

Optimal Fiscal Policy with Heterogeneous Agents ^{*}

Marco Bassetto
Northwestern University[†]

First version: May 1996
This version: February 1999

Abstract

The aim of this paper is to study how the intertemporal behavior of taxes affects the wealth distribution. The optimal-taxation literature has often concentrated on representative-agent models, in which it is optimal to smooth distortionary taxes. When tax liabilities are unevenly spread in the population, deviations from tax smoothing lead to interest rate changes that redistribute wealth. When a “bad shock” hits the economy, the optimal policy will then call for smaller or larger deficits depending on the political power of different groups. The model is applied to war financing and the introduction of a balanced-budget policy.

1 Introduction

The tax structure in an economy is in part the result of a struggle over the distribution of resources. The aim of this work is to study one aspect of this struggle, the choice of an optimal intertemporal tax plan. The intertemporal aspect of fiscal policy is important because any government constantly faces fiscal shocks. These may come from a wide variety of sources: business cycles, financial crises, the transition from a centralized to a decentralized economic system and wars. In any of these cases, the government must choose among various policies for accommodating the shock. For example, a negative shock can be absorbed by an increase in taxes, a low return on previously issued state-contingent debt, or new issues of debt to be repaid with future taxes.

Different groups of agents in the economy have different preferences over these policies, and the goal of this paper is to study how these differences are resolved.

On the normative side, I am interested in studying the characteristics of second-best tax policies, which will trace a (second-best) Pareto frontier. A benevolent government would choose one of these policies; which one depends on its relative preferences for the different classes of

^{*}JEL Classification numbers: E62, H21 **keywords:** Optimal taxation, heterogeneous agents, asset prices, distortion, net trade

[†]Department of Economics, Northwestern University, 2003 Sheridan Road, Evanston, IL 60208. Phone: (847) 491-8233, fax (847) 491-7001, email: m-bassetto@nwu.edu. I am indebted to Robert Lucas, Casey Mulligan, José Scheinkman, Nancy Stokey and especially Thomas Sargent for helpful comments. I also benefited from conversations with Fernando Alvarez, Gary Becker, Mariacristina De Nardi, Larry Jones, Kenneth Judd, Adriano Rampini, Andrew Sellgren and François Velde. Any remaining errors are my own. Financial support from the Alfred P. Sloan Foundation through a dissertation fellowship is gratefully acknowledged.

agents in our economy. Along the Pareto frontier, we can also inquire whether there is a trade-off between “equity and efficiency.” For extreme values for the Pareto weights, the incentive to redistribute wealth may lead the government to impose significant distortions in the economy.

In addition, although in reality the political process is more complicated than a benevolent planner, a good political system will select policies close to the Pareto frontier. For this reason, my analysis is also likely to have some positive implications. Moreover, these are more evident the sharper the welfare differences among the possible policies, i.e., in cases of large fiscal shocks, such as wars.

A striking example of different experiences in war financing comes from England and France in the 17th and 18th centuries. As discussed in Sargent and Velde [18], England relied heavily on debt to finance its wars, while France made heavy use of temporary tax increases. These differences cannot be easily explained by representative-agent models, but I show through an example that they can be accounted for by the theory I am proposing.

The main conflict I study in this paper opposes the “taxpayers,” who bear the burden of taxes, and the “rentiers,” simply identified as all the other agents in the economy. In the model I present, I concentrate on labor-income taxation by considering an economy without capital, similar to the setup of Lucas and Stokey [13]: output is produced by a constant-returns-to-scale (CRS) technology requiring only labor supplied by the taxpayers; labor income is taxed by the government in order to finance an exogenous stream of public spending. The rentiers live off their asset income, which could possibly come from transfers to which the government is committed.¹ Government spending is the main source of uncertainty; the desirability of random policies is explored by introducing a further stochastic process that plays the role of a public randomization device. The government and the two classes of agents trade in complete markets. For most of the paper I look at Ramsey equilibria, i.e., I assume that the government commits at the beginning to a (state-contingent) tax policy. Only in section 7 do I allow for the possibility that the government chooses its policies sequentially, and I study conditions that guarantee time consistency of the Ramsey equilibria.

With constant government spending, I show that tax smoothing is beneficial to the taxpayers, but that the rentiers would achieve a higher welfare if the government varied tax rates over time. The traditional, representative-agent analysis of this model concludes that a constant tax rate is optimal because it minimizes the distortions caused by taxation. With heterogeneous agents, when the tax rate is higher the taxpayers borrow from the rentiers, whereas the opposite happens when the tax rate is low. A varying tax rate, furthermore, affects the prices in a way that favors the rentiers and harms the taxpayers. This is the source of the distributional conflict.

When government spending varies over time, the taxpayers favor large deficits in periods of high spending. Government spending is ultimately paid for by the taxpayers; when it is temporarily high, they need to borrow, either directly or through government debt backed by future tax revenues. A lower tax rate (i.e., a larger deficit) affects prices so that the taxpayers borrow at more favorable terms.

The plan of the paper is the following: section 2 discusses the related literature; section 3 presents the general model; section 4 specializes it to the 2-class economy I study in greater detail. Section 5 presents numerical examples of the solution and discusses the main results.

¹Allowing the rentiers to supply some work would not change our results, if they were subject to a low or zero tax rate: the important difference between the two classes comes from the fact that the taxpayers provide all (or most) of the government revenues.

Section 6 contrasts the results I obtain with the literature on uniform commodity taxation and explains the similarities and the sources of differences. Section 7 contains some results on the time consistency of the optimal tax plans when the assumption of full commitment is relaxed, and section 8 concludes.

2 The Literature

Many papers have studied how the tax burden should optimally be distributed over time. A seminal paper by Barro [3] established a benchmark to assess the relevance of this question. The central result of this paper is the proposition known as Ricardian equivalence: in a world of identical agents (or dynasties) and lump-sum taxes, the timing of taxes has no effect on prices, the allocation, and welfare in the economy. Therefore, any theory of optimal taxation over time must consider either distortionary taxes or heterogeneous agents.

The literature that has focused on distortionary taxes has mainly studied representative-agent economies.

Barro [4] studied a simple model of convex distortions from taxation. The paper derives tax smoothing as the principal policy prescription: the optimal tax plan spreads distortions over time and finances temporary increases in government spending by issuing debt.

Turnovsky and Brock [21] is, to my knowledge, the first paper that addresses the issue of optimal income taxation in a dynamic general equilibrium model with distortions. However, their paper is cast in a deterministic context and does not address the issue of the optimal fiscal response to shocks, which is Barro's main focus.

Lucas and Stokey [13] (L-S from now on) build a dynamic general equilibrium framework to address Barro's issue; their work generalizes the work by Ramsey [16] on optimal commodity taxation in a static microeconomic environment. L-S show that the government can smooth labor income tax revenues even more than in Barro's setup by "purchasing insurance" against unfavorable shocks to government spending. This can be achieved by issuing state-contingent bonds whose return is higher in favorable states and lower when a "bad shock" occurs.² More examples along this line are provided in Sargent and Velde [19], who specialize this problem to a linear-quadratic framework.

Several papers have generalized the results of L-S. For instance, Chari, Christiano and Kehoe [8] and Bohn [7] abandon the assumption of a pure exchange economy and study the optimal intertemporal taxation when capital and investment decisions are taken into account.

Most of these papers have studied representative-agent economies and have thus disregarded the distributional effects that different tax plans have in a world of heterogeneous agents.

An exception is Persson and Svensson [15], who study optimal taxation in an open economy, where heterogeneity is given by the presence of domestic and foreign consumers. However, their emphasis is mainly on the conditions required for time consistency of the optimal policy; my research is instead aimed at characterizing the optimal policy and studying the economic forces that drive it.

The choice of optimal distortionary taxes in a static environment with heterogeneous agents has been addressed by Atkinson and Stiglitz [2]. One contribution of their paper is to provide

²As Chari, Christiano and Kehoe [8] suggest, this is equivalent to capital income taxation on government bonds. Following this interpretation, total tax revenues would then be usually more volatile than government spending, rather than smoother. The crucial result is however that labor income tax revenues are smoothed.

conditions in which uniform commodity taxation is optimal. I will show why these conditions are violated in our dynamic environment; in doing so, I will also clarify similarities and differences between the static and dynamic models of optimal taxation that are left implicit by Lucas and Stokey [13].

Among the papers that have studied the distributional implications of different intertemporal tax plans, several have focused on the conflict between the elderly, the young, and the unborn that arises in overlapping-generations models when the altruistic links are weak, unlike in Barro [3].³ In this setup, different tax plans lead to different prices and allocations even when taxes are lump-sum because the agents are short-lived and government debt shifts the tax burden to future generations.

I focus on other sources of heterogeneity. In particular, an important one stems from the observation that the burden of taxes is spread unevenly in the population. In the past, some classes in the population were largely exempt from taxes (e.g., the nobility or the clergy); the model I develop resembles most closely this extreme case. In modern times, the burden of labor income taxes is unevenly spread because of progressive income taxation, which hits the most productive agents disproportionately, and because of the different importance of labor income in the budget of different households in the economy.

Similar sources of heterogeneity have been studied especially by authors interested in capital income taxation, as Judd [12]. Judd, however, studies mainly what happens when some of the agents (the “workers”) are denied access to capital markets; for the case in which all agents have access to capital markets, only an asymptotic result is derived. These limitations are common. Ben-Gad [6] studies the effect of the timing of taxes in two-period models, where the major impact of government debt is due to the presence of incomplete markets. Conklin [9] studies the time-consistency properties of an optimal tax plan when costless default is allowed; his paper shares the same two-class division I will assume in the numerical examples, but it is not well suited to address the effect on interest rates of different policies because the taxpayers are denied access to any financial market in his model. As a consequence, the government acts mainly as a financial intermediary between the taxpayers and the rentiers.

As Judd [12] points out, optimal taxation may be very different when all agents are allowed access to financial markets. We will focus on the impact on prices and the allocation that distortionary taxation can have in a complete-market environment. In choosing the optimal path of taxes, the government affects the asset-pricing kernel. This can be used both for reducing the burden of taxation and for redistributing wealth.

3 The Model

In this section we present the general setup of the economy: we introduce the preferences, the technology, and the government; we define the equilibrium concept to be used throughout most of the paper; and we provide a few general results that will be useful for characterizing the solution.

We consider an economy populated by N agents, which may differ by their preferences, their initial wealth, and their productivity while at work.

³See, e.g., Weil [23], Tabellini [20].

3.1 Endowment and Technology

There is an exogenous stream of public spending; public spending does not affect the utility of the agents.⁴ We define g_t to be public spending in period t ; we will also use the convenient notation $g_t^s \equiv \{g_u\}_{u=t}^s$ and $g^s \equiv \{g_u\}_{u=0}^s$. $\{g_t\}_{t=0}^\infty$ is a stochastic process with values in a measurable space (G, \mathcal{G}) , where G is a subset of the real line and \mathcal{G} is its Borel σ -algebra.

For reasons that will become clear later, it is convenient to introduce a “public randomization” or “sunspot” process $\{h_t\}_{t=0}^\infty$ with values in (H, \mathcal{H}) .

Let J_t denote the σ -algebra generated by (g^t, h^t) : J_t contains hence all information available up to time t . Given any random variable y_t measurable with respect to J_t , we define $E_s y_t$ as the expectation conditional on J_s .⁵

There is no storage, and only one consumption good. Output is produced through a CRS technology: in each period and each state, 1 unit of time spent working by agent i produces w^i units of output. Each agent is endowed with 1 unit of time. Each agent must choose a plan for consumption and leisure $\{(c_t^i, x_t^i)\}_{t=0}^\infty$ adapted to the filtration $\{J_t\}_{t=0}^\infty$, where c_t^i is consumption of the i -th agent at time t and x_t^i is leisure of the i -th agent at time t .

The government can levy proportional taxes on (or provide subsidies to) the labor income of each agent in the economy. We assume that the tax rate is constrained to be equal across agents and that the marginal tax rate is constant on all labor income (i.e., there is proportional taxation). The former is an “anonymity” assumption: the government cannot directly distinguish among the different agents, so it is not possible to tailor the tax rate to the individual agents. The latter assumption is mainly aimed at simplifying the analysis, but it may also be necessary to avoid lump-sum taxation when we have a small number of types of agents.⁶ We assume that the tax rate must be adapted to $\{J_t\}_{t=0}^\infty$. This implies that at each date and in each state the consumer knows the tax rate before implementing the consumption/leisure decision for that period. We do not allow the government to “toss a coin” to determine the amount of taxes due *after* production has taken place.

There are complete markets, both for privately-issued and publicly-issued securities; the government is not allowed to default on previously issued debt instruments, so privately- and publicly-issued claims are perfect substitutes.

We will define ${}_s b_t^i$ to be the amount of government-issued contingent claims payable at time t that the i -th agent holds at the beginning of period s ; if this is a negative number, then it will mean that the i -th agent owes to the government. We will also define ${}_s \eta_t^i$ to be the amount of privately-issued contingent claims payable at time t that the agent holds at the beginning of period s . Both ${}_s b_t^i$ and ${}_s \eta_t^i$ are hence random variables adapted to $\{J_t\}_{t=0}^\infty$.

⁴As usual, the results do not change if public spending does enter in the utility function of the individuals, but only in a separable way.

⁵We allow $t = +\infty$ to account for infinite summations.

⁶If the government were able to choose freely how the tax rate varied with income, it would try to set it at arbitrarily high values on inframarginal income and at 0 or very low at the income levels chosen by each of the agents in the economy.

3.2 Preferences

The preferences of the i -th consumer are described by:

$$U^i \equiv E_0 \sum_{t=0}^{\infty} \beta^t u^i(c_t^i, x_t^i). \quad (1)$$

We assume u^i is strictly concave, is continuously differentiable, and satisfies Inada conditions.

The preferences of the government are described by the following social welfare function:

$$W \equiv \sum_{i=1}^N \omega^i U^i, \quad (2)$$

where ω^i is the Pareto weight of the i -th agent (a single individual, or the representative agent of the i -th group).

3.3 Equilibrium

We will consider Ramsey equilibria. In the Ramsey problem the government sets a contingent policy at time 0 and is never allowed to revise it. The agents take the policy parameters as given. Since markets are complete, we can assume that they choose their optimal contingent plans at time 0 based on a single Arrow-Debreu budget constraint. More precisely, the timing of the economy is described by the following:

- (i) The economy starts at time 0 with some given level of public spending g_0 ; each agent inherits some claims from the past: the i -th agent starts with ${}_0b_t^i$ government-issued claims and ${}_0\eta_t^i$ privately-issued claims. Consistency requires this initial condition to satisfy:

$$\sum_{i=1}^N {}_0\eta_t^i \equiv 0 \quad \forall t \geq 0 \text{ a.s.} \quad (3)$$

We assume that both ${}_0b_t^i$ and ${}_0\eta_t^i$ are adapted to the information generated by (the history of) g_t alone: the coupon payments inherited at time 0 do not depend on the realizations of the process h_t , which is why we may call it a “sunspot.”

- (ii) The government sets a contingent path for the tax rate $\{\tau_t\}_{t=0}^{\infty}$.
- (iii) The consumers make their optimal (contingent) plans given the government policy, by choosing $\{(c_t^i, x_t^i)\}_{i=1}^N\}_{t=0}^{\infty}$. Note that the consumers behave atomistically, so the outcome would not change if we considered their decision to be sequential.
- (iv) h_0 is realized.

In the subsequent periods, the production and consumption plans are implemented according to the decisions made at time 0.

The i -th agent has the following Arrow-Debreu budget constraint:

$$E_0 \sum_{t=0}^{\infty} \beta^t \{p_t [c_t^i - {}_0\eta_t^i - {}_0b_t^i - (1 - \tau_t)w^i(1 - x_t^i)]\} \leq 0, \quad (4)$$

where p_t is the asset pricing kernel.

The government budget constraint for the Ramsey problem is:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ p_t \left[g_t + \sum_{i=1}^N ({}_0b_t^i - \tau_t w^i (1 - x_t^i)) \right] \right\} = 0. \quad (5)$$

Definition 1 (competitive equilibrium) A competitive equilibrium is a policy $\{\tau_t\}_{t=0}^{\infty}$, an allocation $\{(c_t^i, x_t^i)\}_{i=1}^N\}_{t=0}^{\infty}$, a price system $\{p_t\}_{t=0}^{\infty}$ and initial conditions $\{({}_0b_{t,0}^i, \eta_t^i)\}_{i=1}^N\}_{t=0}^{\infty}$ s.t.

- (i) given the price system, the government policy and the initial conditions, the allocation maximizes the utility of each consumer subject to her budget constraint described by (4);
- (ii) the government budget constraint (5) is satisfied;
- (iii) the markets clear, i.e.,

$$\sum_{i=1}^N c_t^i + g_t = \sum_{i=1}^N w^i (1 - x_t^i) \quad \forall t \geq 0 \text{ a.s.} \quad (6)$$

Definition 2 (Ramsey equilibrium) A Ramsey equilibrium is any competitive equilibrium for which (2) is maximized by choice of a policy function.

We first look at the conditions for a competitive equilibrium. The first-order conditions for the private agents of our economy are the following:

$$\frac{u_x^i(c_t^i, x_t^i)}{u_c^i(c_t^i, x_t^i)} \geq w^i (1 - \tau_t) \quad \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \quad (7)$$

and

$$\frac{u_c^i(c_t^i, x_t^i)}{u_c^i(c_0^i, x_0^i)} = \frac{p_t}{p_0} \quad \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N, \quad (8)$$

where (7) must hold with equality if $x_t^i < 1$, i.e., when the agent is supplying a positive amount of labor. The non-negativity constraints on consumption and leisure are never binding because of the Inada conditions. A competitive equilibrium is characterized by equations (4), (6), (7) and (8). Note that one budget constraint is redundant, so the government budget constraint (5) is automatically satisfied given the previous equations.

From standard comparative statics, we see that a change in p_t will increase the welfare of an agent i when $c_t^i - {}_0\eta_t^i - {}_0b_t^i - (1 - \tau_t)w^i(1 - x_t^i) < 0$ and decrease it when the reverse inequality holds. An agent will benefit from an increase in the relative price of the (possibly contingent) good she is relatively more endowed with. When $c_t^i - {}_0\eta_t^i - {}_0b_t^i - (1 - \tau_t)w^i(1 - x_t^i) < 0$, the i -th agent is a net seller of the considered good: her endowment, either from production or from initial financial claims, is larger than her consumption.

In a world of heterogeneous agents, it will often happen that some agents are net buyers and some agents are net sellers of any given good. By affecting the pricing kernel, the government will thus achieve redistribution among these agents.

In our model, preferences and technology are common knowledge, and the government would take into account the net trade position of each agent in each state-contingent good. In practice, our results will be interesting when a specific source of heterogeneity determines a clear pattern of trade among private agents: it is in this case that the distributional goals of the government will play a significant role in our analysis.

3.4 General Properties

We will now prove some properties that will be very useful in characterizing and computing Ramsey equilibria.

Theorem 1 *For any competitive equilibrium of the described economy, there exist functions $\{C^i\}_{i=1}^N, \{X^i\}_{i=1}^N, P : G \times \mathbb{R} \rightarrow \mathbb{R}$ such that:*

$$\begin{cases} c_t^i = C^i(g_t, \tau_t) & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \\ x_t^i = X^i(g_t, \tau_t) & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \\ p_t = P(g_t, \tau_t) & \forall t \geq 0 \text{ a.s.} \end{cases} \quad (9)$$

Proof. See Appendix.

Note that the functions C^i, X^i and P depend on which competitive equilibrium we are in. The theorem compares consumption-leisure choices *within* a given competitive equilibrium, not *across* competitive equilibria.

Theorem 1 states that, in any given competitive equilibrium, the consumption and leisure choices of all agents in the economy will be the same in all periods and/or states in which government spending and the tax rate are the same.

Theorem 2 *For given initial conditions $\{(ob_{t,0}^i, \eta_t^i)\}_{i=1}^N\}_{t=0}^\infty$ and a given process $\{g_t\}_{t=0}^\infty$ for government spending, let $\{\tau_t\}_{t=0}^\infty$ and $\{\tilde{\tau}_t\}_{t=0}^\infty$ be two policies satisfying the following requirements:*

$$\begin{aligned} & \sum_{t=0}^{\infty} \beta^t \text{Prob}((g_t, \{(ob_{t,0}^i, \eta_t^i)\}_{i=1}^N, \tau_t) \in A) \\ & = \sum_{t=0}^{\infty} \beta^t \text{Prob}((g_t, \{(ob_{t,0}^i, \eta_t^i)\}_{i=1}^N, \tilde{\tau}_t) \in A) \quad \forall A \in \mathcal{G} \times \mathcal{B}^{2N+1}. \end{aligned} \quad (10)$$

Let

$$\begin{cases} c_t^i = C^i(g_t, \tau_t) & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \\ x_t^i = X^i(g_t, \tau_t) & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \\ p_t = P(g_t, \tau_t) & \forall t \geq 0 \text{ a.s.} \end{cases} \quad (11)$$

describe an allocation and a price system that form a competitive equilibrium given the initial conditions, the spending process and the policy $\{\tau_t\}_{t=0}^\infty$. Then the same functions:

$$\begin{cases} c_t^i = C^i(g_t, \tilde{\tau}_t) & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \\ x_t^i = X^i(g_t, \tilde{\tau}_t) & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \\ p_t = P(g_t, \tilde{\tau}_t) & \forall t \geq 0 \text{ a.s.} \end{cases} \quad (12)$$

describe an allocation and a price system that form a competitive equilibrium given the initial conditions, the spending process and the policy $\{\tilde{\tau}_t\}_{t=0}^\infty$. Furthermore, the utility of each agent is the same in either equilibrium, and hence the same is true for government welfare.

Proof. See Appendix.

Definition 3 (Policy equivalence) We call two policies equivalent whenever (10) holds.

Theorem 2 justifies the definition of equivalence. Intuitively, it does not matter what kind of randomization over taxes the government chooses, or its distribution over time: in an Arrow-Debreu economy, it only matters how often it takes a given value and how it moves in line with the “fundamentals,” i.e., government spending and the coupon payments.

As an example of equivalent policies, consider a world where government spending is constant and all outstanding claims at time 0 are annuities, so $(g_t, \{(0b_{t,0}^i \eta_t^i)\}_{i=1}^N)$ are constant and deterministic. The first policy sets the tax rate to some level τ_1 in even periods and to some other level τ_2 in odd periods. The second policy sets the tax rate permanently to either τ_1 or τ_2 , depending on the outcome of h_0 ; the policy is designed in such a way that the probability of the tax rate being τ_1 is $\frac{1}{1+\beta}$. It is easy to see that (10) holds for these policies. While the behavior over time of the economy under these two policies is very different, they look very similar ex ante from an Arrow-Debreu point of view: in one case the agents and the government will trade claims to consumption over time, whereas in the other case they will trade claims to consumption across states. What matters, though, is that the marginal rate of substitution between goods in even and in odd periods in the former case and between goods in the first and in the second state in the latter are equal whenever the consumption levels are. This result arises from the presence of Arrow-Debreu markets and from the fact that our preferences are additive both with respect to time (strong time separability) and with respect to different events (a property of Von Neumann-Morgenstern preferences).

The government would be indifferent between two equivalent policies, and so would each of the private agents. Furthermore, the allocation and the price system are described by the same functions in the competitive equilibria associated with the two different policies. Therefore, if we solve for the competitive equilibrium associated with a given policy, we can infer immediately the allocation and the price system that form a competitive equilibrium with any policy that is equivalent to it.

Guided by Theorem 2, we will now restrict our attention to a simpler set of policies.

Corollary 1 Let $\{\tau_t\}_{t=0}^\infty$ be a policy adapted to the information generated by $\{g_t\}_{t=0}^\infty$ and any sunspot process $\{h_t\}_{t=0}^\infty$. Let $\{\tilde{h}_t\}_{t=0}^\infty$ be the following sunspot process: $(\tilde{H}, \tilde{\mathcal{H}}) = ([0, 1], \mathcal{B}([0, 1]))$, \tilde{h}_0 is distributed according to a uniform distribution and is independent of $\{g_t\}_{t=0}^\infty$; $\tilde{h}_t = \tilde{h}_0 \quad \forall t \geq 0$ a.s.. Then there exists a policy $\{\tilde{\tau}_t\}_{t=0}^\infty$, equivalent to $\{\tau_t\}_{t=0}^\infty$, such that $\tilde{\tau}_t$ can be expressed as a (measurable) function of g_t , $\{(0b_{t,0}^i \eta_t^i)\}_{i=1}^N$ and \tilde{h}_0 .

Proof. See Appendix.

Corollary 2 Let $\{h_t\}_{t=0}^\infty$ be a sunspot process described as follows: $(H, \mathcal{H}) = ([0, 1], \mathcal{B}([0, 1]))$, h_0 is distributed according to a uniform distribution and is independent of $\{g_t\}_{t=0}^\infty$; $h_t = h_0 \quad \forall t \geq 0$ a.s.. Let $\{\tau_t\}_{t=0}^\infty$ be the best policy among those adapted to the information generated by $\{g_t\}_{t=0}^\infty$

and the sunspot process $\{h_t\}_{t=0}^\infty$, i.e., the one that leads to the competitive equilibrium with the highest value W for the government. Then $\{\tau_t\}_{t=0}^\infty$ achieves a payoff which is greater than or equal to the payoff that the government can achieve using the best policy adapted to the information generated by g_t and any sunspot process $\{\hat{h}_t\}_{t=0}^\infty$.

Proof. By Corollary 1, the best policy that is adapted to the information generated by $\{g_t\}_{t=0}^\infty$ and a sunspot process $\{\hat{h}_t\}_{t=0}^\infty$ is equivalent to some policy adapted to the information generated by $\{g_t\}_{t=0}^\infty$ and $\{h_t\}_{t=0}^\infty$. The implication then follows trivially. QED.

From now on we will assume the sunspot process to be the one described in Corollary 2. Since the process is simply constant after time 0, we will drop the time subscript and use just h to indicate the single random variable h_0 . We will study the Ramsey equilibrium where the allocation and the asset-pricing kernel only depend on $(g_t, \{(0b_{t,0}^i \eta_t^i)\}_{i=1}^N, h)$. This is done simply for convenience; using Theorem 2, we can characterize Ramsey equilibria in which the government follows different (but equivalent) policies, such as deterministic variations of the tax rate over time even when the fundamentals are constant.

From now on we will thus define the measure m as:

$$m(A) \equiv \sum_{t=0}^{\infty} \beta^t \text{Prob}((g_t, \{(0b_{t,0}^i \eta_t^i)\}_{i=1}^N, h) \in A) \quad \forall A \in \mathcal{G} \times \mathcal{B}^{2N+1}, \quad (13)$$

and we will use this measure in evaluating (1), (2), (4) and (5).

4 The Two-class Economy

In this section we specialize the general framework presented above. We consider an economy populated by two types of agents: a measure N_1 of agents of type 1 and (by normalization) a measure 1 of agents of type 2. Type 1 agents are “rentiers.” Their productivity is 0, so they always choose $x_t^1 = 1$. They have no labor income, and live only out of their assets.⁷ Type-2 agents are identified as the “taxpayers,” as they are the only ones having labor income and thereby paying taxes. We normalize their productivity to be $w^2 = 1$.

We assume agents to be completely homogeneous within groups. When $\omega^1 = 0$, i.e., when the government maximizes the welfare of the taxpayers only, we can interpret this as an open economy where the government does not have the authority to tax foreigners; in this case our setup is basically the same as that in Persson and Svensson [15].⁸

We assume the agents have the following utility functions:

$$u^1(c_t^1, x_t^1) = \frac{(c_t^1)^{1-\gamma} - 1}{1-\gamma} \quad (14)$$

⁷As mentioned in the introduction, the important aspect for our analysis is not that the rentiers do not work, but that they are not subject to taxes. We could easily adjust the analysis to allow for labor income to be earned by the rentiers, with little difference for the results, as long as their income was not subject to taxes.

⁸Our economy is closest to the second setup in Persson and Svensson [15], in which they allow for perfect capital mobility. The only difference is that the rentiers (i.e., the foreign agents) are endowed with an independent stream of revenues, whereas in our case they only live off assets held against the government or the taxpayers (the domestic agents). We could also allow for an independent stream of revenues without altering the results substantially.

and

$$u^2(c_t^2, x_t^2) = \frac{(c_t^2)^{1-\gamma} - 1}{1-\gamma} + \xi \frac{(x_t^2)^{1-\sigma} - 1}{1-\sigma}. \quad (15)$$

The form of the utility function for the leisure component for type-1 agents is irrelevant, since they will always choose $x_t^1 \equiv 1$.⁹

We define $v \equiv (h, g_t, {}_0b_{t,0}^1, {}_0b_{t,0}^2, {}_0\eta_t^1, {}_0\eta_t^2)$. By the theorems of section 2, the allocation, the policy, and the price system in the Ramsey equilibrium will be functions of v . We will also denote $g_t = g(v)$ and so on, where these are just selector functions that take the appropriate component of the vector v .

We define $e(v)$ to be aggregate private consumption in the Ramsey equilibrium. The theorems and corollaries of section 2 established that in the Ramsey equilibrium aggregate consumption must be a function of the mentioned variables only. Note that we dropped ${}_0\eta_t^1$ since consistency requires $N^1 {}_0\eta_t^1 + {}_0\eta_t^2 = 0 \quad \forall t \geq 0$ a.s.. We now show that, in a competitive equilibrium, knowing $e(v)$ is enough to infer uniquely the allocation, the policy, and the price system.

Using (8) we find that in a competitive equilibrium:

$$c^i(v) = k^i e(v) \quad \forall t \geq 0 \text{ a.s.} \quad i = 1, 2, \quad (16)$$

and the asset-pricing kernel is given by

$$p(v) = e(v)^{-\gamma} \quad \forall t \geq 0 \text{ a.s.} \quad (17)$$

We can compute k^1 from the budget constraint of the agents of type 1 after substituting (16) and (17) and the definition of the measure m :

$$k^1 = \left[\int ({}_0b^1(v) + {}_0\eta^1(v)) e(v)^{-\gamma} dm(v) \right] \left[\int e(v)^{1-\gamma} dm(v) \right]^{-1}. \quad (18)$$

Since aggregate private consumption is $e(v) \equiv N^1 c^1(v) + c^2(v)$, we can compute k^2 from the requirement $N^1 k^1 + k^2 = 1$:

$$k^2 = 1 - N^1 k^1. \quad (19)$$

We use the market-clearing condition (6) to determine leisure:

$$x^2(v) = 1 - c^2(v) - N^1 c^1(v) - g(v) = 1 - e(v) - g(v). \quad (20)$$

We finally determine the tax policy using (7):

$$\xi x^2(v)^{-\sigma} = (1 - \tau(v)) c^2(v)^{-\gamma}$$

$$\tau(v) = 1 - \xi (1 - e(v) - g(v))^{-\sigma} (1 - N^1 k^1)^\gamma e(v)^\gamma. \quad (21)$$

The previous equations show that specifying aggregate private consumption is all we need to characterize a competitive equilibrium, given the initial conditions. We now wish to establish which functions $e(v)$ are compatible with a competitive equilibrium. In order for this to happen, the following must hold:

⁹Although the functional form we chose is different from that in Chari, Christiano and Kehoe [8], it is consistent with their baseline preferences if we set $\gamma = \sigma = 1$ and $\xi = 3$.

- (i) Equations (16)–(21) must have a well-defined solution in the admissible range; if this happens, (7), (8) and (6) will be satisfied, and so will (4) for type-1 agents. In (16), we need $c^i(v) \geq 0 \quad \forall t \geq 0$ a.s. $i = 1, 2$; this requires $e(v) \geq 0 \quad \forall t \geq 0$ a.s. $i = 1, 2$ and $k^i \geq 0 \quad i = 1, 2$, which can be rewritten as requiring $k^1 \in [0, \frac{1}{N_1}]$. We also need $x^2(v) \geq 0 \quad \forall t \geq 0$ a.s., which requires $e(v) + g(v) \leq 1 \quad \forall t \geq 0$ a.s..
- (ii) Either (4) for type-2 agents or (5) must hold (the other one will hold by Walras' law).

We will use the budget constraint of type-2 agents (4). Using equations (16)–(21), this constraint can be rewritten, after some algebra, as:

$$\begin{aligned} & \log \xi + \gamma \log(1 - N^1 k^1) \\ & + \log \left[\int (1 - e(v) - g(v))^{-\sigma} dm(v) - \int (1 - e(v) - g(v))^{1-\sigma} dm(v) \right] \\ & - \log \left[\int e(v)^{1-\gamma} dm(v) - \int e(v)^{-\gamma} (N^1 {}_0 b^1(v) + {}_0 b^2(v)) dm(v) \right] = 0. \end{aligned} \quad (22)$$

We assume that, given $\{\{({}_0 \eta_t^i, {}_0 b_t^i)\}_{i=1,2}\}_{t=0}^\infty$, there exist functions $e(v)$ that satisfy (i) and (ii). Intuitively, this requires the rentiers not to be so poor that their wealth is negative under any government policy, nor so rich that their wealth exceeds the value of the highest attainable output net of government spending; it also requires government spending not to be too large and the government not to be too heavily indebted against private agents. These requirements are necessary for existence of a competitive equilibrium given the initial conditions and the spending process.

Assuming thus that we have at least one competitive equilibrium, we wish to find now the one that maximizes the objective function of the government, which is the Ramsey equilibrium.

Using (16)–(20) the expected utility of type-1 agents is:

$$U^1 = (k^1)^{1-\gamma} \int \frac{e(v)^{1-\gamma}}{1-\gamma} dm(v), \quad (23)$$

the expected utility of type 2 agents is:

$$U^2 = (1 - N^1 k^1)^{1-\gamma} \int \frac{e(v)^{1-\gamma}}{1-\gamma} dm(v) + \int \frac{(1 - e(v) - g(v))^{1-\sigma}}{1-\sigma} dm(v), \quad (24)$$

and the objective function of the government is thus

$$\begin{aligned} W = & \left[\omega^1 N^1 (k^1)^{1-\gamma} + \omega^2 (1 - N^1 k^1)^{1-\gamma} \right] \int \frac{e(v)^{1-\gamma}}{1-\gamma} dm(v) \\ & + \omega^2 \int \frac{(1 - e(v) - g(v))^{1-\sigma}}{1-\sigma} dm(v). \end{aligned} \quad (25)$$

Condition 1 Either of these properties is met:

- $\gamma \geq 1$ and there exists a policy that satisfies the government budget constraint (22) and $e > 0$ a.s.; or:

- $(g_t, N^1 {}_0b_t^1 + {}_0b_{t,0}^2, \eta_t^1 + {}_0b_t^1) \neq \mathbf{0}$ a.s.

Condition 1 is not very important for our results, but it is useful in ruling out $e(v) = 0$ as a possible choice the government might consider in some cases.¹⁰

Theorem 3 *Assume that there exists a solution to maximizing (25) subject to the constraints (22), $e(v) \in [0, 1 - g]$ and $k^1 \in [0, \frac{1}{N+1}]$. Let $\hat{e}(v)$ be any such solution. Assume condition 1 holds. Then \hat{e} , as a function of h , assumes almost surely at most two values for each value of $(g_{t,0}, {}_0b_{t,0}^1, {}_0b_{t,0}^2, \eta_t^1)$.*

Proof. See Appendix.

While the proof of Theorem 3 is lengthy and requires dealing with several technical details, the intuition behind the result is much simpler and more general. Both the objective function and the constraint can be described as operators mapping the space of admissible functions e into real numbers; if we consider them as a joint operator, they map an infinite-dimensional space into a two-dimensional space. Since the dimensionality of the domain is much higher than the dimensionality of the range, we expect the operator in general to be non-singular, i.e., local perturbations around any e will be onto a neighborhood of the values attained at e . Whenever this happens, it is then possible to find a local perturbation that improves on the objective function without violating the constraint. Most of the local perturbations that we consider involve increasing e by a differential amount de in a neighborhood of some value h_1 for h and increasing or decreasing it by some multiple of de in a neighborhood of some other value h_2 . If e takes many values as a function of h , this leads to many degrees of freedom and the mapping is nonsingular; but if e takes few values, then the effect of most of these perturbations is the same, as there are few independent choices for the points h_1 and h_2 . Theorem 3 shows rigorously what “few values” means in our case: e can take at most 2 values as a function of h .

Theorem 3 is very important for searching numerically for Ramsey equilibria: it means that the government will at most randomize (or alternate) between two tax rates for each value taken on by the “fundamentals” of our economy.

We now look at the first-order conditions for the Ramsey problem to get some intuition for the results we will discuss more through numerical examples. After some algebra, the first-order condition for the function e can be written as:

$$\begin{aligned}
& \omega^2 \left[(k^2)^{-\gamma} e(v)^{-\gamma} (k^2 + N^1 k^1) - \xi (1 - e(v) - g(v))^{-\sigma} \right] \\
& + \left[(\omega^1 (k^1)^{-\gamma} - \omega^2 (k^2)^{-\gamma}) ({}_0b^1(v) + {}_0\eta^1(v) - k^1 e(v)) \right] e(v)^{-\gamma-1} N^1 \gamma \\
& + \frac{N^1 \gamma \lambda_0}{k^2 I_1} \left[k^1 e(v)^{-\gamma} + \gamma e(v)^{-\gamma-1} ({}_0b^1(v) + {}_0\eta^1(v) - k^1 e(v)) \right] \\
& + \frac{\lambda_0}{I_4 - I_3} \left[\sigma (1 - e(v) - g(v))^{-\sigma-1} + (1 - \sigma) (1 - e(v) - g(v))^{-\sigma} \right] \\
& + \frac{\lambda_0}{I_2 - I_1} \left[\gamma (N^1 {}_0b^1(v) + {}_0b^2(v)) e(v)^{-\gamma-1} + (1 - \gamma) e(v)^{-\gamma} \right] = 0 \quad \text{a.s.},
\end{aligned} \tag{26}$$

¹⁰ Although I conjecture Theorem 3 holds in its present form even without condition 1, I was not able to prove it. It is however possible to prove a slightly generalized form, where e can take 0 as a third value in the states in which $(g_t, N^1 {}_0b_t^1 + {}_0b_{t,0}^2, \eta_t^1 + {}_0b_t^1) = \mathbf{0}$.

where λ_0 is the Lagrange multiplier associated with the constraint (22), and I used the following definitions:

$$I_1 \equiv \int e(v)^{1-\gamma} dm(v); \quad (27)$$

$$I_2 \equiv \int (N^1 {}_0b^1(v) + {}_0b^2(v))e(v)^{-\gamma} dm(v); \quad (28)$$

$$I_3 \equiv \int (1 - e(v) - g(v))^{1-\sigma} dm(v) \quad (29)$$

and

$$I_4 \equiv \int (1 - e(v) - g(v))^{-\sigma} dm(v). \quad (30)$$

Equation (26) has the following interpretation. The first line is the derivative of the Negishi-aggregated utility function of the two agents at the equilibrium wealth levels; these terms would thus give us the undistorted, first-best solution. The second line describes distortions the government introduces to redistribute resources across agents:¹¹ $\omega^1(k^1)^{-\gamma} - \omega^2(k^2)^{-\gamma}$ measures, up to a constant, the value to the government of transferring one unit of resources from the taxpayers to the rentiers; ${}_0b^1(v) + {}_0\eta^1(v) - k^1e(v)$ measures whether for the considered state v the rentiers are consuming more or less than what they are entitled to based on the maturing financial claims they start with. For example, let both $\omega^1(k^1)^{-\gamma} - \omega^2(k^2)^{-\gamma}$ and ${}_0b^1(v) + {}_0\eta^1(v) - k^1e(v)$ be positive. This means that the government would like to transfer more resources from the taxpayers to the rentiers, if it could do so by means of lump-sum transfers. Furthermore, in the state v the rentiers are net sellers of goods. In this case the government has an incentive to reduce aggregate consumption in the state v ; by doing so, the government distorts upwards the price of goods in such states, which increases the wealth of the rentiers.

The remaining terms capture the distortions the government introduces because of the price effects on its budget constraint. These terms are harder to interpret, as more effects come into play. It is easy to show however that λ_0 measures the marginal benefit of being able to substitute lump-sum taxes for distortionary taxes.

5 Numerical Examples

In this section we analyze the characteristics of the Ramsey equilibria of the two-class economy by looking at some numerical examples.

In all the examples I present, I chose the following parameters: $\gamma = 2$, $\sigma = 1.1$, $\beta = 0.95$, $N_1 = 1$.

I tried different values for all of these parameters, but this did not alter the results significantly. Only the magnitude of changes in the interest rate will be greatly amplified by choosing higher values for the risk aversion.

In each case we can trace the entire Pareto frontier by parameterizing $(\omega^1, \omega^2) = (\alpha, 1 - \alpha)$, $\alpha \in [0, 1]$.

¹¹This term corresponds to what Persson and Svensson [15] call an “optimal intertemporal tariff” in the open economy setup.

5.1 Example 1: No Government Spending

Let $g_t \equiv 0 \quad \forall t \geq 0$, ${}_0b_t^i \equiv 0 \quad \forall t \geq 0 \quad i = 1, 2$, ${}_0\eta_t^i \equiv \hat{\eta}_0^i \quad i = 1, 2$. In this example the government has no public spending to finance and no debt to repay (nor credit to distribute). Furthermore, the only outstanding private claims are annuities that pay a fixed amount every period.

In this setup the government has no need to raise taxes ever; in a representative-agent model, the government would achieve a first best by setting $\tau_t \equiv 0 \quad \forall t \geq 0$. With this tax policy, we would have:

$$c_t^1 = \hat{\eta}_0^1 \quad \forall t \geq 0; \quad (31)$$

$$u_c^2 = u_x^2 \quad \forall t \geq 0 \quad (32)$$

and

$$c_t^2 + x_t^2 = 1 + \hat{\eta}_0^2 \quad \forall t \geq 0. \quad (33)$$

Equation (32) comes directly from (7) and describes the allocation of resources between leisure and consumption given that labor income is not taxed.

Equation (33) states that in each period the sum of each agent's consumption and leisure is equal to her time endowment and the income (which may be negative) from the annuities she holds. The price system in this competitive equilibrium is $p_t = 1 \quad \forall t \geq 0$. Given this price system, the choice of a constant profile of consumption and leisure implied by (31)–(33) is optimal, as it is implied by (8); furthermore the budget constraints of each agent (4) are satisfied, and so is the market clearing condition (6), provided $\hat{\eta}_0^1 = -\hat{\eta}_0^2$, which is required by (3).

This competitive equilibrium is associated with a function $e(v)$ which is constant. By forming the Lagrangean of (25) subject to (22), it is easy to check that the choice of a constant e satisfies the first-order conditions for an optimum even with heterogeneous agents.

While the no-tax solution always satisfies the first-order conditions in this example, it is not the optimal solution when the rentiers have a sufficiently large weight in the government. To understand why this is the case, we concentrate on the welfare of the rentiers.

Let us consider deviations from the no-tax policy that involve one tax rate in all even periods and another one in all odd periods.¹² Figure 1 shows what happens in this case. The no-tax solution is represented by the point C_0 : the rentiers consume in each period exactly the amount of resources they are owed by the taxpayers, i.e., $c_t^1 = \hat{\eta}_0^1 \quad \forall t \geq 0$. The line from A_0 to B_0 represents the Arrow-Debreu budget constraint of the rentiers in the no-tax policy: since $p_t = 1 \quad \forall t \geq 0$, its slope is $-\frac{1}{\beta}$, as we assume the first period to be even (period 0). The indifference curve through C_0 is tangent to the budget constraint, reflecting the optimality of C_0 when the pricing kernel is constant. Suppose now the government varies the tax rates in odd and even periods. The rentiers are not affected directly by the change in the tax rate; they are only affected indirectly, as different tax rates lead to different relative prices between odd and even periods. As a consequence the budget constraint of the rentiers rotates, e.g., to A_1, B_1 ; however, it still goes through C_0 , as $c_t^1 = \hat{\eta}_0^1 \quad \forall t \geq 0$ is always feasible. Since the utility function is assumed to be strictly concave, the indifference curves are strictly convex, and the rentiers are strictly better off when the relative

¹²As we argued previously, this policy is equivalent to a policy that sets two different constant tax rates depending on whether h is larger or smaller than $\frac{\beta}{1+\beta}$.

price of goods varies in either direction: the new choice is C_1 , which lies on a higher indifference curve. This welfare improvement is locally of second-order magnitude, and this is why the no-tax policy satisfies the first-order conditions.

By taxing labor in odd periods and subsidizing it in even periods (or vice versa), the government generates an artificial scarcity of some goods with respect to others, and this is beneficial to the rentiers. Of course, this policy is very costly to the taxpayers. In the economy we consider, if the government were allowed to transfer resources directly between the two agents, it would never choose to distort prices, as a constant consumption stream for both agents would be Pareto efficient. The taxpayers, therefore, pay both for the gains of the rentiers and for the distortions introduced by taxes and subsidies. These losses are also of second order in a neighborhood of the no-tax policy.

To compute when the government would resort to randomization, I solved numerically for the optimal policy. Based on my computations, choosing different tax rates when all exogenous variables (government spending and maturing coupons) are the same is a very costly way of redistributing wealth among the agents. For instance, consider a case in which ${}_0\eta_t^1 = -{}_0\eta_t^2 = 1/3 \quad \forall t \geq 0$ and $\xi = 3^{\gamma-\sigma}$. In this case taxpayers and rentiers reach the same consumption when the government implements a no-tax policy. Deviating from this policy becomes desirable for the government only when the Pareto weight of the rentiers exceeds 0.65, i.e., approximately when the government gives to the rentiers twice the weight attributed to the taxpayers.

5.2 Example 2: Constant Government Spending

This example is identical to Example 1 except that we consider now a positive level of government spending: Let $g_t \equiv \hat{g} \quad \forall t \geq 0$, ${}_0b_t^i \equiv 0 \quad \forall t \geq 0 \quad i = 1, 2$, ${}_0\eta_t^i \equiv \hat{\eta}_0^i \quad i = 1, 2$. In this example there is still no uncertainty, public spending is constant, and private agents are still trading only in annuities, so there is still a solution to the first-order conditions that implies the same allocation in all periods. The only difference with Example 1 is the fact that now the government has to resort to distortionary taxation to finance a stream of spending.

When the Ramsey equilibrium of this model implies a competitive equilibrium with a constant allocation over time, we can solve for the allocation using the following equations:

$$\frac{u_x^2(c^2, x^2)}{u_c^2(c^2, x^2)} = 1 - \tau; \quad (34)$$

$$c^2 + \hat{\eta}_0^1 = (1 - \tau)(1 - x^2); \quad (35)$$

$$c^1 = \hat{\eta}_0^1 \quad (36)$$

and

$$\tau(1 - x^2) = g. \quad (37)$$

In this example we have one more equation than we had in the previous one, since we need to find the tax rate that allows the government to finance its expenditures in each period.

In the tax-smoothing policy the government sets a constant tax rate exactly sufficient to raise revenues covering public spending; the consumption of each private agent is equal to post-tax

labor income plus the coupons from the annuities she owns (which may be a negative amount if she is short on the annuities).

The same considerations as in Example 1 hold when the government considers a deviation from the tax-smoothing policy. Figure 1 still represents what happens to the welfare of the rentiers. It is again true that the rentiers receive a second-order benefit and the taxpayers a second-order loss from local deviations from tax smoothing.

We now turn to more-interesting examples, where government spending is not constant. In this case, deviations from the policy that minimizes distortions will bring about costs that are locally second-order but distributional benefits that are potentially first-order, so government policy will be affected to a much larger extent.

The algorithm I used in the numerical simulations proceeds as follows. Given an example, we first specify what are the possible values that

$$\tilde{v} \equiv (g_{t,0} b_{t,0}^i \eta_{t,0}^1 \eta_t^2) \quad (38)$$

can take at any date and in any state. We will restrict this to be a finite number of possibilities. Let (v^1, \dots, v^M) be the possible values for \tilde{v} .

In this case, the integrals in (25) and (22) become finite sums.

Then the algorithm computes the first-order conditions for the problem of maximizing (25) subject to (22). The algorithm solves for the best policy that does not involve randomization between tax rates when the value of \tilde{v} is the same; it then checks whether the second-order conditions are satisfied locally, and it also checks whether the optimum achieved is preferable to all points on a (coarse) grid spanning all the admissible 2-point random policies. I present here the solutions only for the range in which the optimal policy does not involve randomization; this happens if the Pareto weight of the rentiers is not too large.¹³ The policies the government would follow in the range where the Pareto weight of the rentiers is very high are empirically implausible.

In the following example we illustrate what happens when government spending is not constant.

5.3 Example 3: “France vs. Britain”

This example is suggested by an observation in Sargent and Velde [18] about war financing in France and Britain in the 17th and 18th centuries. The clearest description of the difference between the two regimes is their quote of Montyon, a senior civil servant in the French Finance Ministry in 1770s. He pointed out that

Great Britain finances by taxation neither all nor part of the costs of war, it finances them by loans (...). In wartime it is our habit to increase taxes (...). Indeed in wartime the country suffers enough from the labor withdrawn from agriculture and manufactures to be sent into the army, the navy, and into the production activities necessitated by war.

¹³When randomization is desirable, I found that in many examples there is no optimal policy: the government will achieve the supremum by driving the tax rate to 1 with probability ever closer to 0.

Montyon wrote to express his dissatisfaction with the French policies. As Sargent and Velde argue convincingly, that dissatisfaction was one of the factors that led eventually to the French revolution.

France and Britain had very different political regimes at that time; the noble class had much more clout in France than in Britain. In terms of our model, we interpret this as meaning that the rentiers had a higher Pareto weight in France than in Britain.

In this numerical example the government has thus to finance a war. I consider a single war, starting in period 2 and lasting 6 periods (years).¹⁴ Peacetime government spending is 0.2, whereas wartime spending is 0.4.¹⁵

Figures 2-3 show the Ramsey allocation depending on the Pareto weight of the rentiers in the range where the government does not wish to randomize its policy.¹⁶ Note that the figures plot the allocation during peacetime and wartime only: the peacetime allocation is going to be the same in the periods preceding the war and in all periods following the war, and the wartime allocation will be the same in all war periods.

The main conclusion that these pictures suggest is that the optimal way of financing a war is definitely influenced by the attitude of the government with respect to redistribution. Furthermore, the effect is consistent with the mentioned pattern of war finance in Britain and France: a government run by the taxpayers will run large deficits in wartime, whereas the preferred policy for the rentiers involves a large increase in taxes during the war.

The intuition of this result is the following. Sooner or later, a war will have to be financed in this model by raising funds through taxes. This means that, implicitly (i.e., through the government) or explicitly (i.e., through direct transactions) the taxpayers will have to borrow from the rentiers to smooth their consumption stream. When the government “sides with the taxpayers,” it will try to get the best possible deal to raise the funds it needs. This can be achieved by distorting prices so that the price of the consumption good is kept low during wartime relative to peacetime. To achieve this, the best strategy for the government is to run a huge deficit by cutting taxes during the war. The deficit is then covered by increasing taxes during peacetime and gradually repaying the large debt accumulated during the war. When the government is influenced more by the rentiers, it will respond to a war by running a much smaller deficit and having a higher tax rate during wartime. The scarcity of goods will then be more acute. The rentiers will pay a lower interest rate on the war chest being built before the war and will demand a higher interest rate on the debt they subscribe during wartime.

The vertical dotted line shows the policy chosen by a government whose ratio of the Pareto weights coincides with the ratio of the marginal utilities of the two agents. For such a government, there is no incentive at the margin to redistribute wealth; its policy is thus just dictated by efficiency considerations. For the preferences we specified, this policy implies higher taxes in

¹⁴We could have considered a stochastic process for wars and peaceful periods, but this would not have altered the results. The quantitative results are actually very similar to the ones of a model where there is a Markov process for war and peace, with the average war lasting 6 periods, peace as the initial condition, and the invariant distribution attributing probability .7 to peace.

¹⁵The other parameters I chose are the following: $\xi = 4^{\gamma-\sigma}$, ${}_0\eta_t^1 = 0.25 \quad \forall t \geq 0$ a.s..

¹⁶The government will randomize its policy when the Pareto weight of the rentiers is really large, but this range does not produce empirically plausible predictions. When the government has such a high desire to redistribute wealth from the taxpayers to the rentiers, it is likely that it will try to adopt alternative, less-costly schemes.

wartime.¹⁷ In our case, the “neutral” wartime deficit would thus be only 30% of the cost of the war; the associated tax rates are 33% in peacetime and 43% in wartime. When the taxpayers have all the weight, the government finances 71% of the war by issuing debt, and the tax rate is lowered sharply when a war occurs: the tax rate in this case is 39% in peacetime and only 32% in wartime.

We saw that in Examples 1 and 2 the optimal tax policy is unaffected by the different Pareto weights the government can attach to the agents over a large range. That happened because in the Ramsey policy neither agent was “borrowing” in some state and lending in some other state; the rentiers were consuming exactly their “endowment” stemming from the maturing coupons on the annuities, and the taxpayers were consuming their proceeds from labor supply net of taxes and the coupon payment on the annuities. Because of this, deviating from the Ramsey policy had only second-order distributional effects. In Example 3, this is no longer the case. The rentiers are now lending to the taxpayers in wartime and being repaid in peacetime. Figure 3 shows that their consumption is below the level implied by their coupons in wartime, and it is above in peacetime. A change in the relative tax rates the government applies during the war and in peace brings thus first-order distributional effects. Because of this, the optimal tax rates change when the Pareto weights change.

The rightmost part of the graphs, showing the Ramsey allocation when the Pareto weight of the rentiers is very high, gives results that may be empirically irrelevant but illustrate well the forces at work. When the government has a strong desire for redistributing wealth in favor of the rentiers, it may even run a surplus during wartime, raising the tax rate to a level that will cause extreme scarcity of goods during wartime. The taxpayers will then have to borrow at extremely bad terms, giving up substantial rights to future consumption.

At the extreme opposite, when $\omega_1 = 0$, we have the open-economy solution, in which the taxpayers are interpreted as being the domestic agents and the rentiers are the foreign agents. As Persson and Svensson [15] suggest, in an open economy a negative shock such as a war is absorbed by a trade deficit. In deciding the optimal tax plan, however, the government takes into account that a larger trade deficit implies a worse relative price of current imports against future exports that will offset the current deficit. It therefore has an incentive to abandon the tax plan that minimizes tax distortions to reduce strategically the size of the trade deficit and improve the terms of trade. Accordingly, the optimal plan calls for the government to reduce taxes and run a large budget deficit during the war to stimulate output.

We have so far argued that the volume of public borrowing in France and Britain was consistent with the qualitative predictions of the model. It would be interesting to examine whether the other predictions are consistent with the data. In particular, we would predict that the real interest rate during the war was higher in France than in England, both for private and for public loans; we would also predict that, compared with Britain, France had a larger flow of financial resources from the rentiers to the taxpayers in wartime and a much larger flow from the taxpayers to the rentiers in peacetime.

While in principle these hypotheses are empirically testable, in practice the available data do not allow sharp conclusions.

Velde and Weir [22] show that France paid substantially higher interest rates than Britain did; however, France also defaulted frequently on its debt. They also show that observed interest rates

¹⁷The tax choice that minimizes distortions requires equality of the marginal tax distortions across states. Depending on the preferences, this may imply higher or lower taxes during wartime. See Lucas and Stokey [13].

on government debt oscillated mainly in anticipation of government defaults. The appropriate comparison for our purposes should thus adjust the interest rates for the expected defaults. There seems to be evidence that the premium paid by France was more than enough to offset the default risk, with the exception of the Law Affair.¹⁸ Further research is required though to study whether the interest rates, net of the default premium, were higher in France than in Britain in wartime and lower in periods preceding a war.

As for the private credit markets, there are very few studies on microeconomic data that would allow us to distinguish flows between social classes. Rosenthal [17] studies credit in a rural area, where the shocks to local agriculture seem to be much more important than wars or other government intervention. Wars seem much more important for Paris,¹⁹ but yearly data have been estimated only for aggregate series. The aggregate volume of credit is not a good measure of the series we are interested in because the net position of the different social classes on the market was about even.²⁰

5.4 Example 4: “Taxation over the Business Cycle”

The following example looks at the optimal policy in presence of small shocks, such as the ones that may arise over the business cycle. We are interested in this example because it allows us to evaluate the impact of a balanced-budget constraint that requires the government to raise enough taxes in each period to pay for current spending.

To model a business cycle, we should slightly alter our setup and allow for shocks on the income (or productivity) process. In a recession, a given tax rate would then lead to smaller output and smaller revenues. For our purposes, this has an effect that is equivalent to an increase in government spending, which reduces output available for private consumption and increases the need for the government to raise revenues.²¹ We assume thus that government spending follows a Markov process with support on two values: 0.23 in booms and 0.26 in recessions. We calibrate the transition probabilities so that the average boom lasts 4 periods (years) and the average recession lasts 2 periods.²² We assume the economy starts in a boom period, with no government debt.

The results from this simulation are presented in Figures 4-5.²³ While the results are qualitatively similar, the response to such a small shock is quite modest over most of the range. At the policy chosen by the government that is neutral with respect to distributional issues (the “neutral” policy), the tax rate is 33.3% in booms and 34.6% in recessions, and the deficit is about

¹⁸Cf. Hoffman, Postel-Vinay and Rosenthal [11].

¹⁹Cf. Hoffman, Postel-Vinay and Rosenthal [10].

²⁰Had this not been the case, we could have estimated the flows from measures of aggregate volume of credit. For example, suppose that aristocrats were mainly lending, whereas the bourgeois were mainly borrowing. In this case, we would have expected the bourgeois to borrow even more in wartime, thereby increasing the size of the aggregate volume of private credit. The reverse effect would have arisen if the bourgeois had been the lenders.

²¹I ran a numerical simulation where productivity was allowed to vary and the results were similar; I stick to a government-spending shock in this version in order to be consistent with the notation in the rest of the paper.

²²All other parameters are the same as in Example 3.

²³Labor supply is higher in “recession” than in boom because in our case a recession is interpreted as an increase in government spending. If we considered an example where the recession is dictated by a reduction in productivity, the labor supply would be lower in recession than in boom; however, all the other variables would behave in the same way as in Figures 4 and 5, and especially the optimal tax policy would have the same characteristics as the one described in Figure 4.

one third of the spending increase; the policy preferred by the taxpayers implies a deficit that is twice the size of the neutral policy; the tax rate for this policy is 34.1% in booms and 33.1% in recessions.

While the choice of an intertemporal tax policy cannot do much to benefit the taxpayers, it still can hurt them substantially. However, this would happen only if the government ran large deficits during booms.

Let us now consider the effect of a balanced-budget restriction on the government. As we see from Figure 4, compared with the neutral policy, a balanced-budget policy slightly favors the rentiers. If business cycle shocks were the only source of fluctuations in government revenues and expenses, a balanced-budget requirement would imply very modest costs and benefits for the agents in the economy: in our example, the rentiers would achieve a benefit of 0.02% of their consumption, while the taxpayers would incur a loss of about 0.03% of their consumption.

However, the government balance may be subject to much larger shocks from other sources. For instance, the aging of the population implies a large increase in spending for pensions and publicly-subsidized medical expenses. These shocks evolve much slower than wars or the business cycle, but they have an impact that may be as large as those caused by wars. It is therefore likely that the intertemporal choices on taxation may imply a much larger distributional conflict than the small business-cycle shocks. An important feature of these alternative shocks is that they often involve transfers that are at least partly discretionary. To properly consider them, it will thus be necessary to endogenize government spending, which is left for future research.

6 On Uniform Commodity Taxation

In a seminal paper on optimal taxation in a static environment, Atkinson and Stiglitz [2] (A-S from now on) provide conditions under which it is inefficient to distort relative prices for redistribution purposes.

The purpose of this section is to contrast their results with ours.

Due to the presence of complete markets and the Ramsey assumption on the timing of the government policy, we can view the consumption good at different dates and in different states of the world as many different commodities, to which the results on optimal taxation in a static framework can be applied. A-S show that access to a sufficiently flexible income tax schedule is enough to guarantee optimality of a uniform commodity tax if preferences are (weakly) separable in leisure and the other goods. When the subutility from the other goods is the same for everybody and it is homothetic, a uniform commodity tax is optimal in their environment. The preferences we assumed in sections 4 and 5 satisfy this requirement.

While there is no reason to think that uniform commodity taxation would be optimal in the general environment of section 3, the additional assumptions made in sections 4 and 5 lead us much closer to an environment where it is optimal to tax all goods at the same rate. As I explain in more detail in the appendix, our income tax plays both the role of a tax on a factor of production and that of a tax on a commodity; this is due to the fact that there is no substitutability in production between goods and/or factors in different periods and states of nature.

The appendix shows that a constant tax on labor income is not optimal even in the environment I described in sections 4 and 5 for the following reasons.

- (i) In A-S, leisure is just one good. In our dynamic environment, there are many types of

leisure, one for each time and state of nature. The taxpayers are endowed with one unit of each type of leisure, but consume different amounts of the various types when the labor supply is not constant. Because of this, a uniform tax on all factors of production will not be optimal. The deviations from a constant tax that we obtain through this channel are similar to what happens in L-S and in other papers that have looked at dynamic optimal taxation in a representative-agent framework.²⁴

- (ii) In A-S, no private agent starts with an initial endowment of any good except her time. This is sometimes justified by assuming that the government can seize the initial endowment of any good, except time; however, this hypothesis is not equivalent to the former and would not always imply uniform commodity taxation. We assumed instead that the government and the private agents start with some initial level of financial claims that cannot be taxed away directly. If we had not assumed this, in our case the government would have been able to achieve lump-sum redistribution merely by taxing or subsidizing the initial credit or debit positions. Differences in the initial endowment of goods over time and across states play a very important role in our analysis: the government will distort prices to increase the value of the initial endowment of the group it wants to favor. In our setup, the initial endowment is given by the time of the taxpayers and the initial financial claims for both groups.

The appendix also shows that ruling out lump-sum taxes and transfers *is not* a reason for our deviations from a uniform tax. My paper treats the two classes of households as homogeneous and does not deal with distribution issues *within* any given class; lump-sum taxes and transfers would affect the distribution within classes much more than the price distortions I study here, and this is the reason I focus on price distortions as the main channel of redistribution *across* different classes. Nonetheless, the qualitative results are similar if lump-sum transfers are allowed.²⁵ Lump-sum transfers would be desirable as a way of transferring resources from the taxpayers to the rentiers, and would be used if the government wished to redistribute resources from the taxpayers to the rentiers. However, due to the reasons in (ii), it would still be optimal for the government to deviate from uniform commodity taxation and use price distortions whenever the lump-sum transfers have to be financed through distortionary taxes. When lump-sum transfers are admitted, the magnitude of the price distortions that the government optimally uses to redistribute from the taxpayers to the rentiers is smaller, and we would not observe the extreme results that we get in section 5 for this case.

7 Time Consistency of the Optimal Fiscal Policy

In the previous sections we have characterized Ramsey equilibria that arise when the government commits to its policy at time 0 and is never allowed to revise it. In this section we consider a

²⁴A-S's result in the case of heterogeneous agents is connected to what they established in a previous paper (Atkinson and Stiglitz [1]), where they provided conditions that lead to uniform commodity taxation in a representative-agent environment where lump-sum taxes are ruled out. Separability and homotheticity play there the same role they play in their later paper.

²⁵The introduction of lump-sum taxes would instead have a large impact: over a large range of parameter values, the government would be able to raise revenues much more efficiently. However, even in this case some price distortions would be optimally used.

different environment, where the government chooses its policies sequentially. We provide some sufficient conditions that ensure time consistency of the Ramsey policy when the government chooses sequentially. This exercise is similar to section 2.4 in L-S and to sections III-V in Persson and Svensson [15]. As in their papers, we maintain the assumption that the government is committed to repay its debt obligations, but it can choose sequentially the tax rate on labor income.

In a representative-agent, non-monetary economy, L-S show that a sufficient condition for ensuring time consistency of the Ramsey policy is that the government be allowed to choose an appropriate structure for its debt. At each point in time, the government will have to readjust its portfolio of state-contingent bonds across different maturities and different risk profiles.

Persson and Svensson [15] extend that theory to an environment with an open economy. They show that, when private capital is perfectly mobile, the maturity structure of government debt that ensures time consistency will depend on the maturity structure of net private claims against the rest of the world.²⁶

Our result is similar to that of Persson and Svensson. In general, there is a structure of government debt that ensures time consistency. However, this structure depends on who is holding the debt and on the structure of private claims traded by the taxpayers and the rentiers. Time consistency will thus require the government to react to the arrangements in the private capital markets.

Since we are only interested in providing sufficient conditions for time consistency, we will compute a new Ramsey equilibrium starting from period 1 and check whether this coincides with the time-0 Ramsey equilibrium in any possible state of the world at time 1. If this happens, then the Ramsey equilibrium is time consistent, as argued in L-S.²⁷

The first step for our analysis is to describe the evolution of the economy when policy decisions are taken sequentially. We therefore introduce a dynamic budget constraint for the agents and the government. In our previous sections we relied only on the single budget constraint at time 0; this was justified because dynamics were irrelevant in an Arrow-Debreu economy, where we could simply reinterpret the same good at different dates as many different goods, and the tax rates in different dates and states as tax rates on different production processes leading to different goods.

We will assume that there are complete contingent markets at each date and in each state; this is a much stronger assumption than market completeness. It implies that at any moment the entire set of contingent claims is open. While this provides the private agents with redundant securities, we will see that this is crucial for the government to be able to guarantee time consistency of the Ramsey equilibrium. Trading in financial markets happens in every period after g_t (but before h is revealed at time 0), i.e., at the same time as production takes place.

The sequential budget constraint of type i agents at time 0 is the following:

$$p_0[c_0^i - \eta_0^i - w^i(1 - \tau_0)(1 - x_0^i)] - E_0 \sum_{t=1}^{\infty} \beta^t \{p_t[b_t^i + \eta_t^i - b_t^i - \eta_t^i]\} = 0. \quad (39)$$

²⁶The result I mentioned is contained in the section where perfect capital mobility is allowed. This is the relevant case for our analysis, since the rentiers (the foreigners, in the limiting case of an open economy) and the taxpayers (the domestic agents) trade in perfect capital markets.

²⁷It is possible to show that this condition is necessary and sufficient for the Ramsey equilibrium to be a Markov equilibrium.

Equation (39) states that the time-0 value of the financial assets maturing at time 1 or later that agent i owns at the end of period 0 must exceed their value at the beginning of time 0 by an amount equal to gross savings at time 0, defined as the sum of labor income and maturing assets minus current consumption. Each agent will also face a transversality condition, requiring:

$$\lim_{s \rightarrow +\infty} \sum_{t=s}^{\infty} \beta^t p_t [{}_s b_t^i + {}_s \eta_t^i] = 0 \quad \text{a.s.} \quad (40)$$

This is equivalent to requiring the agent to meet in each period the budget constraint in the Arrow-Debreu form. For period 1, this requires:

$$E_1 \sum_{t=1}^{\infty} \beta^t p_t [c_t^i - {}_{-1} b_t^i - {}_{-1} \eta_t^i - w^i (1 - \tau_t)(1 - x_t^i)] = 0 \quad \text{a.s.} \quad (41)$$

In an Arrow-Debreu economy we only require clearing in the market for state-contingent goods. Since we now deal explicitly with the trade in financial assets, we also need a market-clearing condition in the markets for assets:

$$\sum_{i=1}^N {}_1 \eta_t^i = 0 \quad \forall t \geq 1 \forall g^t. \quad (42)$$

The government will also have a dynamic budget constraint and a transversality condition. However, as long as the budget constraints and the transversality conditions of all other agents are met, the government ones will be met automatically, so they are redundant for our purposes.

Let $\{(c_t^1, c_t^2, x_t^2)\}_{t=0}^{\infty}$ be the Ramsey equilibrium allocation as of time 0. Suppose now that the government is given at time 1 the (unexpected) opportunity to revise its policies. We wonder under what circumstances the optimal choice would be not to revise the previous plan. To answer the question, we compute the first-order conditions as of time 1, as well as the constraints the government faces. We inquire whether it would be possible for the time 0 Ramsey allocation to satisfy the constraints and the first-order conditions.

As in L-S, we will rely on the first-order conditions to look at the optimal policies as of period 1. From the previous sections we know these conditions may not be sufficient. In the numerical examples, we therefore check whether the solution we find is indeed an optimum.

The Ramsey problem the government would face as of time 1 can be rewritten exactly as we did in section 3 for the time-0 problem. There will be a new measure for government spending and coupons: because one period is passed and possibly some uncertainty has been revealed, the process for government spending the economy starts from is now different; coupon payments on financial claims may be different for the same reason, and they may also be different because the government and the private agents may have reshuffled their portfolios during period 0.

The new vector describing the economy is ${}_1 v \equiv (h, g_{t,1}, {}_1 b_{t,1}^1, {}_1 b_{t,1}^2, \eta_{t,0}^2, {}_1 b_{t,0}^1, {}_1 b_{t,0}^2, \eta_t^2)$. If we were simply solving for the time 1 Ramsey problem, the time-0 financial commitments $({}_0 b_{t,0}^1, {}_0 b_{t,0}^2, \eta_t^2)$ would not matter. However, our goal is to establish when the time-0 Ramsey allocation coincides with the solution to the Ramsey problem at time 1. Since the optimal allocation as of time 0 depends on $({}_0 b_{t,0}^1, {}_0 b_{t,0}^2, \eta_t^2)$, it is useful to retain these variables in our “state.” Their effect at time 1 will of course have to be mediated through other variables, notably the financial commitments in place at the beginning of period 1, i.e., $({}_1 b_{t,1}^1, {}_1 b_{t,1}^2, \eta_t^2)$.

The new relevant measure as of time 1 will thus be defined by:

$$m_1(A) \equiv \sum_{t=0}^{\infty} \beta^t \text{Prob}\left((g_{t,1} b_{t,1}^1 b_{t,1}^2 \eta_{t,1}^1 b_{t,1}^1 b_{t,1}^2 \eta_{t,1}^2, h)\right) \in A \quad \forall A \in \mathcal{G} \times \mathcal{B}^7. \quad (43)$$

The Ramsey problem as of time 1 is thus:

$$\begin{aligned} \max_{e, {}_1k^1} & \left[\omega^1 N^1 ({}_1k^1)^{1-\gamma} + \omega^2 (1 - N^1 {}_1k^1)^{1-\gamma} \right] \int \frac{e({}_1v)^{1-\gamma}}{1-\gamma} dm_1({}_1v) \\ & + \omega^2 \int \frac{(1 - e({}_1v) - g({}_1v))^{1-\sigma}}{1-\sigma} dm_1({}_1v) \end{aligned} \quad (44)$$

s.t.

$$\begin{aligned} \xi (1 - N^1 {}_1k^1)^\gamma & \left[\int (1 - e({}_1v) - g({}_1v))^{-\sigma} dm_1({}_1v) - \int (1 - e({}_1v) - g({}_1v))^{1-\sigma} dm_1({}_1v) \right] \\ & = \int e({}_1v)^{1-\gamma} dm_1({}_1v) - \int e({}_1v)^{-\gamma} (N^1 {}_1b^1({}_1v) + {}_1b^2({}_1v)) dm_1({}_1v), \end{aligned} \quad (45)$$

with

$${}_1k^1 \left[\int e({}_1v)^{1-\gamma} dm_1({}_1v) \right] = \left[\int ({}_1b^1({}_1v) + {}_1\eta^1({}_1v)) e({}_1v)^{-\gamma} dm_1({}_1v) \right]. \quad (46)$$

Equation (46) is the budget constraint of the rentiers. When we solved the problem at time 0, we substituted it into the objective function. In this case it will be more convenient to treat it as a constraint with its Lagrange multiplier. It is also more convenient not to take the logarithm in the constraint (45).²⁸

The first-order conditions for this problem are:

$$\begin{aligned} & \left[\omega^1 N^1 ({}_1k^1)^{1-\gamma} + \omega^2 (1 - N^1 {}_1k^1)^{1-\gamma} \right] e({}_1v)^{-\gamma} - \omega^2 \xi (1 - e({}_1v) - g({}_1v))^{-\sigma} \\ & + \lambda_1 \xi (1 - N^1 {}_1k^1)^\gamma \left[\sigma (1 - e({}_1v) - g({}_1v))^{-\sigma-1} + (1 - \sigma) (1 - e({}_1v) - g({}_1v))^{-\sigma} \right] \\ & + \lambda_1 \left[(1 - \gamma) e({}_1v)^{-\gamma} + \gamma (N^1 {}_1b^1({}_1v) + {}_1b^2({}_1v)) e({}_1v)^{-\gamma-1} \right] \\ & - \gamma \mu ({}_1b^1({}_1v) + {}_1\eta^1({}_1v)) e({}_1v)^{-\gamma-1} - (1 - \gamma) \mu {}_1k^1 e({}_1v)^{-\gamma} = 0 \quad \text{a.s.} \end{aligned} \quad (47)$$

and

$$\begin{aligned} & \left[\omega^1 N^1 ({}_1k^1)^{-\gamma} - \omega^2 N^1 (1 - N^1 {}_1k^1)^{-\gamma} \right] {}_1I_1 + \lambda_1 \xi N^1 \gamma (1 - N^1 {}_1k^1)^{\gamma-1} \\ & \cdot \left[\int (1 - e({}_1v) - g({}_1v))^{-\sigma} dm_1({}_1v) - \int (1 - e({}_1v) - g({}_1v))^{1-\sigma} dm_1({}_1v) \right] - \mu {}_1I_1 = 0, \end{aligned} \quad (48)$$

with ${}_1k^2 = 1 - N^1 {}_1k^1$ and

$${}_1I_1 \equiv \int e({}_1v)^{1-\gamma} dm_1({}_1v). \quad (49)$$

²⁸The formulation we used for the time-0 problem is more convenient when we have to solve for e ; now we want instead to establish when a given solution for e can be supported for an appropriate choice of other variables. The form we choose here is convenient because the first-order conditions and the constraints are linear in $({}_1b_{t,1}^1, {}_1b_{t,1}^2, \eta_t^2)$.

Here λ_1 is the Lagrange multiplier associated with the constraint (45), and μ is the Lagrange multiplier associated with (46).

In our model, as in L-S, the potential for time inconsistency stems from the distortions the government introduces. The distortions that it is optimal to introduce at time 1 are in general different from the ones that were optimal at time 0.

In L-S the only potential source of such distortions arises from the budget constraint of the representative agent. If the structure of debt is not chosen appropriately, the government will in general set a new tax policy to distort the price system in its favor. With heterogeneous agents we may have a second source due to the constraint that in a competitive equilibrium all agents must have the same marginal rates of substitution across goods in different dates and states. The government can only distort the price system faced by each of the private agents in the same way; as this leads to a worse outcome, the government will in general be tempted to revise its policy to achieve further redistribution.

While there are two sources of time inconsistency in our setup, the government's tax instruments remain the same as in L-S. This will be important in generalizing their findings.

We are now ready to state the main result of this section.

Result 1 To ensure time consistency of the Ramsey policy, it will be sufficient in general for the government to be able to adjust during time 0 one of the three bilateral positions ${}_1b_t^1$, ${}_1b_t^2$, or ${}_1\eta_t^2$ after having observed the two others.

As a first step in establishing our result, notice that we need only to check whether the government will optimally choose the same function e at time 1 as it did at time 0. If this happens, then the allocation, the price system, and the government policy will be the same. To see this, we will repeat the same steps which led us to the formulation of the time-0 Ramsey problem as one of maximizing (25) subject to (22).

We start by observing that the price system will have to be the same if e is the same, because the price system is still determined by equation (17). Given that the price system is the same as the one implied by the time-0 Ramsey equilibrium, the consumption of the rentiers will be the same as well, since the optimal plan of the rentiers is uniquely determined by the price system. It follows thus that $k^1 = {}_1k^1$, which in turn implies that $k^2 = {}_1k^2$ and that the consumption of the taxpayers must be the same. Equations (20) and (21) establish then that leisure and the tax rate must also be the same as in the time-0 Ramsey equilibrium.

Up to the first-order conditions, all the government needs to do to guarantee time consistency of its policy is to trade in financial claims, or affect private trade in financial claims, in such a way that equations (45), (47), and (48) are satisfied at time 1 almost surely.

Equation (45), as (22), is derived from the budget constraint of the private agents. Equations (39) and (40) imply that the agents will trade in such a way that their present-value budget constraint as of time 1 (equation (41)) is always satisfied; it follows that, if the government does not unexpectedly change its choice of e , the constraints (45) and (46) will automatically be satisfied. In order to have an equilibrium, however, the government must pick ${}_1b_t^1$ and ${}_1b_t^2$ so that its budget constraint is satisfied, which is given by the following:

$$E_1 \sum_{t=0}^{\infty} \beta^t \{ p_t [g_t + {}_1b_t^1 + {}_1b_t^2 - \tau_t(1 - x_t^2)] \} = 0. \quad (50)$$

In the representative-agent setup of L-S, there is a unique debt structure that ensures that the time-0 Ramsey allocation satisfies the first-order conditions as of time 1 as well. This structure is such that it leaves the government with exactly no incentives (at the margin) to revise the policy laid out at time 0.

By looking at equation (47), we can immediately notice two main differences with the representative-agent economy. First, in our case the government can try to influence more than one debt structure. In general, all the bilateral positions ${}_1b_t^1$, ${}_1b_t^2$ and ${}_1\eta_t^2$ will matter for the first-order condition of the government. Second, if some of these variables are beyond the control of the government, it is necessary for the government to be able to observe or correctly anticipate them and to react accordingly in order to assure time consistency.

Now, let λ_1 and μ be given. Then equation (47) is a linear equation in each of the bilateral positions, and can be solved for each position given the others, provided the coefficient on each position is nonzero. We will later discuss in more detail when these coefficients might be zero. We can thus derive one of the bilateral positions as a function of λ_1 , μ , and the other positions. However, this solution will not satisfy the budget constraint of some of the agents and the first-order condition (48) for generic values of λ_1 and μ . We will then need to find a value for the multipliers λ_1 and μ such that the budget constraints and (48) are satisfied. This requires solving two nonlinear equations in two unknowns. In the computer simulations I ran (see below), such a solution always existed.²⁹

To illustrate these results more concretely, we go back to some of the numerical examples we looked at previously. We will assume that the government takes as given the transactions among private agents, and markets an appropriate mix of government debt to be sold exclusively to the rentiers.

We will thus proceed as follows. We will first specify the values for ${}_1\eta_t^2$. We then substitute these values and the Ramsey allocation into equation (47); we jointly solve (47), (48) and (45) for λ_1 , μ , and ${}_1b^1(1v)$. We will only look at the range of values for the Pareto weights that do not imply randomization in the government policy.

7.1 Time-Consistency in Examples 1 and 2

In these examples, government spending is constant and there is no difference between time 0 and time 1, provided the private agents continue to trade only in annuities. It is trivial to check that the time-consistent structure of public debt is to keep the same debt structure as at time 0, i.e., to set ${}_1b_t^i = {}_0b_t^i = 0$: the government should continue not to issue any debt claims (nor purchase any). This is consistent with the intuition: the structure of claims that makes the policy time consistent is the one that avoids any temptation to distort prices in the future; if private agents do not offer any temptation by trading only in annuities, so should the government. Since the government has 0 net wealth, its position in the annuity market should accordingly be 0. Furthermore, the government will not have any incentives to randomize its policy whenever it did not have any at time 0.

Suppose now instead that the private agents switch from the annuity market that was open at the beginning of time 0 to short-term debt market. At the beginning of period 0 the agents

²⁹I do not have a general proof that such a system always has a solution, unlike L-S. This is why I stated my proposition as a “result” rather than a “theorem.”

start with a net financial position given by ${}_0\eta_t^1 = -{}_0\eta_t^2 = \frac{1}{3} \quad \forall t \geq 0$. In Example 1 we have $g_t = 0 \quad \forall t \geq 0$, whereas for Example 2 we set $g_t = 0.2 \quad \forall t \geq 0$.

We now require ${}_1\eta_t^i = 0 \quad \forall t \geq 2$, whereas ${}_1\eta_1^i$ is set so that the budget constraint of type-2 agents is met at the Ramsey allocation: at time 0 type-1 agents sell their annuities that give them the right to $1/3$ unit of consumption forever from time 1 on and purchase an equivalent amount of short-term assets against type-2 agents. We wish to study how the government should react to this change in order to preserve time consistency.

Figures 6 and 7 show λ_1 and μ . Here μ measures the value of redistributing one unit of wealth (measured in time-1 units) from the taxpayers to the rentiers, whereas the latter is a measure of the value for the government of switching to lump-sum taxation, i.e., of redistributing resources from the taxpayers to its own budget. In the first example $\lambda_1 = 0$ independently of the Pareto weights of the 2 agents: since the tax rate is 0, at the margin taxes are not distortionary, and hence (at the margin) the government can costlessly transfer consumption units from the taxpayers to itself. The government would like to transfer funds to type-1 agents when ω^1 is large, if it had an instrument to do so, whereas the reverse happens when ω^1 is small. In the second example λ_1 is positive, reflecting the necessity of distortionary taxation.

Figures 6 and 7 show also the structure of debt at time 1 that makes the Ramsey allocation time consistent. In the first example, when $\lambda^1 = 0$, the correct choice of the government is to simply undo the bad incentives generated by the maturity imbalance in the private debt by offsetting with matching issues of public debt. Accordingly, no matter what the Pareto weights of the planner are, the correct choice is to lend short-term to the rentiers and finance these loans by issuing consoles that pay from period 2 on. In the private markets the rentiers purchase the right to 6.6667 units of consumption payable at time 1; the government lends to them by buying claims to 6.3333 units of consumption payable at time 1, and this is financed by issuing consoles (purchased by the rentiers) that pay 0.3333 units of consumption in each period from time 2 on. This arrangement leaves the rentiers with net claims to 0.3333 units of consumption in each future period. The maturity structure of the net position of the government and of the taxpayers is unbalanced, but this is irrelevant, because at the margin the government is not willing to distort prices to transfer funds from its budget to the taxpayers' budgets (or vice versa).

When public spending is positive, the optimal structure of debt is quite different. Now the government is always willing to distort prices to reduce the tax burden, so the strategy of simply offsetting the maturity imbalance in the private market would not work. Looking at Figure 7, we can divide the optimal debt structure into 3 ranges.

In the first range, when the Pareto weight of the taxpayers is high, μ is negative and the government would like to redistribute wealth from the rentiers to the taxpayers. In this range the government issues annuities and lends short-term to the rentiers. In this way the government leaves two opposite and offsetting incentives for period 1. By reducing taxes in period 1 and raising them from then on, the government at time 1 would favor the taxpayers, since the price of the consumption good in time 1 would decrease in terms of future consumption, so the maturing short-term claims of type-1 agents would be worth less. However, this temptation will be offset by the fact that this change in price increases the value of government liabilities, thereby requiring more taxation.

At the boundary between the first and the second range, when $\mu = 0$, the government has no incentive to redistribute wealth. In this case the optimal debt structure entails abstaining from trading any claims: since there is no need of offsetting perverse redistribution temptations, the

government avoids creating any other source of temptation by setting ${}_1b_t^1 \equiv 0$.

In the second range $0 < \mu < \lambda_1$. Here the government would like to redistribute wealth from the taxpayers to the rentiers, but because of distortionary taxation, it still values more the resources in its own budget than the ones of the rentiers, so it is still willing to distort prices in its favor to extract resources from type-1 agents. In this range the government would have an incentive to raise taxes in period 1 and lower them in future periods to increase the value of short-term claims of the rentiers. To offset this, the time-consistent debt structure requires the government to issue short-term debt and purchase annuities.

At the end of the second range, when $\mu = \lambda_1$, the system becomes singular. Since the government has no incentive to distort prices to transfer funds to or from its budget to the rentiers' budget, the debt structure becomes irrelevant, and it is not possible to offset the incentive to redistribute wealth in favor of the rentiers. In this case the only solution for the government is to target type-2 agents and issue debt to them instead. Note that, near this point, the time-consistent maturity structure of debt becomes extremely unbalanced: since the incentives provided by public debt are small, it takes huge imbalances to offset a given incentive to redistribute wealth among private agents.

In the final range, when the Pareto weight of the rentiers is sufficiently high, the government is willing to increase its liabilities to transfer resources to the rentiers. In this case the appropriate choice for the government is again to issue annuities and lend short-term: the imbalance in the private sector would provide the government with an incentive to raise taxes in period 1 to raise the value of claims maturing in that period, but this is counterbalanced by the transfer of resources from type-1 agents to the government that would arise from the imbalance in the structure of public debt.

8 Conclusions and Directions for Future Research

I have used my model to explore some of the ways in which distributional motives affect the choice of an intertemporal tax plan, interacting with efficiency considerations.

In presence of real shocks, the possibility of distorting intertemporal prices gives the government an important redistribution tool. This seems especially relevant for large shocks, such as wars. We have shown that the size of the deficit a government chooses to run during a war will be heavily influenced by its constituency. A government that draws main support from the people that pay taxes should optimally run much larger deficits and wait for the end of the war to levy the taxes necessary to repay for the defense expenses. On the other hand, a king supported by a privileged class of rentiers, largely exempt from taxes, should run a much smaller deficit and force the taxpayers to borrow at bad terms from the rentiers.

Future research is needed to establish how the incentives to distort interest rates for distributional purposes are affected in the presence of other taxes, especially when capital is introduced. It would also be interesting to consider different sources of heterogeneity among consumers, in particular the importance of age and retirement in overlapping-generations models. To this end, it would be interesting to study how endogenous transfers and government spending contribute to achieving redistribution. This issue is addressed in Bassetto [5].

A Proofs

A.1 Proof of Theorem 1

From the first-order conditions of the consumers we have

$$\begin{cases} u_c^i(c_t^i, x_t^i) = \nu^i p & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \\ u_x^i(c_t^i, x_t^i) \geq \nu^i p(1 - \tau_t)w^i, & = \quad \text{if } x_t^i < 1 \end{cases} \quad (51)$$

where ν^i are the Lagrange multipliers associated with the budget constraints of the agents. Because of the strict concavity of u , (51) can be inverted to get

$$\begin{cases} c_t^i = \hat{C}^i(\nu^i, p_t, (1 - \tau_t)w^i) & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \\ x_t^i = \hat{X}^i(\nu^i, p_t, (1 - \tau_t)w^i) & \forall t \geq 0 \text{ a.s.} \quad \forall i = 1, \dots, N \end{cases} \quad (52)$$

where both \hat{C}^i and \hat{X}^i are strictly decreasing in p_t . We now use (52) in the feasibility constraints, which give us

$$\sum_{i=1}^N \hat{C}^i(\nu^i, p_t, (1 - \tau_t)w^i) + g_t = \sum_{i=1}^N w^i \left[1 - \hat{X}^i(\nu^i, p_t, (1 - \tau_t)w^i) \right] \quad \forall t \geq 0 \text{ a.s.} \quad (53)$$

Given the monotonicity properties of \hat{C}^i and \hat{X}^i , (53) is an implicit equation that can at most have one solution for p_t as a function of $(\{\nu^i, w^i\}_{i=1}^N, g_t, \tau_t)$. Since we are considering an allocation and a price system that form a competitive equilibrium, the asset pricing kernel must be a solution of (53) given the Lagrange multipliers.³⁰ We can thus define the function P as the unique solution to (53). By substituting this function into (52) we get consumption and leisure as a function of (g_t, τ_t) given the value of ν^i : this defines the functions C^i and X^i . QED.

Note that the functions C^i , X^i and P we just derived depend on which competitive equilibrium we are in, since the Lagrange multipliers do.

As a technical remark, the functions C^i , X^i and P are measurable and identified up to sets of measure 0. This is because the equations (53) and (52) that define them involve only measurable functions and are valid almost surely.

A.2 Proof of Theorem 2

We need to prove that the allocation and the price system described by (12), together with the initial conditions, the spending process and the policy $\{\tilde{\tau}_t\}_{t=0}^\infty$, satisfy equations (4), (6), (7) and (8).

Define a measure M on $(G \times \mathbb{R}^{2N+1}, \mathcal{G} \times \mathcal{B}^{2N+1})$ by

$$M(A) \equiv \sum_{t=0}^{\infty} \beta^t \text{Prob}((g_t, \{(0b_{t,0}^i, \eta_t^i)\}_{i=1}^N, \tau_t) \in A) = \sum_{t=0}^{\infty} \beta^t \text{Prob}((g_t, \{(0b_{t,0}^i, \eta_t^i)\}_{i=1}^N, \tilde{\tau}_t) \in A) \quad (54)$$

³⁰If we had to actually compute the competitive equilibrium, we should take into account the fact that the Lagrange multipliers depend on the price system. Our problem here is however different: given that we are in a competitive equilibrium with some multipliers ν^i , we want to show that in this equilibrium there can be only one level of p_t associated with any level of (g_t, τ_t) .

Given any measurable function $f : G \times \mathbb{R}^{2N+1} \rightarrow \mathbb{R}$, (54) implies that

$$E \sum_{t=0}^{\infty} \beta^t f(g_t, \{(ob_{t,0}^i \eta_t^i)\}_{i=1}^N, \tau_t) = \int_{G \times \mathbb{R}^{2N+1}} f dM = E \sum_{t=0}^{\infty} \beta^t f(g_t, \{(ob_{t,0}^i \eta_t^i)\}_{i=1}^N, \tilde{\tau}_t) \quad (55)$$

It then follows immediately that (4) is satisfied for the policy $\{\tilde{\tau}_t\}_{t=0}^{\infty}$ whenever it is satisfied for $\{\tau_t\}_{t=0}^{\infty}$. In the same way we can prove that the expected utility of each agent is the same in both equilibria.

Let us now consider equation (6). Assume by contradiction that it does not hold almost surely for the policy $\{\tilde{\tau}_t\}_{t=0}^{\infty}$. Then there is some time \hat{t} such that

$$\text{Prob}\left(\sum_{i=1}^N C^i(g_{\hat{t}}, \tilde{\tau}_{\hat{t}}) + g_{\hat{t}} \neq \sum_{i=1}^N w^i(1 - X^i(g_{\hat{t}}, \tilde{\tau}_{\hat{t}}))\right) > 0$$

It then follows that

$$\sum_{t=0}^{\infty} \beta^t \text{Prob}\left(\sum_{i=1}^N C^i(g_t, \tilde{\tau}_t) + g_t \neq \sum_{i=1}^N w^i(1 - X^i(g_t, \tilde{\tau}_t))\right) > 0$$

which implies

$$\sum_{t=0}^{\infty} \beta^t \text{Prob}\left(\sum_{i=1}^N C^i(g_t, \tau_t) + g_t \neq \sum_{i=1}^N w^i(1 - X^i(g_t, \tau_t))\right) > 0 \quad (56)$$

Equation (56) contradicts equation (6), which must hold almost surely for all periods t given the policy $\{\tau_t\}_{t=0}^{\infty}$ because of the assumptions of the theorem.

In the same way we can prove that (7) and (8) hold for the policy $\{\tilde{\tau}_t\}_{t=0}^{\infty}$. QED

A.3 Proof of Corollary 1

What we need to prove is that we can find a measurable function $f : G \times \mathbb{R}^{2N+1}$ such that, whenever we choose $\tilde{\tau}_t = f(g_t, \{(ob_{t,0}^i \eta_t^i)\}_{i=1}^N, \tilde{h}_0)$

$$\begin{aligned} M(A) &\equiv \frac{1}{1-\beta} \sum_{t=0}^{\infty} \beta^t \text{Prob}\left((g_t, \{(ob_{t,0}^i \eta_t^i)\}_{i=1}^N, \tau_t) \in A\right) \\ &= \frac{1}{1-\beta} \sum_{t=0}^{\infty} \beta^t \text{Prob}\left((g_t, \{(ob_{t,0}^i \eta_t^i)\}_{i=1}^N, \tilde{\tau}_t) \in A\right) \equiv \tilde{M}(A) \quad \forall A \in \mathcal{G} \times \mathcal{B}^{2N+1} \end{aligned} \quad (57)$$

We scaled the measures in (57) so that they are probability measures; this is just for convenience, as we can now call ‘‘conditional expectation’’ the projection operator.

Notice that $M(A)$ and $\tilde{M}(A)$ coincide by their definition on all events that do not depend on τ and $\tilde{\tau}$, i.e. on all sets of the form $A = A_1 \times \mathbb{R}$, $A_1 \in \mathcal{G} \times \mathcal{B}^{2N}$: this is because we are keeping the same spending process and the same initial conditions under both policies.

It is natural to call $(g, \{(ob_{t,0}^i \eta_t^i)\}_{i=1}^N, \tau)$ the random vector whose probability distribution is $M(A)$.

We can then decompose $\tau = E^M(\tau|(g, \{(ob^i, \eta^i)\}_{i=1}^N) + \tau^\perp$. Let F_{τ^\perp} be the c.d.f. of τ^\perp . Let us then choose the function f as follows:

$$\tilde{\tau}_t = f(g_t, \{(ob_{t,0}^i, \eta_t^i)\}_{i=1}^N, \tilde{h}_0) = E^M(\tau|(g, \{(ob^i, \eta^i)\}_{i=1}^N) + \hat{f}(\tilde{h}_0) \quad (58)$$

with

$$\hat{f}(x) \equiv \min\{y : x \leq F_{\tau^\perp}(y)\} \quad (59)$$

Note that this choice implies

$$\text{Prob}(f(\tilde{h}_0) \leq x) = F_{\tau^\perp}(x) \quad \forall x \in \mathbb{R} \quad (60)$$

To prove that \tilde{M} coincides with M , it is enough to show that they coincide on all sets in the following π -system:

$$\begin{aligned} \mathcal{A} \equiv & \{A : \{z \in G \times \mathbb{R}^{2N+1} : z_i \leq \bar{z}_i \quad i = 1, \dots, 2N+1, z_{N+2} \\ & \leq E^M(\tau|z_1 \leq \bar{z}_1, \dots, z_{2N+1} \leq \bar{z}_{2N+1}) + \bar{z}_{2N+2} \text{ for some } \bar{z} \in \mathbb{R}^{2N+2}\} \end{aligned} \quad (61)$$

By construction, given a set $A \in \mathcal{A}$ characterized by a vector \bar{z} , we have $M(A) = M((-\infty, \bar{z}_1] \times \dots \times (-\infty, \bar{z}_{2N+1}] \times \mathbb{R})F_{\tau^\perp}(\bar{z}_{2N+2})$.

Furthermore, for such sets

$$\begin{aligned} \tilde{M}(A) & \equiv \frac{1}{1-\beta} \sum_{t=0}^{\infty} \beta^t \text{Prob}\left((g_t, \{(ob_{t,0}^i, \eta_t^i)\}_{i=1}^N, \tilde{\tau}_t) \in A\right) = \frac{1}{1-\beta} \cdot \\ & \cdot \sum_{t=0}^{\infty} \beta^t \text{Prob}\left(g_t \leq \bar{z}_{1,0}, b_t^1 \leq \bar{z}_{2,0}, \eta_t^1 \leq \bar{z}_3, \dots, b_t^N \leq \bar{z}_{2N,0}, \eta_t^N \leq \bar{z}_{2N+1}, \hat{f}(\tilde{h}_0) \leq \bar{z}_{2N+2}\right) \\ & = \sum_{t=0}^{\infty} \beta^t \text{Prob}\left(g_t \leq \bar{z}_{1,0}, b_t^1 \leq \bar{z}_{2,0}, \eta_t^1 \leq \bar{z}_3, \dots, b_t^N \leq \bar{z}_{2N,0}, \eta_t^N \leq \bar{z}_{2N+1}\right) F_{\tau^\perp}(\bar{z}_{2N+2}) \cdot \\ & \cdot \frac{1}{1-\beta} = M((-\infty, \bar{z}_1] \times \dots \times (-\infty, \bar{z}_{2N+1}] \times \mathbb{R})F_{\tau^\perp}(\bar{z}_{2N+2}) = M(A) \end{aligned} \quad (62)$$

QED.

A.4 Proof of Theorem 3

Since the proof of theorem 3 is rather cumbersome, it is useful to break it into several lemmas.

We first rewrite the objective function in terms of the integrals I_1, I_2, I_3, I_4 , which we already defined in equations (27), (28), (29) and (30), and

$$\hat{I}_2 \equiv \int (ob^1(v) + \eta^1(v))e(v)^{-\gamma} dm(v) \quad (63)$$

In terms of these equations, our problem becomes

$$\max_{e(v)} \left[\omega^1 N^1 I_1^{\gamma-1} \hat{I}_2^{1-\gamma} + \omega^2 (1 - N^1 (1-\gamma)^{-1} I_1^{-1} \hat{I}_2)^{1-\gamma} \right] \frac{I_1}{1-\gamma} + \omega^2 \frac{I_3}{1-\sigma} \quad (64)$$

subject to

$$\log \xi + \gamma \log(1 - N^1 I_1^{-1} \hat{I}_2) + \log(I_4 - I_3) - \log(I_1 - I_2) = 0 \quad (65)$$

Note that I_3 has a monotone effect on the objective function, and that I_4 enters in the constraint but not in the objective function.

If we find a perturbation to a given trajectory \hat{e} that changes I_3 , but not I_1 , I_2 , \hat{I}_2 or $I_3 - I_4$, we can improve on \hat{e} while keeping the constraint holding. Therefore, at an optimum this cannot happen.

We will proceed as follows:

- (i) Lemma 1 first restricts the cases in which the optimal solution \hat{e} may lie on a boundary, i.e. at 0 or at $1 - g$.
- (ii) Under the special case of constant and deterministic government spending and coupon payments, lemma 2 shows that \hat{e} can take at most 3 values, except on a set of m -measure 0.
- (iii) Under the special case of lemma 2, lemma 3 strengthens the result and proves that \hat{e} may take at most 2 values only, except on a set of m -measure 0.
- (iv) We generalize the proof by removing the assumption that led us to study the special case. As we will see, the proof of the general case will work by “conditioning” on the values of $(g_{t,0} b_{t,0}^1 b_{t,0}^2 \eta_t^1)$ and reducing the problem to our special case. This step will not be trivial but will not contain any further insights.

A.4.1 Statement and Proof of Lemma 1

Lemma 1 *Assume that condition 1 holds. Then the optimal choice \hat{e} for maximizing (64) s.t. (65) may not be equal to 0, except on sets of m -measure 0, and it can only be equal to $1 - g$ if it is equal to that value almost everywhere with respect to the measure m .*

Proof. We first consider the boundary $\hat{e}(v) = 0$. If $\gamma \geq 1$, choosing $\hat{e}(v) = 0$ with positive measure leads both consumers to infinite negative utility; the government will never pick such a policy whenever an alternative policy is available.

Consider now the case $\gamma < 1$, $(g_t, N^1 {}_0b_t^1 + {}_0b_{t,0}^2 \eta_t^1 + {}_0b_t^1) \neq \mathbf{0}$. Since $g < 1$ by our assumptions on the public spending process, equation (20) implies $x^2(v) = 1 - g(v) > 0$. The leisure of type 2 agents can be positive while their consumption is 0 only if the tax rate is 100% in the state we are considering, i.e. $\tau = 1$. In this case, these agents will not work at all, which implies $x^2(v) = 1$: this can be consistent with market clearing only if $g(v) = 0$. Since the price of the goods in the states with no consumption is infinite, the budget constraints of the agents can hold only if $N^1 {}_0b^1(v) + {}_0b^2(v) = 0$ and ${}_0\eta^1(v) + {}_0b^1(v) = 0$. These requirements violate condition 1. We thus proved that $e > 0$ except at most on sets of m -measure 0.

We now look at the consequences of $e(v) = 1 - g(v)$. From (21), this can happen in two cases: either $k^2 = 0$ or $\tau(v) = -\infty$. The latter case is easily shown to be incompatible with the government budget constraint. If $k^2 = 0$, it follows that $c^2(v) = 0$ a.s. and hence, from (7), $x^2(v) = 0$ a.s. as well. We thus obtain $e(v) = 1 - g(v)$ a.s. from equation (20). QED.

A.4.2 Statement and Proof of Lemma 2

Lemma 2 *Assume that $(g_{t,0} b_{t,0}^1 b_{t,0}^2 \eta_t^1)$ take a single value almost everywhere with respect to the measure m . Assume that condition 1 holds. Then the optimal choice \hat{e} for maximizing (64) s.t. (65) may take at most 3 values, except on a set of m -measure 0.*

Lemma 2 corresponds to the case of no uncertainty, constant government spending and constant coupon payments among all the agents in the economy.

Proof. We reason by contradiction. Let \hat{e} be the optimal choice by the government. We ruled out that $\hat{e}(v) = 0$ or $\hat{e}(v) = 1 - g$ with positive probability, unless $\hat{e}(v) = 1 - g$ a.s., in which case our statement holds.³¹ Therefore, if \hat{e} takes more than 3 values, we can find an open set $S \subset (0, 1 - g)$ such that \hat{e} takes more than 3 values in S .³² Let $V \equiv \hat{e}^{-1}(S)$ be the set of realizations of v such that $e(v)$ falls into S . We wish to prove that there exist a function e that satisfies the constraint (65) and leads to a higher value for the objective. We restrict our search to the following space:

$$\mathcal{S} \equiv \left\{ e : e \text{ is } m\text{-measurable} \wedge e(v) = \hat{e}(v) \forall v \in [0, 1] \setminus V \wedge e(v) \in S \forall v \in V \right\} \quad (66)$$

It is easy to see that $\hat{e} \in \mathcal{S}$.

The space \mathcal{S} allows perturbations of \hat{e} only in the range where the function lies in S . The reason for this is to be sure that a Fréchet differential is properly defined.

Note first that, if $e \in \mathcal{S}$, then its restriction to V ($\hat{e}|_V$) belongs to $L_m^1(V)$, which is a Banach space. Furthermore, the space of all the restrictions to V of functions in \mathcal{S} is an open subset of $L_m^1(V)$. Since all the perturbations we consider coincide outside of V by our construction of \mathcal{S} , we only consider their restriction on V .

We can treat I_1, I_2, \hat{I}_2, I_3 and I_4 as functions of $e|_V$. It is more convenient to replace I_2 and \hat{I}_2 in our analysis with

$$\tilde{I}_2 \equiv \int e(v)^{-\gamma} dm(v) \quad (67)$$

In the case we are considering here, we have $I_2 = (N^1 {}_0b^1 + {}_0b^2)\tilde{I}_2$ and $\hat{I}_2 = ({}_0b^1 + {}_0\eta^1)\tilde{I}_2$: both I_2 and \hat{I}_2 are simply proportional to \tilde{I}_2 , so that a perturbation that does not affect the latter integral will not affect the two former either.

Let thus $I \equiv (I_1, I_2, I_3, I_4) : \mathcal{S} \rightarrow \mathbb{R}^4$. I is Fréchet differentiable, and its Fréchet differential is given by

$$\delta I(e|_V; h) = \begin{bmatrix} (1 - \gamma) \int_V h(v) e(v)^{-\gamma} dm(v) \\ -\gamma \int_V h(v) e(v)^{-\gamma-1} dm(v) \\ (1 - \sigma) \int_V h(v) (1 - e(v) - g)^{-\sigma} dm(v) \\ -\sigma \int_V h(v) (1 - e(v) - g)^{-\sigma-1} dm(v) \end{bmatrix} \quad (68)$$

³¹For simplicity of notation, we can here drop the dependence of $g, {}_0b^1, {}_0b^2$ and ${}_0\eta^1$ on v , since they are constant functions almost everywhere with respect to the measure m .

³²Note that we require S to be a strict subset of $(0, 1 - g)$. This is convenient to ensure that all our integrals will be properly defined.

We know that, if \hat{e} is a regular point for the mapping I , then we can find a perturbation that will leave I_1, I_2 and $I_3 - I_4$ unchanged while increasing or decreasing I_3 .³³ This would imply that it is possible to improve upon the choice of \hat{e} , and therefore \hat{e} would not be optimal.

We therefore need to show that \hat{e} is a regular point for the mapping I whenever it takes more than three values in V with positive measure m . \hat{e} will be a regular point for I whenever its Fréchet differential is onto \mathbb{R}^4 .

Since the function h is an arbitrary function in $L_m^1(V)$, $\delta I(e|_V; h)$ will not be onto \mathbb{R}^4 if and only if there is a non-zero vector $a \equiv (a_1, a_2, -a_3, -a_4)$ such that

$$a \cdot \begin{bmatrix} e(v)^{-\gamma} \\ e(v)^{-\gamma-1} \\ (1 - e(v) - g)^{-\sigma} \\ (1 - e(v) - g)^{-\sigma-1} \end{bmatrix} = 0 \quad (69)$$

for all $v \in V$, except at most a set of m -measure 0.

The remainder of the proof of lemma 2 shows that equation (69) can never hold in more than 3 points. To do this, we define

$$f_1(y) \equiv a_1 y^{-\gamma} + a_2 y^{-\gamma-1} \quad (70)$$

$$f_2(y) \equiv a_3(1 - y - g)^{-\sigma} + a_4(1 - y - g)^{-\sigma-1} \quad (71)$$

and we look for the maximum number of intersections between f_1 and f_2 in $(0, 1-g)$. By enumerating and studying each possible sign that each component of a can take, it is possible to show that in no case there can be more than 3 intersections between f_1 and f_2 . Note that, by linear homogeneity, we can restrict our attention to $a_1 = 1$ or $a_1 = 0, a_2 = 1$. I only present here the analysis of the most complicate case, i.e. $a_1 = 1, a_2 < 0, a_3 > 0, a_4 < 0$. The other cases, as well as the details of the algebra, are available from the author upon request.³⁴

In this case, f_1 is negative³⁵ for $x < -a_2$, strictly increasing for $x < -\frac{a_2(\gamma+1)}{\gamma}$, strictly concave for $x < -\frac{a_2(\gamma+2)}{\gamma}$ and it has a strictly positive third derivative for $x < -\frac{a_2(\gamma+3)}{\gamma}$.

f_2 is strictly positive for $x < 1 - g + \frac{a_4}{a_3}$, strictly increasing for $x < 1 - g + \frac{a_4(\sigma+1)}{a_3\sigma}$, strictly convex for $x < 1 - g + \frac{a_4(\sigma+2)}{a_3\sigma}$ and it has a strictly positive third derivative for $x < 1 - g + \frac{a_4(\sigma+3)}{a_3\sigma}$.

We will prove that $f_1' - f_2'$ has at most two roots over $(0, 1 - g)$. This is enough to establish that $f_1 - f_2$ has at most 3 roots. We distinguish 7 subcases.

1. $-\frac{a_2(\gamma+1)}{\gamma} < 1 - g + \frac{a_4(\sigma+2)}{a_3\sigma}$. In all of these subcases, $f_1' - f_2'$ has exactly 1 root in the interval $(0, -\frac{a_2(\gamma+1)}{\gamma}]$: $f_1' - f_2'$ is strictly decreasing; it converges to $+\infty$ as $y \rightarrow 0$ and it is strictly negative in $-\frac{a_2(\gamma+1)}{\gamma}$. Furthermore, $f_1' - f_2'$ has no roots $(-\frac{a_2(\gamma+1)}{\gamma}, 1 - g + \frac{a_4(\sigma+1)}{a_3\sigma}]$, since it is strictly negative in this interval.

³³This is an application of theorem 1 in section 9.2 of Luenberger [14].

³⁴All the other cases are considerably simpler, and most of them are trivial. In particular, this is the only case for which we need to study derivatives of up to the third order!

³⁵The complete statement would say that f_1 is strictly negative for $x < -a_2$, 0 for $x = -a_2$ and positive for $x > -a_2$. In this statement and all the following ones we will leave the equality and the other side of the inequality implicit. This is just for brevity.

- 1a. $-\frac{a_2(\gamma+1)}{\gamma} < -\frac{a_2(\gamma+2)}{\gamma} \leq 1 - g + \frac{a_4(\sigma+1)}{a_3\sigma}$. In this case $f'_1 - f'_2$ is strictly increasing on $(1 - g + \frac{a_4(\sigma+1)}{a_3\sigma}, 1 - g)$. It is strictly negative at the lower bound, and tends to $+\infty$ at the upper bound; we therefore have exactly one intersection. In subcase (1a), we therefore have exactly 2 intersections between f'_1 and f'_2 in $(0, 1 - g)$.
- 1b. $-\frac{a_2(\gamma+1)}{\gamma} < 1 - g + \frac{a_4(\sigma+1)}{a_3\sigma} < -\frac{a_2(\gamma+2)}{\gamma} < 1 - g$. Let us first consider the interval $(1 - g + \frac{a_4(\sigma+1)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$. In this interval, f'_1 is strictly convex, whereas f'_2 is strictly concave; it follows that $f'_1 - f'_2$ is strictly convex. Since $f'_1 - f'_2$ is strictly negative at the lower bound of the interval, it can have either 0 or 1 roots in the interval, depending on the sign it takes at the upper bound. In the interval $(-\frac{a_2(\gamma+2)}{\gamma}, 1 - g)$ $f'_1 - f'_2$ is strictly increasing, and its limit at $1 - g$ is $+\infty$; if it is nonnegative at the lower bound, this implies that there was exactly one root in $(1 - g + \frac{a_4(\sigma+1)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$ and there is no root in $(-\frac{a_2(\gamma+2)}{\gamma}, 1 - g)$; if it is strictly negative at the lower bound, then there was no root in $(1 - g + \frac{a_4(\sigma+1)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$ and there is exactly one root in $(-\frac{a_2(\gamma+2)}{\gamma}, 1 - g)$. It follows that in subcase (1b) we have exactly 2 intersections between f'_1 and f'_2 on $(0, 1 - g)$.
- 1c. $-\frac{a_2(\gamma+1)}{\gamma} < 1 - g + \frac{a_4(\sigma+1)}{a_3\sigma} < 1 - g \leq -\frac{a_2(\gamma+2)}{\gamma}$. In this case $f'_1 - f'_2$ is convex over the interval $(1 - g + \frac{a_4(\sigma+1)}{a_3\sigma}, 1 - g)$; it is strictly negative at the lower bound, and it converges to $+\infty$ at the upper bound, so that it has exactly 1 intersection in the considered interval. In subcase (1c) we thus have exactly 2 intersections between f'_1 and f'_2 on $(0, 1 - g)$.
2. $1 - g + \frac{a_4(\sigma+2)}{a_3\sigma} < -\frac{a_2(\gamma+1)}{\gamma}$.
- 2a. $0 < 1 - g + \frac{a_4(\sigma+2)}{a_3\sigma} < -\frac{a_2(\gamma+2)}{\gamma} < 1 - g$. On $(0, 1 - g + \frac{a_4(\sigma+2)}{a_3\sigma}]$, $f'_1 - f'_2$ is strictly decreasing; its limit at the lower bound is $+\infty$. There are no roots if $f'_1 - f'_2$ is positive at the upper bound, and exactly one root if $f'_1 - f'_2$ is nonpositive at the upper bound. On $(1 - g + \frac{a_4(\sigma+2)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$, $f'_1 - f'_2$ is strictly convex. If $f'_1 - f'_2$ is positive at the lower bound, it can have 0, 1 or 2 roots in the interval $(1 - g + \frac{a_4(\sigma+2)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$ and thus the same number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$. If $f'_1 - f'_2$ is nonpositive at $1 - g + \frac{a_4(\sigma+2)}{a_3\sigma}$, it can have 0 or 1 roots in $(1 - g + \frac{a_4(\sigma+2)}{a_3\sigma}, -\frac{a_2(\gamma+2)}{\gamma}]$ and thus it will have either 1 or 2 roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$. On $(-\frac{a_2(\gamma+2)}{\gamma}, 1 - g)$, $f'_1 - f'_2$ is strictly increasing, and its limit at the upper bound is $+\infty$. If $f'_1 - f'_2$ has an even number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, then it is nonnegative at $-\frac{a_2(\gamma+2)}{\gamma}$ and thus there are no roots in $(-\frac{a_2(\gamma+2)}{\gamma}, 1 - g)$; if it has an odd number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, then it is negative at $-\frac{a_2(\gamma+2)}{\gamma}$ and there is exactly one root in $(-\frac{a_2(\gamma+2)}{\gamma}, 1 - g)$. It follows that in subcase (2a) $f'_1 - f'_2$ has either 0 or 2 roots over $(0, 1 - g)$.
- 2b. $1 - g + \frac{a_4(\sigma+2)}{a_3\sigma} \leq 0 < -\frac{a_2(\gamma+2)}{\gamma} < 1 - g$. On $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, $f'_1 - f'_2$ is strictly convex, and its limit at 0 is $+\infty$. Therefore it can have 0, 1 or 2 roots in this interval. On $(-\frac{a_2(\gamma+2)}{\gamma}, 1 - g)$, $f'_1 - f'_2$ is strictly increasing, and its limit at the upper bound is $+\infty$. If $f'_1 - f'_2$ has an even number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, then it is nonnegative at

- $-\frac{a_2(\gamma+2)}{\gamma}$ and thus there are no roots in $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g)$; if it has an odd number of roots in $(0, -\frac{a_2(\gamma+2)}{\gamma}]$, then it is negative at $-\frac{a_2(\gamma+2)}{\gamma}$ and there is exactly one root in $(-\frac{a_2(\gamma+2)}{\gamma}, 1-g)$. Therefore in subcase (2b) $f'_1 - f'_2$ has either 0 or 2 roots in $(0, 1-g)$.
- 2c. $0 < 1-g + \frac{a_4(\sigma+2)}{a_3\sigma} < 1-g \leq -\frac{a_2(\gamma+2)}{\gamma}$ On $(0, 1-g + \frac{a_4(\sigma+2)}{a_3\sigma}]$, $f'_1 - f'_2$ is strictly decreasing; its limit at the lower bound is $+\infty$. There are no roots if $f'_1 - f'_2$ is positive at the upper bound, and exactly one root if $f'_1 - f'_2$ is nonpositive at the upper bound. On $(1-g + \frac{a_4(\sigma+2)}{a_3\sigma}, 1-g)$, $f'_1 - f'_2$ is strictly convex, and its limit at $1-g$ is $+\infty$. If $f'_1 - f'_2$ is positive at $1-g + \frac{a_4(\sigma+2)}{a_3\sigma}$, then there can be either 0 or 2 roots in this interval. If $f'_1 - f'_2$ is nonnegative at $1-g + \frac{a_4(\sigma+2)}{a_3\sigma}$, since it is strictly decreasing in that point, it follows that it will have exactly one root in $(1-g + \frac{a_4(\sigma+2)}{a_3\sigma}, 1-g)$. We thus have that in subcase (2c) there can be either 0 or 2 roots for $f'_1 - f'_2$ in $(0, 1-g)$.
- 2d. $1-g + \frac{a_4(\sigma+2)}{a_3\sigma} \leq 0 < 1-g \leq -\frac{a_2(\gamma+2)}{\gamma}$ In this case $f'_1 - f'_2$ is convex over the whole interval $(0, 1-g)$; furthermore, its limits at both bounds are $+\infty$. It follows that it can have either 0 or 2 roots in the interval.

QED.

A.4.3 Statement and Proof of Lemma 3

Lemma 3 *Assume that $(g_{t,0}, b_{t,0}^1, b_{t,0}^2, \eta_t^1)$ take a single value almost everywhere with respect to the measure m . Assume that condition 1 holds. Then the optimal choice \hat{e} for maximizing (64) s.t. (65) may take at most 2 values, except on a set of m -measure 0.*

Lemma 3 still considers the case of no uncertainty, constant government spending and constant coupon payments among all the agents in the economy. It starts from the result of lemma 2 and strengthens it by considering a different class of perturbations.

Proof. From lemma 2, we know that the optimal choice \hat{e} is a step function with at most 3 values, aside from sets of m -measure 0; from lemma 1, we know that each of the three values lies in $(0, 1-g)$, unless \hat{e} is the constant $1-g$.

We reason again by contradiction. Suppose \hat{e} takes three values, each with positive m -measure; let $e_1 < e_2 < e_3$ be the three values. The integrals I can then be rewritten as

$$I_1 = m_1 e_1^{1-\gamma} + m_2 e_2^{1-\gamma} + \left(\frac{1}{1-\beta} - m_1 - m_2\right) e_3^{1-\gamma} \quad (72)$$

$$\tilde{I}_2 = m_1 e_1^{-\gamma} + m_2 e_2^{-\gamma} + \left(\frac{1}{1-\beta} - m_1 - m_2\right) e_3^{-\gamma} \quad (73)$$

$$I_3 = m_1 (1-g-e_1)^{1-\sigma} + m_2 (1-g-e_2)^{1-\sigma} + \left(\frac{1}{1-\beta} - m_1 - m_2\right) (1-g-e_3)^{1-\sigma} \quad (74)$$

$$I_4 = m_1 (1-g-e_1)^{-\sigma} + m_2 (1-g-e_2)^{-\sigma} + \left(\frac{1}{1-\beta} - m_1 - m_2\right) (1-g-e_3)^{-\sigma} \quad (75)$$

where $m_i \equiv m(\{v : e(v) = e_i\})$, $i = 1, 2$.

As we already observed, \hat{e} cannot be optimal if we can perturb in either direction I_3 while holding I_1 , \tilde{I}_2 and $I_3 - I_4$ constant. For this proof, we treat I as a function of $(e_1, e_2, e_3, m_1, m_2)$. If \hat{e} takes all three values with positive measure, we have $m_i > 0$, $i = 1, 2$ and $m_1 + m_2 < \frac{1}{1-\beta}$. In this case, therefore, I is now a mapping from \mathbb{R}^5 to \mathbb{R}^4 ; given our previous observations, the mapping is well defined and differentiable in an open neighborhood of $(e_1, e_2, e_3, m_1, m_2)$. By the same theorem we applied in lemma 2, \hat{e} cannot be optimal if $(e_1, e_2, e_3, m_1, m_2)$ is a regular point of the mapping I , i.e. if the differential of I as a function of $(e_1, e_2, e_3, m_1, m_2)$ is onto \mathbb{R}^4 . We now prove that $(e_1, e_2, e_3, m_1, m_2)$ is indeed a regular point of I when all three points are distinct and all measures strictly positive. To do this, we will just perturb (e_1, e_2, e_3, m_2) while we will hold m_1 fixed: we will show that the differential with respect to just the four elements already spans \mathbb{R}^4 . The Jacobian of the mapping I is given by³⁶

$$J = \begin{bmatrix} e_1^{-\gamma} & e_1^{-\gamma-1} & (1-e_1-g)^{-\sigma} & (1-e_1-g)^{-\sigma-1} \\ e_2^{-\gamma} & e_2^{-\gamma-1} & (1-e_2-g)^{-\sigma} & (1-e_2-g)^{-\sigma-1} \\ e_3^{-\gamma} & e_3^{-\gamma-1} & (1-e_3-g)^{-\sigma} & (1-e_3-g)^{-\sigma-1} \\ \frac{e_2^{1-\gamma}-e_3^{1-\gamma}}{1-\gamma} & -\frac{e_2^{-\gamma}-e_3^{-\gamma}}{\gamma} & \frac{(1-e_2-g)^{1-\sigma}-(1-e_3-g)^{1-\sigma}}{1-\sigma} & -\frac{(1-e_2-g)^{-\sigma}-(1-e_3-g)^{-\sigma}}{\sigma} \end{bmatrix} \quad (76)$$

which can be rewritten as

$$J = \begin{bmatrix} e_1^{-\gamma} & e_1^{-\gamma-1} & (1-e_1-g)^{-\sigma} & (1-e_1-g)^{-\sigma-1} \\ e_2^{-\gamma} & e_2^{-\gamma-1} & (1-e_2-g)^{-\sigma} & (1-e_2-g)^{-\sigma-1} \\ e_3^{-\gamma} & e_3^{-\gamma-1} & (1-e_3-g)^{-\sigma} & (1-e_3-g)^{-\sigma-1} \\ \int_{e_2}^{e_3} y^{-\gamma} dy & \int_{e_2}^{e_3} y^{-\gamma-1} dy & \int_{e_2}^{e_3} (1-y-g)^{-\sigma} dy & \int_{e_2}^{e_3} (1-y-g)^{-\sigma-1} dy \end{bmatrix} \quad (77)$$

The Jacobian J can only be singular if there exists a nonzero vector (a_1, a_2, a_3, a_4) such that

$$a_1 e_i^{-\gamma} + a_2 e_i^{-\gamma-1} + a_3 (1-e_i-g)^{-\sigma} + a_4 (1-e_i-g)^{-\sigma-1} = 0 \quad i = 1, 2, 3 \quad (78)$$

and

$$\int_{e_2}^{e_3} (a_1 y^{-\gamma} + a_2 y^{-\gamma-1} + a_3 (1-y-g)^{-\sigma} + a_4 (1-y-g)^{-\sigma-1}) dy = 0 \quad (79)$$

From lemma 2, we know that the function that we are integrating in (79) can have at most three zeros in $(0, 1-g)$. By (78), the three zeros are e_1, e_2, e_3 , so that the function is never zero in any point of (e_2, e_3) ; since it is a continuous function, it is either always strictly positive, or always strictly negative. It follows that its integral cannot be zero; therefore J is of full rank and \hat{e} cannot be an optimal choice. QED.

We are now ready to prove the main theorem.

A.4.4 Proof of the Main Body of Theorem 3

Note that maximizing (25) subject to (22) can be rewritten as maximizing (64) subject to (65), given the definitions of our integrals. As in lemmas 2 and 3, our proof proceeds by using the fact

³⁶For convenience, we scaled the columns by $\frac{1}{1-\gamma}$, $-\frac{1}{\gamma}$, $\frac{1}{1-\sigma}$ and $-\frac{1}{\sigma}$ respectively. This does not alter the rank of the Jacobian, and it allows us to have a shorter expression.

that we can improve upon \hat{e} if we can find a perturbation that can vary I_3 in either direction while leaving I_1, I_2, \hat{I}_2 and $I_3 - I_4$ unchanged. Recalling the definition of \tilde{v} in (38), we can use Fubini's theorem and rewrite the integrals as follows:

$$I_1 \equiv \int \int_0^1 e(\tilde{v}, h)^{1-\gamma} dh dm(\tilde{v}, [0, 1]) \quad (80)$$

$$I_2 \equiv \int \int_0^1 (N^1 o b^1(\tilde{v}, h) +_o b^2(\tilde{v}, h)) e(\tilde{v}, h)^{-\gamma} dh dm(\tilde{v}, [0, 1]) \quad (81)$$

$$\hat{I}_2 \equiv \int \int_0^1 ({}_o b^1(\tilde{v}, h) +_o \eta^1(\tilde{v}, h)) e(\tilde{v}, h)^{-\gamma} dh dm(\tilde{v}, [0, 1]) \quad (82)$$

$$I_3 \equiv \int \int_0^1 (1 - e(\tilde{v}, h) - g(\tilde{v}))^{1-\sigma} dh dm(\tilde{v}, [0, 1]) \quad (83)$$

$$I_4 \equiv \int \int_0^1 (1 - e(\tilde{v}, h) - g(\tilde{v}))^{-\sigma} dh dm(\tilde{v}, [0, 1]) \quad (84)$$

Consider now the inner integrals. In these integrals we are conditioning on \tilde{v} and integrating with respect to h alone. By the same proof as lemmas 2 and 3, $\hat{e}(\tilde{v}, h)$ must take at most 2 values as a function of h for each \tilde{v} , except at most in sets of Lebesgue measure 0,³⁷ for otherwise we can vary the inner integral in I_3 while holding the inner integrals in I_1, I_2, \hat{I}_2 and $I_3 - I_4$ fixed. Of course, changes in the inner integrals will be reflected in changes in the whole integrals only if they take place on sets that have positive m -measure in the outside integration: therefore $\hat{e}(\tilde{v}, h)$ can take more than two values as a function of h for any given \tilde{v} on sets of m -measure 0, but this cannot happen on sets of positive m -measure. QED.

B Uniform Commodity Taxation: a Formal Analysis

This appendix contains a formal treatment of the claims contained in section 6 of the paper. In this appendix, we adopt a notation that allows us to easily compare the results in the papers with what has already been established in a static framework, in particular by Atkinson and Stiglitz [1, 2].

We indicate by c^i , $i = 1, 2$ the vector of consumption goods consumed by type- i agents. x is the vector of leisure consumed by the taxpayers (agents of type 2). The preferences of the rentiers are described by

$$V^1(\Gamma^1(c^1)) \quad (85)$$

and those of the taxpayers by

$$V^2(\Gamma^2(c^2), \Theta(x)), \quad (86)$$

³⁷Note that h is distributed uniformly, so its measure is the Lebesgue measure.

where Γ^i , $i = 1, 2$ and Θ are linearly homogeneous functions, and all the functions are assumed to be twice continuously differentiable. Equations (85) and (86) capture two of the features that are relevant for our purposes: that preferences are separable between leisure and the consumption goods, and that the subutilities are homothetic. These assumptions are satisfied by the preferences (14) and (15) that we assumed in section 4.

To keep notation simple, we will let c^i and x be finite-dimensional vectors; c_j^i will denote the j -th component of c^i , and x_j will denote the j -th component of x . All the results continue to hold if we switch to the appropriate notation in an infinite-dimensional space.

The technology of the economy is characterized by

$$F(N^1 c^1 + c^2 + g, x) \leq 0. \quad (87)$$

We assume the technology exhibits constant returns to scale. To stay closer to A-S, we first assume F to be twice continuously differentiable, with a strictly positive gradient. This assumption implies that any good (or leisure) can be transformed into another good (leisure), which is violated by our problem; we therefore will later amend this hypothesis and look at the implications of doing so.

Atkinson and Stiglitz work mostly with a small open economy (or a linear technology), in which producer prices are given, although their results are more general; in their case, the function F could be written as

$$F(N^1 c^1 + c^2 + g, x) = \sum_j q_j^* (N^1 c_j^1 + c_j^2 + g_j) + \sum_j w_j^* (x_j - 1) \quad (88)$$

where q_j^* are the international prices of the different consumption goods and w_j^* are the international wages for the various types of leisure.

Let w be the vector of wages corresponding to the different types of leisure and q the vector of producer prices of the consumption goods. If we normalize to 1 the wage rate of time of the first type, profit maximization on the firms' part requires:

$$q_j = -\frac{F_{c_j}}{F_{x_1}} \quad (89)$$

and

$$w_j = \frac{F_{x_j}}{F_{x_1}} \quad (90)$$

The budget constraints of the rentiers and the taxpayers can be written as follows:

$$\sum_j p_j (c_j^1 - \bar{c}_j^1) - T \leq 0 \quad (91)$$

and

$$\sum_j p_j (c_j^2 - \bar{c}_j^2) + \sum_j w_j^a (x_j - 1) - T \leq 0, \quad (92)$$

where \bar{c}^i is the vector of the initial endowment of each type, p is the vector of consumer prices, w^a is the vector of after-tax wages and T is a lump-sum transfer from the government. We already

imposed that the taxpayers start with 1 unit of time of each type; since we are free to adjust the function F , this can be viewed simply as a normalization.

In line with A-S, we assume the government can tax the consumption goods (net of the initial endowment) and the labor supply, but it cannot tax any type of leisure. For a general production function F , this is a richer set of instruments than the one we introduced in the paper, where only the labor supply can be taxed. However, we will show later that taxing consumption in addition to the labor supply is redundant for the particular production function that we use in the paper. One tax rate is redundant, so we can set $q_1 = p_1$.

As in the main text, we will work with the primal problem: we will use the first-order conditions of the consumers and the producers to substitute out the prices and we will look at the Ramsey problem as one of solving for quantities.³⁸ From the budget constraints and the first-order conditions of the consumers we obtain the following implementability constraints:

$$\sum_j \Gamma_j^1 (c_j^1 - \bar{c}_j^1) - \Gamma_1^1 T \leq 0 \quad (93)$$

and

$$V_1^2 \sum_j \Gamma_j^2 (c_j^2 - \bar{c}_j^2) + V_2^2 \sum_j \Theta_j (x_j - 1) - V_1^2 \Gamma_1^2 T \leq 0. \quad (94)$$

In equations (93) and (94) and in what follows a subscript j to a function refers to the partial derivative with respect to the j -th component. We normalized the price of the first consumption good to 1, we multiplied the first equation by Γ_1^1 and the second equation by $V_1^2 \Gamma_1^2$.

In addition to the implementability constraints, the government faces the following further constraints:

- (i) the feasibility constraint, given by equation (87);
- (ii) in a competitive equilibrium, the marginal rates of substitution must be the same for all consumers, i.e.

$$\Gamma_j^1 \Gamma_k^2 = \Gamma_k^1 \Gamma_j^2 \quad \forall j, k \quad (95)$$

Because of (95), the implementability constraint of the rentiers can also be written in the following form, which will be more convenient later:

$$V_1^2 \sum_j \Gamma_j^2 (c_j^1 - \bar{c}_j^1) - V_1^2 \Gamma_1^2 T \leq 0. \quad (96)$$

The first-order conditions for the government are

$$\omega^1 V_1^1 \Gamma_i^1 + \lambda^1 V_1^2 \Gamma_i^2 + \sum_{j>1} \nu_j \left[\Gamma_{1i}^1 \Gamma_j^2 - \Gamma_{ij}^1 \Gamma_1^2 \right] = \mu F_{c_i} \quad \forall i, \quad (97)$$

³⁸ Atkinson and Stiglitz follow the dual approach: they substitute out quantities and solve the problem in terms of prices. The primal approach is easier in our case in which we have an initial endowment of more than one good.

$$\begin{aligned}
& \omega^2 V_1^2 \Gamma_i^2 + \lambda^1 \left\{ V_{11}^2 \Gamma_i^2 \left[\sum_j \Gamma_j^2 (c_j^1 - \bar{c}_j^1) - \Gamma_1^2 T \right] + V_1^2 \left[\sum_j \Gamma_{ij}^2 (c_j^1 - \bar{c}_j^1) - \Gamma_{1i}^2 T \right] \right\} \\
& + \lambda^2 \left\{ V_{11}^2 \Gamma_i^2 \left[\sum_j \Gamma_j^2 (c_j^2 - \bar{c}_j^2) - \Gamma_1^2 T \right] + V_1^2 \left[\sum_j \Gamma_{ij}^2 (c_j^2 - \bar{c}_j^2) - \Gamma_{1i}^2 T \right] + V_1^2 \Gamma_i^2 \right. \\
& \left. + V_{12}^2 \Gamma_i^2 \sum_j \Theta_j (x_j - 1) \right\} + \sum_{j>1} \nu_j \left[\Gamma_1^1 \Gamma_{ij}^2 - \Gamma_j^1 \Gamma_{1i}^2 \right] = \mu F_{c_i} \quad \forall i,
\end{aligned} \tag{98}$$

$$\begin{aligned}
& \omega^2 V_2^2 \Theta_i + \lambda^1 V_{12}^2 \Theta_i \left[\sum_j \Gamma_j^2 (c_j^1 - \bar{c}_j^1) - \Gamma_1^2 T \right] + \lambda^2 \left\{ V_{12}^2 \Theta_i \left[\sum_j \Gamma_j^2 (c_j^2 - \bar{c}_j^2) - \Gamma_1^2 T \right] \right. \\
& \left. + V_{22}^2 \Theta_i \sum_j \Theta_j (x_j - 1) + V_2^2 \sum_j \Theta_{ij} (x_j - 1) + V_2^2 \Theta_i \right\} = \mu F_{x_i} \quad \forall i
\end{aligned} \tag{99}$$

and

$$-\lambda^1 V_1^2 \Gamma_1^2 - \lambda^2 V_1^2 \Gamma_1^2 \geq 0 \implies \lambda^1 \geq -\lambda^2, \quad T \geq 0, \quad (\lambda^1 + \lambda^2)T = 0, \tag{100}$$

where $\lambda^1, \lambda^2, \nu$ and μ are the Lagrange multipliers associated with the constraints (96), (94), (95) and (87) respectively. Given the Ramsey allocation, the conditions for a competitive equilibrium imply the following price system and tax policy:

$$p_i = \frac{\Gamma_i^2}{\Gamma_1^2} \quad \forall i, \tag{101}$$

$$w_i^a = \frac{V_2^2 \Theta_i}{V_1^2 \Gamma_1^2} \quad \forall i, \tag{102}$$

$$q_i = \frac{F_{c_i}}{F_{c_1}} \quad \forall i, \tag{103}$$

$$w_i = \frac{F_{x_i}}{F_{c_1}} \quad \forall i, \tag{104}$$

$$\tau_i^c = \frac{p_i}{q_i} - 1 \quad \forall i \tag{105}$$

and

$$\tau_i^w = 1 - \frac{w_i^a}{w_i} \quad \forall i, \tag{106}$$

where we normalized $p_1 = q_1 = 1$.

We have a uniform commodity tax when $\frac{p_i}{q_i}$ is independent of i , or, equivalently, when $\frac{\Gamma_i^2}{F_{c_i}}$ is independent of i .³⁹ To study what conditions lead to a uniform commodity tax, it is useful to rewrite the first-order conditions as follows:

$$\omega^1 V_1^1 \frac{\Gamma_i^1}{F_{c_i}} + \lambda^1 V_1^2 \frac{\Gamma_i^2}{F_{c_i}} = \mu - \frac{1}{F_{c_i}} \sum_{j>1} \nu_j [\Gamma_{1i}^1 \Gamma_j^2 - \Gamma_{ij}^1 \Gamma_1^2] \quad \forall i, \quad (107)$$

$$\begin{aligned} & \frac{\Gamma_i^2}{F_{c_i}} \left\{ \omega^2 V_1^2 + \lambda^1 \left[V_{11}^2 \sum_j \Gamma_j^2 (c_j^1 - \bar{c}_j^1) \right] + \lambda^2 \left[V_{11}^2 \sum_j \Gamma_j^2 (c_j^2 - \bar{c}_j^2) + V_1^2 + V_{12}^2 \sum_j \Theta_j (x_j - 1) \right] \right\} \\ & = \mu - \frac{1}{F_{c_i}} \left[\lambda^1 V_1^2 \sum_j \Gamma_{ij}^2 (c_j^1 - \bar{c}_j^1) + \lambda^2 V_1^2 \sum_j \Gamma_{ij}^2 (c_j^2 - \bar{c}_j^2) + \sum_{j>1} \nu_j (\Gamma_1^1 \Gamma_{ij}^2 - \Gamma_j^1 \Gamma_{1i}^2) \right] \quad \forall i \end{aligned} \quad (108)$$

and

$$\begin{aligned} & \frac{\Theta_i}{F_{x_i}} \left\{ \omega^2 V_2^2 + \lambda^1 V_{12}^2 \sum_j \Gamma_j^2 (c_j^1 - \bar{c}_j^1) + \lambda^2 \left[V_{12}^2 \sum_j \Gamma_j^2 (c_j^2 - \bar{c}_j^2) + V_{22}^2 \sum_j \Theta_j (x_j - 1) + V_2^2 \right] \right\} \\ & = \mu - \frac{\lambda^2 V_2^2}{F_{x_i}} \sum_j \Theta_{ij} (x_j - 1) \quad \forall i, \end{aligned} \quad (109)$$

together with (100).

Note that all the terms in T dropped out of equations (107), (108) and (109) because of (100). Lump-sum transfers and taxes affect our problem only through the multipliers λ^1 and λ^2 : given such multipliers, the first-order conditions are identical whether T is optimally chosen or is constrained to be 0, as in our main text. While lump-sum taxes and transfers can reduce the incentive to distort prices, they cannot completely offset it unless $\lambda^1 = \lambda^2 = 0$; it is easy to check that in this case the allocation is an unconstrained Pareto optimum, which can only happen if the government does not need to levy distortionary taxes.

We can now focus on the terms that break the optimality of a uniform commodity tax.

- (i) The terms $\sum_j \Gamma_{ij}^2 c_j^1$ and $\sum_j \Gamma_{ij}^2 c_j^2$. If Γ^2 is homogeneous of degree 1, then its derivatives are homogeneous of degree 0 and hence $\sum_j \Gamma_{ij}^2 c_j^2 = 0$: this is the key to the result obtained by Atkinson and Stiglitz [1]. However, it is not enough for Γ^2 (nor Γ^1) to be homogeneous of degree 1 to reduce the first of the two terms to 0. If Γ^1 and Γ^2 are two different homogeneous functions, then the rentiers and the taxpayers will allocate their spending over the consumption goods in different proportions; the government will then be able to favor a group by taxing more lightly the goods it consumes in a larger proportion, which would lead away from optimal taxation. On the other hand, if Γ^1 and Γ^2 are the same function, then equality of the marginal rates of substitution implies that c^1 is proportional to c^2 and both sums are 0: in this case, both groups allocate their spending in equal proportions on the consumption goods and the government would not favor either by deviating from a uniform tax. The environment of sections 4 and 5 satisfies the condition $\Gamma^1 = \Gamma^2$, so the deviations from a uniform commodity tax do not arise from this source in our case.

³⁹Due to (95), this also implies that $\frac{\Gamma_i^1}{F_{c_i}}$ is independent of i .

- (ii) The terms $\sum_j \Gamma_{ij}^2 \bar{c}_j^1$ and $\sum_j \Gamma_{ij}^2 \bar{c}_j^2$. Even if Γ^1 and Γ^2 are homogeneous of degree 1 and are the same function, deviations from a uniform tax follow if \bar{c}^1 and \bar{c}^2 are not proportional to c^2 (and c^1). Since \bar{c}^1 is the only source of resources for the rentiers, if it is proportional to c^1 it must be equal to it as well: we are then in a case in which the rentiers do not trade away from their initial endowment. As we observed, the net trade between the two types is the main determinant of the pattern of taxes we derive in this paper. By distorting prices, taxes change the value of the initial endowment at consumer prices; each group gets a positive [negative] income effect from increases in the prices of goods for which it is a net seller [buyer]. For this reason, a government that wants to favor the rentiers will use some price distortion even if it can raise revenues from the taxpayers and redistribute them lump-sum; the first-order conditions show that it will be optimal to trade off the increased distortions from the necessity of raising additional revenues to rebate lump sum with the price distortions that a nonuniform commodity tax implies.
- (iii) The term $\sum_{j>1} \nu_j (\Gamma_1^1 \Gamma_{ij}^2 - \Gamma_j^1 \Gamma_{1i}^2)$. This term comes from the constraint that marginal rates of substitution should be equal across consumers (equation (95)). We now show that this constraint is not binding if the previous two sources of deviation from uniform commodity taxation are not present. To see this, assume that the functions Γ^1 and Γ^2 are the same and that \bar{c}^1 and \bar{c}^2 are proportional to c^1 and c^2 . If (95) is not binding, then $\nu = \mathbf{0}$. Under these conditions, equation (108) implies that $\frac{\Gamma_i^2}{F_{c_i}}$ is independent of i and hence equation (107) implies the same for $\frac{\Gamma_i^1}{F_{c_i}}$; Γ^1 and Γ^2 are thus proportional to each other, and the constraint (95) is satisfied.

Notice that uniform commodity taxation *does not imply* uniform factor taxation. We could easily repeat the same steps to analyze the taxes on factors; since the initial endowment of each factor is 1, homotheticity will not be enough to establish uniform factor taxation unless the labor supply is constant. This is the reason why the labor tax rate is not constant in the representative-agent economy of Lucas and Stokey [13], even when the separability and homotheticity requirements discussed above hold and when there is no initial government debt.

The previous analysis requires the production function to allow for substitutability of all input and output factors. Our technology is instead described by

$$F_i(N^1 c_i^1 + c_i^2 + g_i, x_i) \leq 0 \quad \forall i. \quad (110)$$

With this production function, the multiplier μ will be a vector rather than a scalar. In this case, the Ramsey allocation identifies uniquely the consumer prices (through the marginal rates of substitution), but not the producer prices: since there is no substitutability among different goods, firms will not be able to change their production in response to changes in relative prices. Analogously, firms cannot substitute different factors of production and hence are not able to react to changes in wages before tax. The firms' profit maximization conditions only link the producer price of a good with the wage before tax in the same period and state of nature. Because of this, the government can then implement the Ramsey allocation by using consumption taxes or labor taxes alone, as we assumed in the text; the only requirement is the following:

$$\frac{1 + \tau_j^c}{1 - \tau_j^w} = \frac{F_{jx} V_1^2 \Gamma_j^2}{F_{jc} V_2^2 \Theta_j} \quad \forall j. \quad (111)$$

Given equation (111), it is always possible to obtain a constant tax rate on consumption goods or a constant tax rate on all factors of production by an appropriate choice. As an example, in the paper we normalized $\tau^c = \mathbf{0}$. However, when the sources of deviations from uniform commodity taxes and/or uniform factor taxes are present, it is not possible in general to have *both* a uniform commodity tax and a uniform factor tax. For this reason, we obtain different tax rates on labor income even under the assumptions of section 4.

References

- [1] A.B. Atkinson and Joseph E. Stiglitz. The Structure of Indirect Taxation and Economic Efficiency. *Journal of Public Economics*, 1:97–119, 1972.
- [2] A.B. Atkinson and Joseph E. Stiglitz. The Design of Tax Structure: Direct versus Indirect Taxation. *Journal of Public Economics*, 6:55–75, 1976.
- [3] Robert J. Barro. Are Government Bonds Net Wealth? *Journal of Political Economy*, 82:1095–1117, 1974.
- [4] Robert J. Barro. On the Determination of the Public Debt. *Journal of Political Economy*, 87:940–971, 1979.
- [5] Marco Bassetto. Political Economy of Taxation in an Overlapping–Generations Economy. Mimeo, Northwestern University, 1998.
- [6] Michael Ben-Gad. *Public Debt with Heterogeneous Agents – A General Equilibrium Approach*. PhD thesis, University of Chicago, 1996.
- [7] Henning Bohn. Optimal State–Contingent Capital Taxation: When is there an Indeterminacy? *Journal of Monetary Economics*, 34:125–137, 1994.
- [8] V.V. Chari, Lawrence J. Christiano, and Patrick J. Kehoe. Optimal Fiscal Policy in a Business Cycle Model. *Journal of Political Economy*, 102:617–652, 1994.
- [9] James Conklin. *Essays on Mechanisms of Sovereign Commitment (Financial Boycott, Supergame Equilibria, Spain)*. PhD thesis, Stanford University, 1994.
- [10] Philip Hoffman, Gilles Postel-Vinay, and Jean-Laurent Rosenthal. Économie et Politique. Les marchés du crédit à Paris, 1750–1840. *Annales HSS*, 49:65–98, 1994.
- [11] Philip Hoffman, Gilles Postel-Vinay, and Jean-Laurent Rosenthal. Redistribution and Long–Term Private Debt in Paris, 1660–1726. *Journal of Economic History*, 55:256–84, 1995.
- [12] Kenneth L. Judd. Redistributive Taxation in a Simple Perfect Foresight Model. *Journal of Public Economics*, 28:59–83, 1985.
- [13] Robert E. Lucas, Jr. and Nancy L. Stokey. Optimal Fiscal and Monetary Policy in an Economy without Capital. *Journal of Monetary Economics*, 12:55–93, 1983.
- [14] David G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, 1969.

- [15] Torsten Persson and Lars E.O. Svensson. International Borrowing and Time-Consistent Fiscal Policy. *Scandinavian Journal of Economics*, 88:273–295, 1986.
- [16] Frank P. Ramsey. A Contribution of the Theory of Taxation. *Economic Journal*, 37:47–61, 1927.
- [17] Jean-Laurent Rosenthal. Rural Credit Markets and Aggregate Shocks: The Experience of Nuits St. Georges, 1756–1776. *Journal of Economic History*, 54:288–306, 1994.
- [18] Thomas J. Sargent and François R. Velde. Macroeconomic Features of the French Revolution. *Journal of Political Economy*, 103:474–518, 1995.
- [19] Thomas J. Sargent and François R. Velde. Optimal Fiscal Policy in a Linear Stochastic Economy. In Ramon Marimon and Andrew Scott, editors, *Computational Methods for the Study of Dynamic Economies*. Oxford University Press, 1999.
- [20] Guido Tabellini. The Politics of Intergenerational Redistribution. *Journal of Political Economy*, 99:335–357, 1991.
- [21] Stephen J. Turnovsky and William A. Brock. Time Consistency and Optimal Government Policies in Perfect Foresight Equilibrium. *Journal of Public Economics*, 13:183–212, 1980.
- [22] François R. Velde and David R. Weir. The Financial Market and Government Debt Policy in France, 1746–1793. *Journal of Economic History*, 52:1–39, 1992.
- [23] Philippe Weil. Love Thy Children – Reflections on the Barro Debt Neutrality Theorem. *Journal of Monetary Economics*, 19:377–391, 1987.

Figure 1

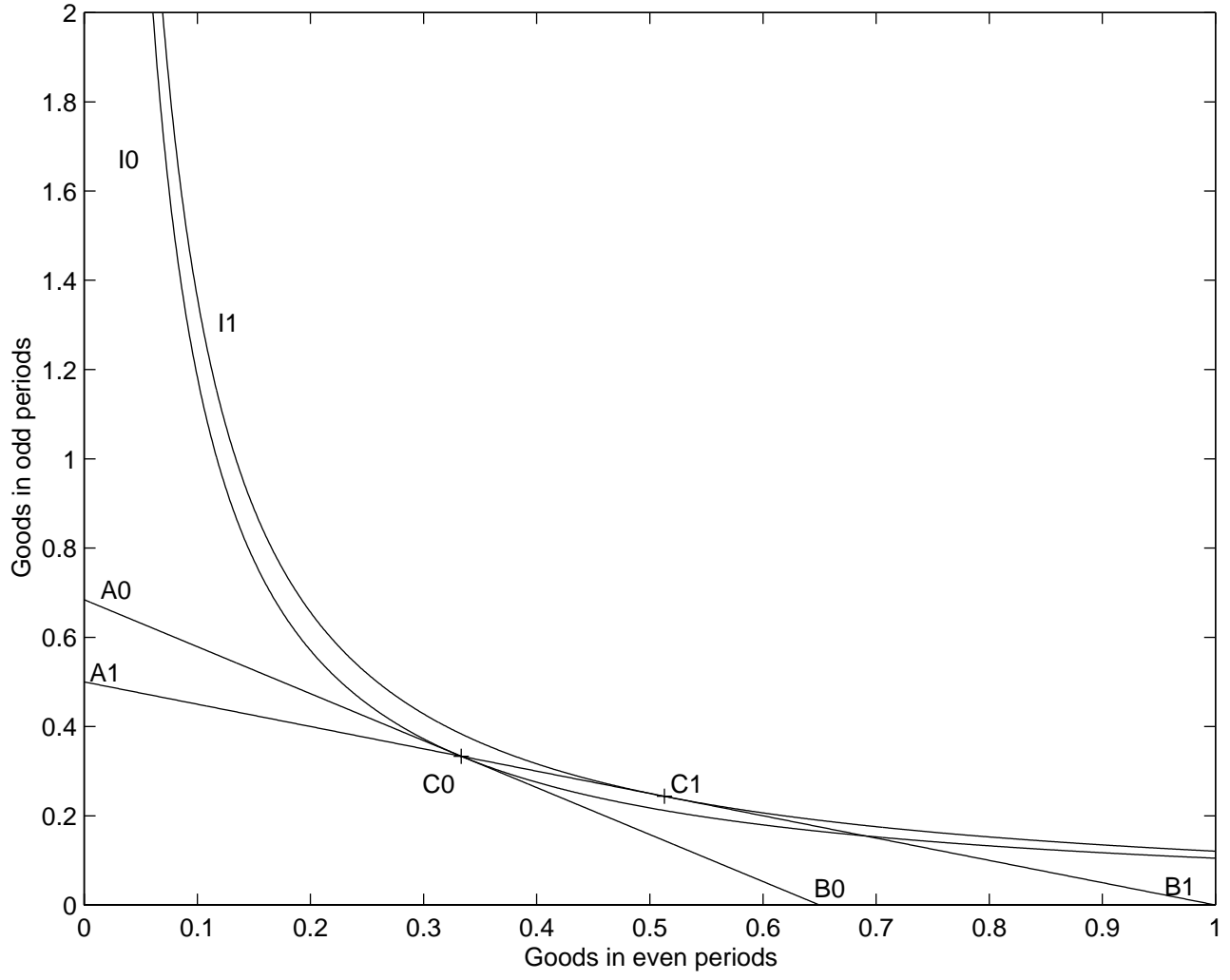


Figure 2 – France vs. Britain example

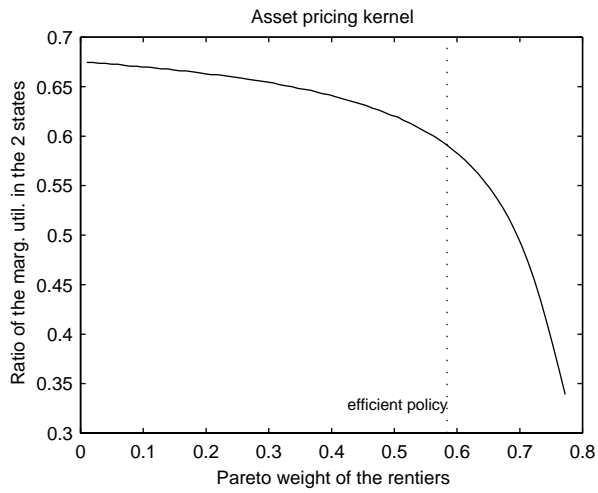
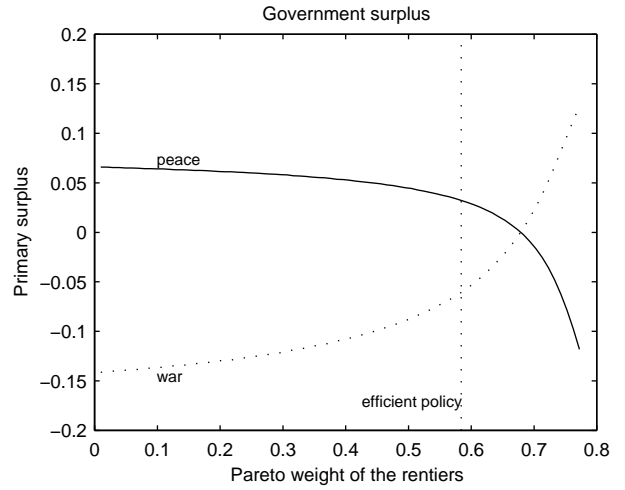
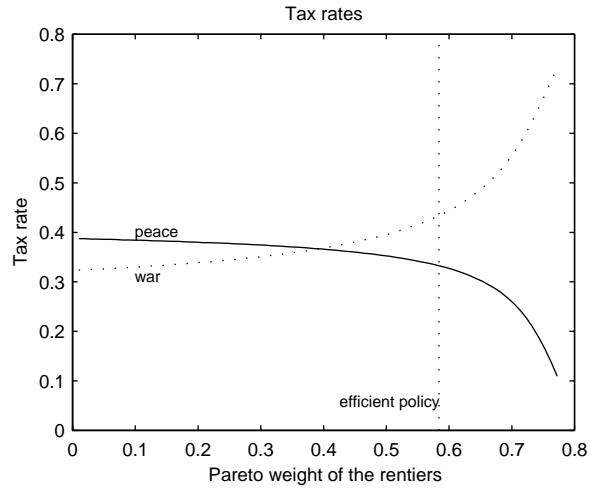


Figure 3 – France vs. Britain example

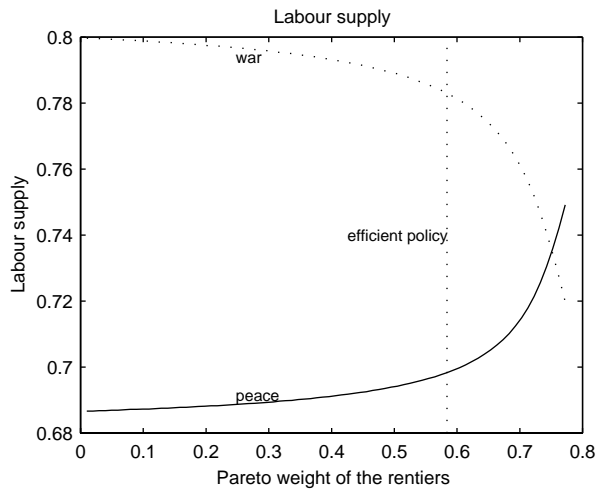
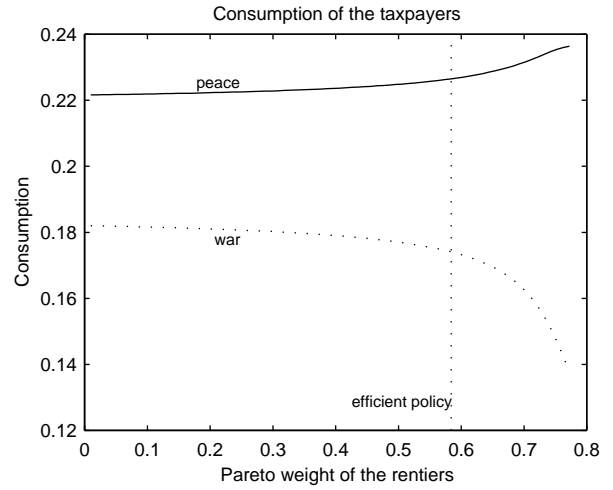
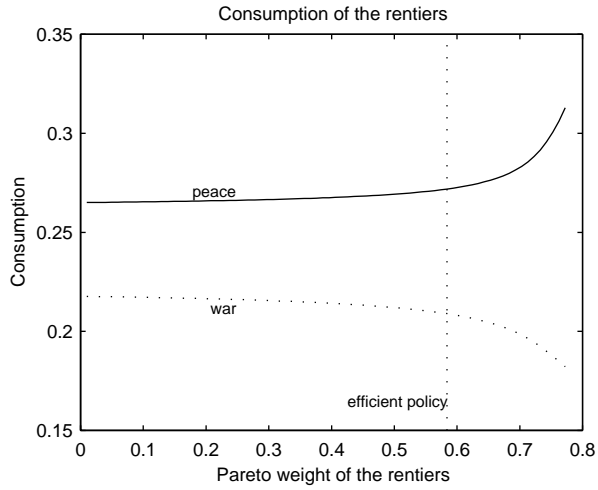


Figure 4 – bus. cycle example

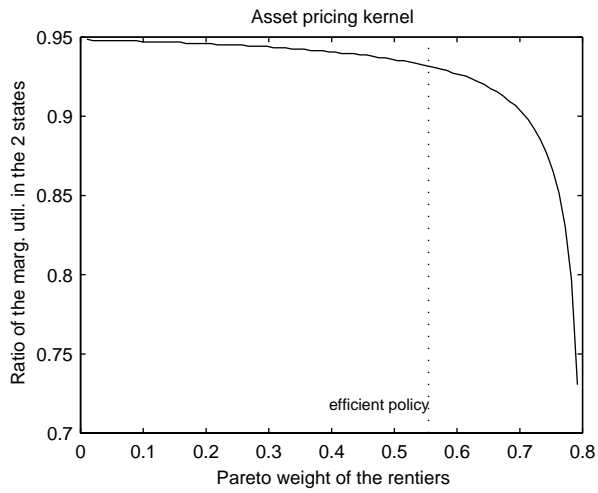
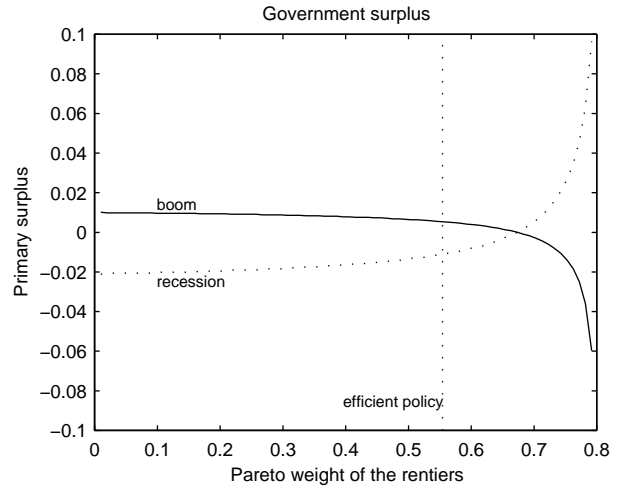
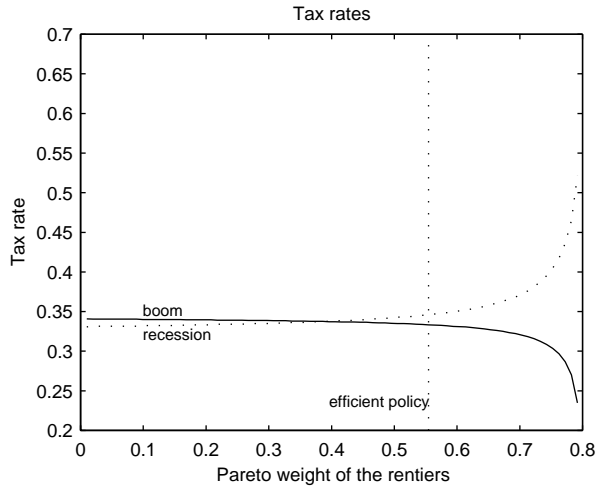


Figure 5 – bus. cycle example

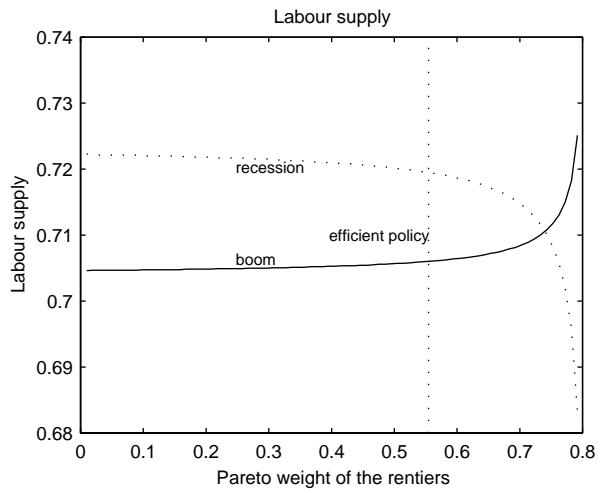
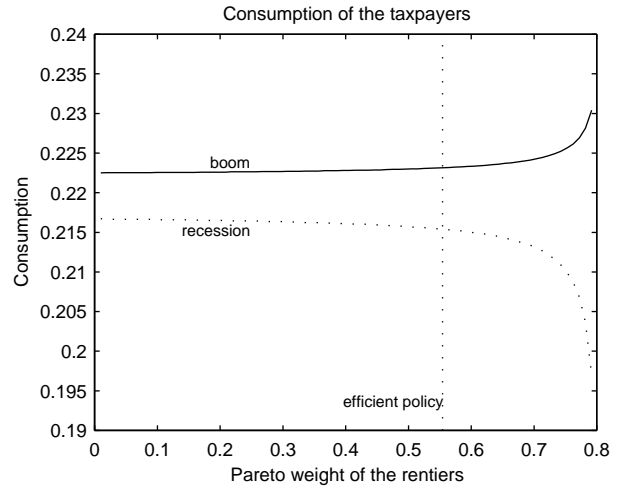
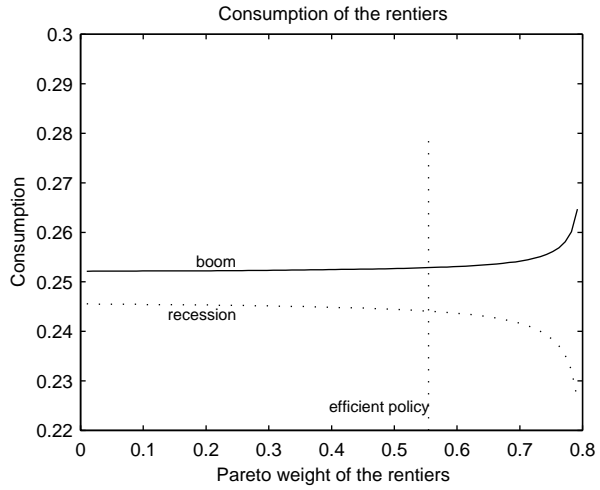


Figure 6: time consistency in example 1

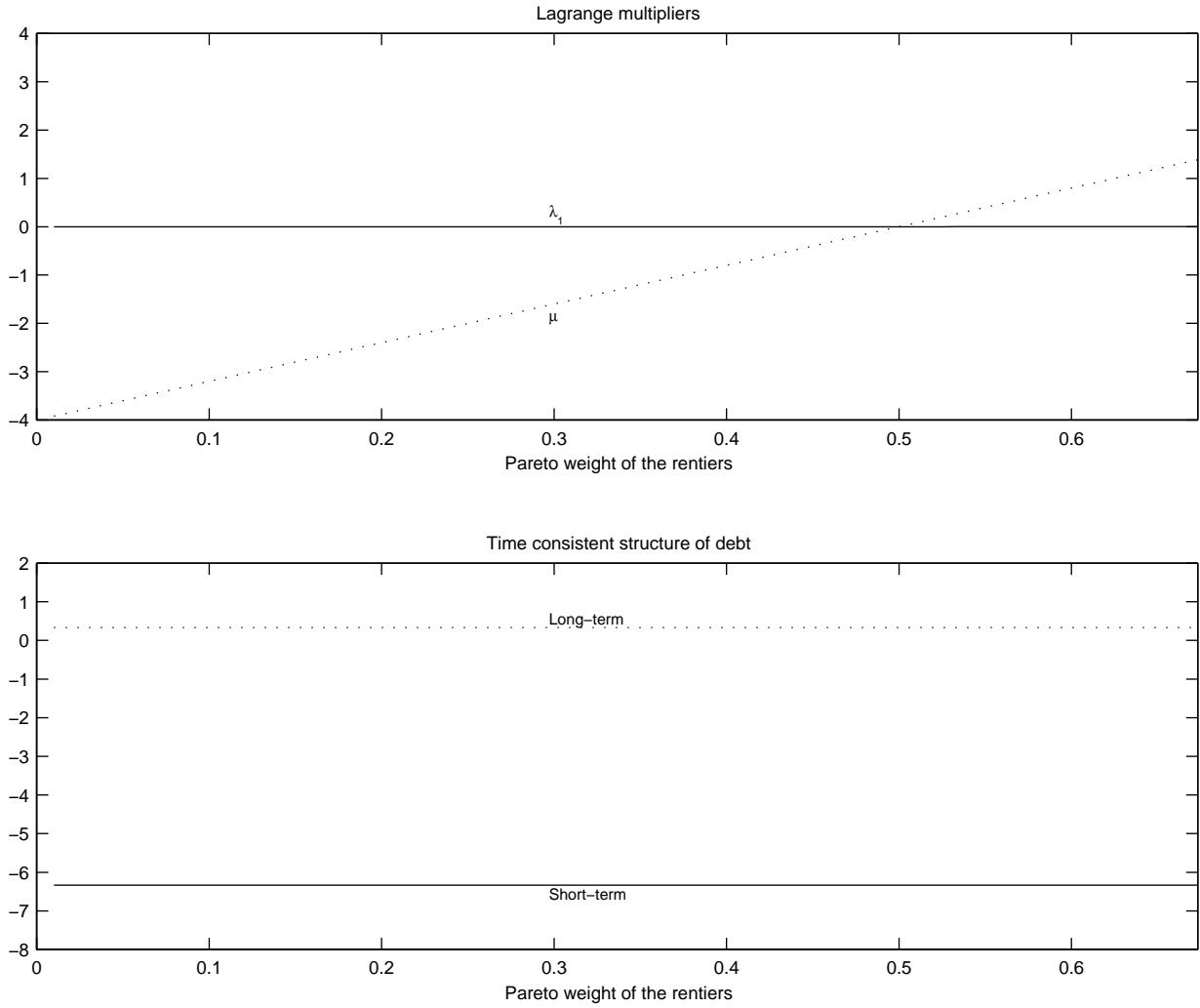


Figure 7: time consistency in example 2

