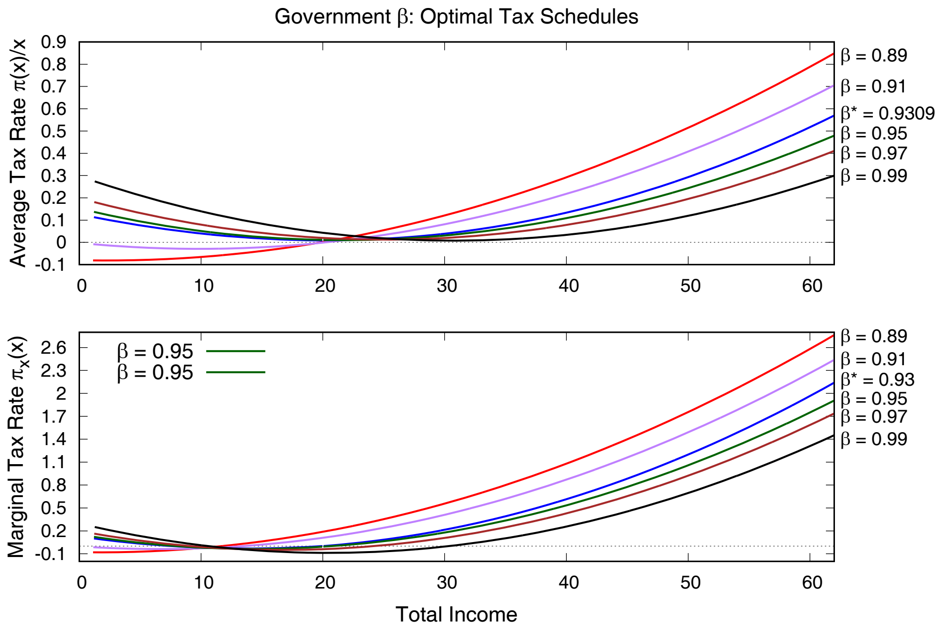

Figure 3: Optimal average and marginal tax rates for different government discount factors.

6.3 Different Welfare Criteria

We now relate our results to recent findings in the private information literature that society should be more patient than individual households.[27] In dynamic models, the social discount factor affects the way how government (social planner) balances efficiency and equality between the current and future generations (or periods). Indeed, the high aggregate capital stock in the constrained optimum in Davila et al. (2012) might arise in an economy whose social planner is more patient than agents. To study this issue, we set the government's discount factor in the calculus of variations Ramsey problem in equation (12) to $\hat{\beta}$ that differs from the benchmark discount factor β used by agents in the Euler equation (10).

Figure 3 shows that when the social discount factor is lower than that of agents, $\hat{\beta} < \beta$, the optimal tax schedule puts more weight on current redistribution and less weight on the long-term general equilibrium effect from the aggregate capital stock accumulation. The optimal tax schedule for a very low $\hat{\beta} = 0.89$ is actually increasing and convex. On the other hand, a very high social discount factor leads to a more convex U-shaped average tax

[27] This issue has been studied in an asymmetric information framework as a tool to overcome the immiseration result of Atkeson and Lucas (1992). Phelan (2006) achieves this by assigning equal weights on all future generations, while Farhi and Werning (2007) place a positive and vanishing Pareto weight on expected welfare of future generations. Farhi and Werning (2010) analyze a dynamic Mirrleesian model with productivity shocks and find that a progressive estate tax implements efficient allocations by providing the necessary mean reversion and insurance across generations. Farhi et al. (2012) study efficient non-linear taxation of labor and capital in a dynamic Mirrleesian model without commitment. In order to lower the capital stock and, therefore, the gains from a deviation, the marginal tax on capital income is progressive.

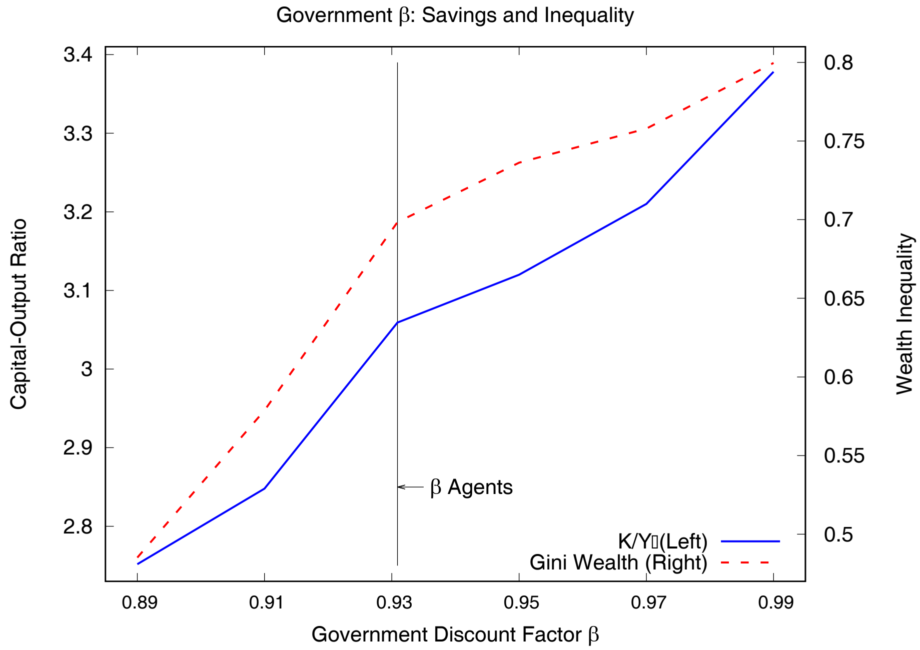

Figure 4: Capital-output ratio (left axis) and wealth inequality (right axis) for different government discount factors.

function needed for saving incentives (see Bakis et al. (2015) for similar results in a dynastic economy). Figure 4 shows that both the capital-output ratio and wealth inequality increase in $\hat{\beta}$. When $\hat{\beta} = 0.99$, the capital-output ratio rises to 3.38 and Gini coefficient of wealth to 0.80. On the other hand, a very low discount rate $\hat{\beta} = 0.89$ decreases wealth inequality to 0.485. In Appendix D we report all steady state and transition results. For $\hat{\beta} < \beta$, the optimal tax reform of the progressive-tax steady state is supported by the majority of the population. This is because for low $\hat{\beta}$s the government prefers a lower aggregate capital stock, so agents can deaccumulate and consume their capital stock during a transition, and vice versa for high social discount factors. For the same reason, the optimal tax reform of the flat-tax economy makes the majority of agents better off.

The shape of the optimal tax schedule is necessarily sensitive to the government's objective function. Heathcote and Tsujiyama (2017) characterize the mapping between a taste for redistribution in a class of Pareto weight functions and the progressivity parameter in Heathcote et al. (2016). When the taste for redistribution increases, the marginal tax schedule moves from being upward sloping to becoming a U-shaped function. In their simulations, the welfare gains from an optimal progressive tax reform are very small. Appendix F discusses these extensions and other results in detail.

6.4 Sensitivity Analysis

Finally, we analyze the optimal limiting tax schedule in the benchmark calibration for different values of risk aversion σ. The U-shape of the optimal limiting average tax schedule holds in all simulated economies. An increasing risk aversion puts more weight on the distributional effect relative to the general equilibrium effect and the average tax schedules become more progressive with higher σ. The Gini coefficient of wealth falls to 0.59 when $\sigma = 4$. In the flat-tax economy, a higher risk aversion induces agents to accumulate more capital as a buffer stock against idiosyncratic risk (see Imrohoroglu (1998)), rising transitional gains as a consequence. Because general equilibrium effects are more important at low risk aversion, the negative slope of the average tax rates is steeper at low incomes.

6.5 A Low Wealth Dispersion Economy

Mirrlees (1971), Mankiw et al. (2009), or Diamond and Saez (2011) note that the shape of the marginal tax schedule is sensitive to the distribution of skill or income. In Appendix E, we evaluate our example in Section 6 for a more traditional calibration based on Aiyagari (1994). The three-state, first order Markov process of the idiosyncratic shock to labor productivity is estimated from the household annual labor income process in the PSID data by a first-order autoregression with a persistence parameter 0.6 and a volatility 0.2.

This calibration of the earnings process leads to a steady state with an unrealistic, low dispersion of wealth. The optimal limiting average tax schedule is also U-shaped with an important upward shift: the average tax rate at the lowest total income is 37%, decreases to a minimum of 24%, and rises to 48% at the highest level of total income. The marginal tax rates are close to zero for low incomes and rise at higher income levels to provide resources for redistribution and insurance. In this low wealth dispersion economy, the earnings process is not very persistent, wealth inequality is low and poor agents can escape poverty easily. This allows the optimal tax schedule to focus more on the general equilibrium price effect, increasing the aggregate capital stock by 18.7%, to a higher level than in the flat-tax steady state. The savings decisions of all types of agents are very sensitive to perturbations of the optimal limiting tax schedule. The degree of persistence in the idiosyncratic labor productivity process has similar effects to those found in Evans (2017). In that paper, the optimal capital tax rate balances the general equilibrium price effect and redistribution. When the productivity shocks are iid, the former effect dominates and the planner subsidizes savings decisions as in Davila et al. (2012). When persistence of the productivity shocks increases, the latter effect is more important and the long-run capital tax is positive. Golosov et al. (2016) and Mankiw et al. (2009) also find U-shaped marginal rates originating from a large fraction of agents with low productivity.[28]

7 Conclusion

In this paper, we provide a solution method for optimal limiting government policies in a general equilibrium economy with incomplete markets. We think of these policies as optimal because they take simultaneously into account their effects on the distribution of agents and equilibrium prices. In our example, we find the optimal limiting tax schedule on total income in a stationary Ramsey problem. As in Diamond (1998), Saez (2001), or Golosov et al. (2011), the optimal limiting average tax schedule is U-shaped. The shape of the optimal limiting tax schedule balances the efficiency-equality trade-off between general equilibrium price effects and redistribution. In the calibrated economy with high wage and wealth dispersion, the distributional effects dominate and the high income agents contribute a large fraction of the total tax revenues. Initially low and decreasing marginal tax rates improve efficiency by lowering distortions in the economy. When a government discounts future more than agents or if the latter are more risk averse, the short-term distributional effect is more important than the long-term, general equilibrium effect from changes in the aggregate capital stock. Compared with Davila et al. (2012), the aggregate capital stock increases only slightly but the optimal limiting tax function significantly reduces inequality. In other words, without agent-specific lump-sum transfers the optimal tax policy would not subsidize rich agents' savings at the cost of not providing insurance to poor agents with high marginal utility. Our results confirm the conjecture in Davila et al. (2012) that simple uniform policies cannot achieve the constrained optimum.

This paper is a first step for analyzing more realistic models with important policy implications. Our example shows that analyzing linear tax functions or restricting functional forms to progressive taxation might miss a large part of efficiency and equality effects. The most important extension of the model would introduce the extensive labor margin. Labor decisions are more relevant for poor agents while savings decisions are more important for wealthy agents. In a simplified two-period model, Davila et al. (2012) find that the pecuniary externality might induce the agents to supply inefficiently low effort. Park (2017) analyzes the constrained optimum with endogenous human capital, motivated by Huggett et al. (2011) who find that the determinants of endogenous human capital dispersion are more important than idiosyncratic shocks as a source of lifetime earnings dispersion. Because the planner in Park (2017) can only alter equilibrium prices, the qualitative results

[28] Heathcote and Tsujiyama (2017) the shape depends on how much fiscal pressure the government faces. When fiscal pressure is low, the marginal tax schedule is monotone increasing. As fiscal pressure increases, the optimal schedule becomes first flatter and then U-shaped.

of Davila et al. (2012) do not change. However, numerical simulations lead to a lower capital-labor ratio and a reduced inequality relative to the economy with exogenous labor. While the optimal limiting tax schedule might motivate poor agents to increase labor effort, the welfare of poor households could be better improved by reducing the risky part of their income through a lower wage. Therefore, the redistribution channel arising from the endogenous labor income dispersion could further reduce inequality and lower the optimal capital-labor ratio. It is also not clear whether the optimal tax schedule would still display increasing marginal rates as they reduce the returns to labor and the returns to human capital investment.

In our future research we also plan to study the optimal limiting tax schedule with life-cycle income profiles as in Golosov et al. (2011) and Farhi and Werning (2012). With respect to the class of optimal government policies, an important extension is in relaxing the assumption on continuity of the tax function. Finally, we would also like to explore different (Rawlsian) welfare functions and the role of government debt. Further restrictions on available policy tools imply additional aggregate and distributional effects. One can also analyze different calibrations with heterogeneous preferences, opportunities to participate in labor or asset markets. A major task would be an extension of our methodology to the private information literature. Finally, we plan to make our methodology applicable to time-varying optimal policies for heterogeneous response of agents to aggregate shocks.