

Political Economy of Ramsey Taxation*

Daron Acemoglu
MIT

Michael Golosov
Yale and NES

Aleh Tsyvinski
Yale and NES

August 2009.

Abstract

We study the dynamic taxation of capital and labor in the Ramsey model under the assumption that taxes and public good provision are decided by a self-interested politician who cannot commit to policies. We show that, as long as the discount factor of the politician is equal to or greater than that of the citizens, the Chamley-Judd result of zero long-run taxes holds. In contrast, if the politician is less patient than the citizens, the best (subgame perfect) equilibrium from the viewpoint of the citizens involves long-run capital taxation.

JEL Classification: H11, H21, E61, P16.

Keywords: capital taxation, fiscal policy, political economy.

*We are grateful to Andrew Atkeson, Tim Besley, V.V. Chari, Stephen Coate, and Pierre Yared for comments. We thank Georgy Egorov and Oleg Itskhoki for research assistance and the National Science Foundation for financial support.

1 Introduction

Atkeson, Chari and Kehoe (1999) summarize the main result of the Ramsey paradigm of dynamic optimal taxation—taxing capital income is a bad idea. When taxes on labor and capital are restricted to be linear and when the government is benevolent and can commit to a complete sequence of tax policies, Chamley (1986) and Judd (1985) result holds—the optimal dynamic tax sequence involves zero capital taxes in the long run. The result is surprisingly general and robust in a variety of settings, including models with human capital accumulation (Jones, Manuelli, and Rossi, 1997), models where capital-holders are distinct from workers (Judd, 1985), and certain overlapping generations models (Atkeson, Chari and Kehoe, 1999, Garriga, 2001, and Erosa and Gervais, 2002). Similar results hold in stochastic versions of the neoclassical growth model (e.g., Zhu, 1992, Chari, Christiano, and Kehoe, 1994) and most quantitative investigations suggest that capital taxes should be zero or very small even in the short run (e.g., Atkeson, Chari, and Kehoe, 1999).¹ These prescriptions of the Ramsey taxation are used to guide policy not only in developed countries but also around the world.

An obvious shortcoming of this paradigm, and of the results that it implies, is that, in practice, taxes are not set by benevolent governments, but by politicians who have objectives different from citizens. Moreover, these politicians are typically unable to commit to complete sequences of future taxes. These two frictions, self-interest and lack of commitment, are at the center of many political economy models (see, e.g., Persson and Tabellini, 2004, Besley and Coate, 1998) and are also the cornerstone of the public choice theory (see, e.g., Buchanan and Tullock, 1962). From a practical viewpoint, it then seems natural to expect that these frictions should also affect equilibrium taxes and what types of tax structures are feasible. A major question for the analysis of dynamic fiscal policy is whether the key conclusions of the Ramsey paradigm generalize to more realistic environments with self-interested politicians and no commitment. This paper presents a simple answer to this question.

The answer has two parts. First, our analysis reveals a simple but intuitive economic mechanism that makes positive capital taxes optimal from the viewpoint of the citizens; positive capital taxes reduce capital accumulation and thus the incentives of politicians

¹A notable exception is the New Dynamic Public Finance literature, which studies dynamic nonlinear taxes and characterizes conditions under which capital taxes need to be positive to provide intertemporal incentives to individuals with private information (see, e.g., Golosov, Kocherlakota and Tsyvinski 2003, Kocherlakota, 2005, Golosov, Tsyvinski, and Werning, 2006).

to deviate from the policies favored by the citizens. Thus, starting from an undistorted allocation a small increase in capital taxes is typically beneficial because it relaxes the political economy constraints. Second, despite this first-order effect, we show that the result that capital taxes should be equal to zero in the long run generalizes to some political economy environments. That is, even when taxes are set by self-interested politicians with no commitment power to future tax sequences, the best sustainable equilibrium may involve zero taxes. In particular, we delineate precise conditions under which capital taxes are positive in the (best) subgame perfect equilibrium of the political economy environment we specify, but then limit to zero in the long run. Conversely, when these conditions are not satisfied, capital taxes are positive and in the long run, thus presenting a possible explanation for the ubiquity of capital taxes in practice.

More specifically, we model the political economy of taxation using a version of the political agency models by Barro (1973) and Ferejohn (1986). In this model, taxes are the outcome of a dynamic game between politicians and citizens. While politicians have the power to set taxes, they are potentially controlled by the citizens, who can remove them from power using elections or other means. We analyze a neoclassical growth model, where self-interested politicians decide on linear taxes on labor and capital income and manage government debt. The amount that is left after servicing debt and financing public goods constitutes the rents for the politician in power. The interactions between citizens and politicians define a dynamic game. We characterize the best subgame perfect equilibrium (SPE) of this game from the viewpoint of the citizens.² We show that this problem is similar to the dynamic taxation problems in the literature except for the addition of a sequence of *sustainability constraints* for politicians, which ensure that politicians are willing to choose a particular sequence of capital and labor income taxes.

Our first result is that despite the self-interested objectives (rent-seeking behavior) of politicians and the lack of commitment to future policies, the best equilibrium will involve zero capital taxes as in the celebrated Chamley-Judd result, provided that politicians have a discount factor equal to or greater than that of the citizens. The intuition for this result is that the society can structure dynamic incentives to politicians in such a way that, in the long-run, rents to the politicians can be provided in a non-distortionary way. This

²Our focus on the best SPE is motivated by our attempt understand what the best feasible tax structures will be in the presence of political economy and no commitment constraints. Naturally, the dynamic game we specify has other equilibria, and many of these exhibit greater inefficiencies than the best SPE characterized here. We believe that focusing on the best SPE highlights the dynamic economic forces affecting capital taxes in the clearest possible way.

result shows that the Chamley-Judd conclusion concerning the desirability of zero capital taxes in the long run has wider applicability than previously considered.

Our second result, however, delineates a specific reason for why positive capital taxes might be desirable. If politicians are more impatient than the citizens (which may be a better approximation to reality than the politicians having the same patience as the citizens, for example, because of exogenous turnover), the best equilibrium involves long-run capital taxes as well as additional distortions on labor supply. The reason for the presence of positive long-run capital taxation in this case is that, when politicians have a lower discount factor than the citizens, the political sustainability constraint remains binding even asymptotically. This increases the marginal cost of saving (and also of supplying labor for the citizens) because any increase in output must now also be accompanied with greater payments to politicians to provide them with the appropriate incentives. Intuitively, starting from a situation with no distortions (and zero capital taxes), an increase in capital taxation has a second-order effect on the welfare of the citizens holding politician rents constant, but reduces the capital stock of the economy and thus the rents that should be provided to politicians by a first-order amount. Consequently, positive capital taxes will be beneficial to citizens when political sustainability constraints are binding. It is also important to emphasize that such an allocation indeed requires distortionary taxes. If capital taxes were equal to zero, each individual would have an incentive to save more and the capital stock would be too high relative to the one that maximizes the utility of the citizens. Therefore, the “second-best allocation” can be decentralized only by using distortionary (linear) taxes.

Overall, our results suggest that the conclusions of the existing literature may have wider applicability than the framework with benevolent government typically considered in the literature. But, they also highlight a new reason for why positive capital taxes might be useful, and thus suggest caution in applying these results in practice, especially when politicians are short-sighted either because electoral controls are imperfect or because of exogenous turnover or other reasons.

Important precursors to our paper include Brennan and Buchanan (1980) and Wilson (1989), who argue for distortionary taxes to be used to curb the negative political economy effects. In a more recent contribution, Becker and Mulligan (2003) argue that inefficient taxes may be beneficial as a way of reducing excessive spending by politicians and provide empirical evidence consistent with this view. Besley and Smart (2007) emphasize the importance of fiscal restraints in political agency models where politicians are controlled

by elections. None of these papers consider the implications of political economy concerns for long-run capital taxation.³

Our analysis builds on earlier work by Chari and Kehoe (1990, 1993), who study dynamic fiscal policy as a game between a benevolent (potentially time-inconsistent) government and citizens, and on Acemoglu, Golosov and Tsyvinski (2008a,b). Acemoglu, Golosov and Tsyvinski (2008a) develop a benchmark framework for the analysis of government policy in the context of a dynamic game between a self-interested government and citizens, but focus on situations in which there are either no restrictions on tax policies. Acemoglu, Golosov and Tsyvinski (2008b) use this framework for the analysis of the political economy of taxation and dynamic Mirrlees economies. Thus the restrictions on taxes in that paper are endogenous and result from incentive compatibility constraints due to incomplete information. Consequently, these papers do not directly make contact with the large body of work on dynamic fiscal policy, which focuses on the canonical Ramsey setup, where government is limited to linear (distortionary) taxes. The setup is the basis of the celebrated Chamley-Judd zero long-run capital tax result. The current paper extends this framework and provides a systematic analysis of how political economy constraints affect the optimality of long-run capital taxes in the canonical Ramsey setup. It thus clarifies the conditions under which the Chamley-Judd result extends in the presence of political economy constraints, and also highlights why this result may not hold because of clinical economy.

Most closely related to our paper is the recent work by Yared (2008), who studies dynamic fiscal policy in a stochastic general equilibrium framework with linear taxes under political economy constraints similar to ours. The main difference is that Yared's analysis does not incorporate capital, which is the focus of the present paper.

Our paper is also related to Benhabib and Rustichini (1997) and to recent work by Reis (2007) on optimal policy with benevolent government without commitment.⁴ Albanesi and Armenter (2007a,b) provides a unified framework for the study of intertemporal distortions, though they do not incorporate explicit political economy considerations.

³We should also note that the optimality of positive capital taxes even in the long run is not an artifact of our model and reflects concerns faced by real world economic policy. In a political economy setup similar to ours, Caballero and Yared (2008) provide a model and evidence on how rent-seeking politicians can affect the composition of debt over the cycle and suggest that distortionary taxation may be useful as a corrective device in such situations.

⁴There is also a large quantitative literature on time-inconsistent tax policies with benevolent politicians (social planners). For example, Klein, Krusell, and Rios-Rull (2007) focus on time consistent Markovian equilibria, while Phelan and Stacchetti (2001) study more general sustainable equilibria in such environments.

Conesa, et. al. (2008) is a numerical study of nonlinear taxation in which the capital tax is positive. Hassler, Krusell, Storesletten and Zilibotti (2008) features both positive long-run taxation, and possibly non-converging cycles in an environment where depreciation rates change systematically with the age of the capital. Other recent work by Aguiar, Amador, and Gopinath (2007a,b) studies the optimal taxation of capital and optimal debt policy in a small open economy without commitment to future policies, but once again without political economy considerations. Finally, Hassler, Krusell, Storesletten and Zilibotti (2005), Song, Storesletten and Zilibotti (2009) and Battaglini and Coate (2008) also study the political economy of dynamic taxation, but focus on Markov Perfect Equilibria.

The rest of the paper is organized as follows. The next section presents our model and the characterization of equilibrium. It presents all of our main theoretical results. Section 3 illustrates these theoretical results using a simple quantitative exercise. Section 4 concludes.

2 Model and Main Result

We start by setting up a neoclassical economy with Ramsey taxation closely following the standard treatment in Chari and Kehoe (1998). We then augment it with the political economy setup of electoral accountability models in which the politician cannot commit and is self interested.

Consider an infinite-horizon discrete-time economy populated by a continuum of measure 1 of identical consumers with preferences

$$\sum_{t=0}^{\infty} \beta^t [u(c_t) - h(l_t)], \quad (1)$$

where $c \geq 0$ denotes consumption, $l \geq 0$ is labor supply, and $\beta \in (0, 1)$ is the discount factor of the citizens. Preferences are assumed to be separable for simplicity. We make the standard assumptions on preferences that $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ are twice continuously differentiable, with derivatives $u'(\cdot)$ and $h'(\cdot)$, are strictly increasing; $u(\cdot)$ is strictly concave and $h(\cdot)$ is strictly convex. In addition, we impose the following standard Inada conditions on preferences:

1. $\lim_{l \rightarrow 0} h'(l) = 0$. Moreover, there exists some $\bar{L} \in (0, \infty)$ such that $\lim_{l \rightarrow \bar{L}} h'(l) = \infty$. This feature implies that the marginal disutility of labor becomes arbitrarily large when individuals supply the maximum amount of labor, \bar{L} .

2. $\lim_{c \rightarrow 0} u'(c) = \infty$ and $\lim_{c \rightarrow \infty} u'(c) = 0$.

These assumptions ensure interior solutions for c and l .

We use subscript i to denote an individual citizen and designate the set of citizens by I . Each citizen starts with an identical initial endowment of capital $k_0 = K_0$ at time $t = 0$. At time t , an amount of public goods g_t needs to be financed, otherwise, production in the economy is equal to zero. For example, one can think of the public goods g_t as expenditure on infrastructure. When the necessary amount of public goods is provided, the unique final good of the economy can be produced via the aggregate production function $F(K, L)$, where $K \geq 0$ denotes the aggregate capital stock, and $L \geq 0$ denotes the aggregate labor provided by all the citizens. We assume that F is strictly increasing and concave in both of its arguments, continuously differentiable (with derivatives denoted by $F_K(\cdot, \cdot)$ and $F_L(\cdot, \cdot)$) and exhibits constant returns to scale. Throughout, to simplify notation, we interpret $F(\cdot, \cdot)$ as the production function inclusive of undepreciated capital. Finally, we also assume that the aggregate production function satisfies the following natural requirements

- a. there exists $\bar{K} < \infty$ such that $F(\bar{K}, \bar{L}) < \bar{K}$. This assumption ensures that the steady-state level of output has to be finite (since by the concavity of F , it also implies that $F(K, \bar{L}) < K$ for all $K \geq \bar{K}$);
- b. $F_K(K, 0) = 0$ for all K . This assumption implies that when there is no employment, the marginal product of capital is equal to 0.

Factor markets are competitive, and thus, as long as the necessary amount of public good is provided, the wage rate and the interest rate (which is also the rental rate of capital) at time t , w_t and r_t , satisfy

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

The only tax instruments available to the government are linear taxes on capital, $\tau_{k,t}$, and labor income, $\tau_{l,t}$. The government can also use one-period non-state contingent bonds for debt management (see below). Taxation and debt management decisions at time t are made by the politician in power. There is a set \mathcal{I} of potential politicians with identical preferences defined on their own consumption, $x_t \geq 0$. In particular, the utility of a typical politician at time $t = 0$ is given by

$$\sum_{t=0}^{\infty} \delta^t v(x_t), \quad (3)$$

where $v(\cdot)$ is strictly increasing, strictly concave, and continuously differentiable, with $v(0) = 0$. Note that the discount factor of politicians, $\delta \in (0, 1)$, is potentially different from that of the citizens, β .

Denote by $\gamma_t \in \{0, 1\}$ whether the government will supply the necessary public goods. Restricting this choice of γ_t to $\{0, 1\}$ is without loss of any generality, since anything less than the full amount of necessary public good provision leads to the same outcome (lack of production). Let $b_t \in \mathbb{R}$ be the debt level of the government at time t (at date t prices), $q_{t+1} \geq 0$ denote the price of date $t + 1$ government bonds at time t , and $\iota_t \in \{0, 1\}$ denote the debt default decision of the government, with $\iota_t = 0$ corresponding to default at time t (which is feasible only when $b_t > 0$, that is, when the government is indebted at time t). Since the population is normalized to 1, all quantities here stand both for aggregates and per capita levels.

The consumption of the politician, x_t , net debt payments, and government expenditures must be financed by taxation and new debt issuance, so the government budget constraint must be satisfied at all t :

$$x_t + \gamma_t g_t + \iota_t b_t \leq \tau_{k,t} r_t K_t + \tau_{l,t} w_t L_t + q_{t+1} b_{t+1} \quad (4)$$

The left-hand side of (4) corresponds to the outlays of the government at time t , while the right-hand side denotes the revenues resulting from taxation of capital and labor income and issuance of new debt.

We introduce the default decision to ensure that (4) does not become infeasible along off equilibrium paths. Notice also that government debt b_t is not specific to a politician. If the politician in power does not default on government debt at time t , but is replaced, the next politician will start period $t + 1$ with debt obligations b_{t+1} . Throughout, we also take the sequence of necessary public good expenditures $\{g_t\}_{t=0}^{\infty}$ as given and assume that this sequence is such that it is feasible to have $\gamma_t = 1$ for all t (this assumption will be stated as a part of the relevant propositions below). Otherwise, the economy would shut down at some point and would produce zero output thereafter.

At any point of time one politician is in power. Citizens decide whether to keep the politician in power or replace him with a new one using elections.⁵ Specifically, the timing of moves in each period is as follows.

⁵Since all citizens have the same preferences regarding politician behavior, we assume that they will all vote unanimously on replacement decisions. See Acemoglu, Golosov and Tsyvinski (2008a) and Persson and Tabellini (2000, Chapter 4) for further discussion of various decision-making processes that citizens can use for replacing politicians.

1. At the beginning of period t , each citizen $i \in I$ chooses labor supply $l_{i,t} \geq 0$ and the output is being produced according to $F(K_t, L_t)$, where $K_t \equiv \int_{i \in I} k_{i,t} di$ and $L_t \equiv \int_{i \in I} l_{i,t} di$, where $k_{i,t} \geq 0$ denotes the capital holding of agent $i \in I$ at time t . Citizen i receives factor payments $w_t l_{i,t}$ and $r_t k_{i,t}$, with w_t and r_t as given in (2).
2. The politician in power chooses linear taxes on capital $\tau_{k,t}$ and labor $\tau_{l,t}$; $0 \leq \tau_{k,t}, \tau_{l,t} \leq 1$, and makes the decisions on public good provision, $\gamma_t \in \{0, 1\}$, and default, $\iota_t \in \{0, 1\}$. In addition, he announces a price $q_{t+1} \geq 0$ for the next period's government bonds at which an unlimited amount of bonds can be purchased or sold by the citizens. Given these choices, the politician's consumption level $x_t \geq 0$ is determined from the government budget constraint (4) (if this constraint has no solution with $x_t \geq 0$ and $\gamma_t = 1$, then necessarily $\gamma_t = 0$).
3. Given the politician's actions $\{\tau_{k,t}, \tau_{l,t}, q_t, x_t, \iota_t, \gamma_t, q_{t+1}\}$,⁶ each citizen $i \in I$ chooses consumption, $c_{i,t} \geq 0$, and capital and government bond holdings for the next period, $k_{i,t+1} \geq 0$ and $b_{i,t+1}$, subject to the individual flow budget constraint

$$c_{i,t} + k_{i,t+1} + q_{t+1} b_{i,t+1} \leq (1 - \tau_{l,t}) w_t l_{i,t} + (1 - \tau_{k,t}) r_t k_{i,t} + \iota_t b_{i,t}. \quad (5)$$

The right-hand side of this equation includes the individual's total income, comprising labor and capital income net of taxes and government bond payments. The left-hand side is the total expenditure of the individual at date t . We also impose the standard no Ponzi condition on individuals—requiring their lifetime budget constraints to be satisfied—for the *equilibrium* sequence of policies. Note, however, that if $b_{i,t} < 0$, the lifetime budget constraint of individuals might be violated for some non-equilibrium future policy sequences (despite the no Ponzi game condition). This can only be an issue when there is a deviation from equilibrium policies, but we still need to specify how the game proceeds if there is such a deviation. We assume that at any date t , each individual must pay the minimum of $b_{i,t}$ or the net present value of his income in the continuation game. This assumption ensures that lifetime budget constraints are never violated.

4. Citizens decide whether to keep the current politician in power or replace him, $\rho_t \in \{0, 1\}$, with $\rho_t = 1$ denoting replacement.

⁶Throughout, we refer to the tuple $\{\tau_{k,t}, \tau_{l,t}, q_t, x_t, \iota_t, \gamma_t, q_{t+1}\}$ as policies or politician's actions. The sequence $\{g_t\}_{t=0}^{\infty}$ is taken as given and we do not explicitly mention it as part of the policies.

The history at every node of the game, h^t , encodes all actions up to that point. Throughout, we look at pure strategy subgame perfect equilibria (SPE). A strategy profile will constitute a SPE if each individual (citizen and politician) plays a best response to all other strategies at each history h^t .⁷ In addition, we will focus on the SPE that maximizes citizens' utility at time $t = 0$ and refer to this as the *best SPE*. The focus on symmetric equilibria is to reduce notation (given the concavity of the utility function in (1), it is clear that the best equilibrium will be symmetric). The focus on the best equilibrium from the viewpoint of the citizens is motivated by our desire to understand the structure of the best sustainable allocations in an environment with self-interested politicians, i.e., to answer the question of what the best allocations are if the political constraints are imposed. The focus on the best SPE also makes our analysis comparable to the traditional models that look for the utility-maximizing allocation from the viewpoint of the citizens. Clearly, other equilibria will feature more inefficiency than the best SPE. From the strict concavity of individuals' problem, it is clear that the best SPE will be symmetric and we use this fact throughout to economize on notation. In particular, we refer to a SPE by the along-the-equilibrium path actions, that is, as $\{\tau_{k,t}, \tau_{l,t}, x_t, \iota_t, \gamma_t, \rho_t, c_t, l_t, b_t, q_{t+1}, k_{t+1}\}_{t=0}^{\infty}$.

The first step in our analysis is to establish a connection between the SPE of the game described here and competitive equilibria (given policies). In particular, recall that even though there is a dynamic political game between the government and the citizens, each individual makes his economic decisions competitively, that is, taking prices as given.

Definition 1 *For a given sequence of policies $\{\tau_{k,t}, \tau_{l,t}, x_t, \iota_t, \gamma_t, q_{t+1}\}_{t=0}^{\infty}$, a **competitive equilibrium** is a sequence of allocations $\{\hat{c}_t, \hat{l}_t, \hat{b}_t, \hat{k}_{t+1}\}_{t=0}^{\infty}$ together with prices $\{\hat{r}_t, \hat{w}_t\}_{t=0}^{\infty}$ that satisfy*

i (utility maximization) $\{\hat{c}_t, \hat{l}_t, \hat{b}_t, \hat{k}_{t+1}\}_{t=0}^{\infty}$ maximizes (1) subject to (5) given $\{\tau_{k,t}, \tau_{l,t}, x_t, \iota_t, \gamma_t, q_{t+1}\}_{t=0}^{\infty}$ and $\{\hat{r}_t, \hat{w}_t\}_{t=0}^{\infty}$.

ii (factor prices) factor prices \hat{w}_t and \hat{r}_t are given by (2) evaluated at $K_t = \hat{k}_t$ and $L_t = \hat{l}_t$ at each t .

iii (government budget constraint) the government budget constraint (4) is satisfied at each t .

⁷For a standard treatment of the SPE in a game between a government and a continuum of citizens, see Chari and Kehoe (1990).

iv (**feasibility**) the feasibility constraint

$$\hat{c}_t + \hat{x}_t + \gamma_t g_t + \hat{k}_{t+1} \leq F(\hat{k}_t, \hat{l}_t) \quad (6)$$

is satisfied at each t .

Given the differentiability and the Inada-type assumptions imposed above, utility maximization requirement of a competitive equilibrium implies that, as long as $\gamma_t = 1$, the following two first-order conditions must hold

$$(1 - \tau_{l,t})\hat{w}_t u'(\hat{c}_t) = h'(\hat{l}_t) \text{ and } (1 - \tau_{k,t})\beta \hat{r}_t u'(\hat{c}_t) = u'(\hat{c}_{t-1}). \quad (7)$$

These are written for aggregates, suppressing the subscript i , for notational convenience. The first condition requires the marginal utility from an additional unit of labor supply to be equal to the marginal disutility of labor, and the second is the standard Euler equation for the marginal utility of consumption between two periods. In addition, no arbitrage implies that whenever there is no default $\iota_t^* = 1$, the value of holding capital and bonds must be the same, thus

$$(1 - \tau_{k,t})\hat{r}_t = q_t^{-1}. \quad (8)$$

If this condition did not hold, individuals would either not invest in physical capital or not hold any government bonds (since one of the two assets would have a higher certain rate of return than the other). Given the concavity of the utility-maximization problem of the citizens, (5), (7) and (8) are not only necessary but also sufficient. In view of this, we can first state the following preliminary result connecting the SPE in which the government does not default and provides the public good to a corresponding competitive equilibrium.

Proposition 1 *Consider any SPE $\{\tau_{k,t}, \tau_{l,t}, x_t, \iota_t, \gamma_t, \rho_t, c_t, l_t, q_{t+1}, k_{t+1}\}_{t=0}^{\infty}$ with $\gamma_t = \iota_t = 1$ for all t . Then there exists a sequence $\{\tau_{k,t}, \tau_{l,t}, x_t\}_{t=0}^{\infty}$ such that $\{c_t, l_t, b_t, k_{t+1}\}_{t=0}^{\infty}$, with associated prices $\{r_t, w_t\}_{t=0}^{\infty}$, is a competitive equilibrium given $\{\tau_{k,t}, \tau_{l,t}, x_t, \iota_t, \gamma_t, q_{t+1}\}_{t=0}^{\infty}$ and $\{g_t\}_{t=0}^{\infty}$.*

Proof. This result follows from the definition of the competitive equilibrium, Definition 1, the conditions on factor prices (2), the first order conditions on capital and labor (7), and the no-arbitrage condition (8). First, the SPE must satisfy the feasibility condition, (6), by construction, thus the feasibility condition (iv) of Definition 1, and it also satisfies the government budget constraint (4) (with or without financing of government expenditures, $\{g_t\}_{t=0}^{\infty}$, since this is already specified by the sequence $\{\tau_{k,t}, \tau_{l,t}, x_t, \iota_t, \gamma_t, \rho_t, c_t, l_t, b_t, q_{t+1}, k_{t+1}\}_{t=0}^{\infty}$), so the government budget constraint

in the competitive equilibrium (iii) is also satisfied. Finally, given $\{c_t, l_t, b_t, k_{t+1}\}_{t=0}^{\infty}$ and $\{r_t, w_t\}_{t=0}^{\infty}$, $\{\tau_{k,t}, \tau_{l,t}\}_{t=0}^{\infty}$ must satisfy the first order conditions on capital and labor (7) and $\{g_{t+1}\}_{t=0}^{\infty}$ must satisfy the no-arbitrage condition (8), since if this were not the case, there would exist some equilibrium-path history h^t , where an individual can deviate and improve his utility. Since (7) and (8) are necessary and sufficient for utility-maximization, the utility maximization condition in the competitive equilibrium (i) of Definition 1 is also satisfied, completing the proof. ■

To make further progress, we use the standard technique in dynamic fiscal policy analysis of representing a competitive equilibrium subject to taxes by introducing an *implementability* constraint (e.g., Atkinson and Stiglitz, 1980, Chari and Kehoe, 1998, or Ljungqvist and Sargent, 2004). This primal approach has the advantage of turning the government (politician) maximization problem into one of choosing allocations rather than taxes.

Proposition 2 *Take the initial capital tax rate $\tau_{k,0} \in [0, 1)$, the initial capital stock $k_0 \geq 0$, and the initial government bond holdings b_0 as given. Suppose that $\gamma_t = \iota_t = 1$ for all t . Then, the sequence $\{\hat{c}_t, \hat{l}_t, \hat{b}_t, \hat{k}_{t+1}\}_{t=0}^{\infty}$ is a competitive equilibrium for some $\{x_t, g_t\}_{t=0}^{\infty}$ if and only if it satisfies (6) and*

$$\sum_{t=0}^{\infty} \beta^t \left[u'(\hat{c}_t) \hat{c}_t - h'(\hat{l}_t) \hat{l}_t \right] = u'(\hat{c}_0) \left[(1 - \tau_{k,0}) F_K \left(0, \hat{l}_0 \right) k_0 + b_0 \right]. \quad (9)$$

Proof. Substitute the necessary and sufficient first-order conditions for utility maximization given in (7) into the individual budget constraint, (5), and rearrange to achieve the required implementability constraint (9). If this condition were not satisfied, it would imply that either at some t , utility-maximization fails or the individual budget constraint is not satisfied. ■

Given Proposition 2, the traditional analysis of optimal fiscal policy proceeds to find a sequence of allocation and the associated taxes that maximize the utility of the citizens while generating sufficient revenue to finance g_t . In our environment with political economy constraints, there are two crucial differences. First, the best SPE must also raise additional resources to finance government (politician) consumption, x_t . In particular, it is straightforward that if we chose $x_t = 0$ for all t , the politician in power would be better off taxing capital and labor at a very high rate and consuming the proceeds today and then being replaced. Second, and related to the previous point, we must make sure that the politician in power never finds it beneficial to deviate from the implicitly-chosen

sequence of allocations. This will be done by introducing another sequence of constraints, the political *sustainability constraints*. The previous argument already suggests what form these sustainability constraints should take. At any point in time, the politician in power can always deviate to $\tau_{l,t} = \tau_{k,t} = 1$, collect all production as tax revenue, and consume all the proceeds. The worst subgame perfect punishment that the citizens can impose is to replace the politician. After replacement, we assume that the politician receives zero consumption and obtains per period utility $v(0) = 0$ in all future dates.⁸ By the standard arguments in dynamic and repeated games (e.g., Abreu, 1988), it is sufficient to look at this worst punishment to characterize the best SPE. This best deviation for the politician combined with the worst punishment on the side of the citizens implies that the sustainability constraint at time t should take the form

$$\sum_{s=0}^{\infty} \delta^s v(x_{t+s}) \geq v(F(k_t, l_t)). \quad (10)$$

We next show that (10) is in fact the relevant sustainability constraint. In particular, the next proposition proves that if the best allocation subject to (10) involves the provision of the public good in all periods, then the best SPE will involve no replacement of the initial politician and no default, and can be characterized as a solution to a simple maximization problem with (10) as the sustainability constraint.

Proposition 3 *Suppose that given the sequence $\{g_t\}_{t=0}^{\infty}$, any solution to the maximization of (1), subject to the feasibility constraint, (6), the implementability constraint (9), and the political sustainability constraint (10) involves provision of the public good, $\gamma_t = 1$. Then, the best SPE $\{\tau_{k,t}^*, \tau_{l,t}^*, b_t^*, l_t^*, \gamma_t^*, \rho_t^*, x_t^*, c_t^*, l_t^*, k_{t+1}^*, q_{t+1}^*\}$ also involves no replacement of the initial politician, public good provision in all periods and no default at all times (that is, $\rho_t^* = 0$ and $\gamma_t^* = l_t^* = 1$ for all t) along the equilibrium path. This best SPE can be characterized as maximizing the utility of the citizens (1), subject to the feasibility constraint, (6), the implementability constraint (9), and the political sustainability constraint (10).*

Proof. First, note that by the argument preceding the sustainability constraint (10), this equation is a necessary condition, since otherwise the politician can improve his utility by deviating. Moreover, the feasibility constraint (6) is necessary by Definition 1 and implementability constraint (9) is necessary by Proposition 2. Therefore, the best

⁸The alternative would be to allow the politician to save and achieve consumption smoothing after the replacement. Whether or not we allow the politician to save after replacement has no effect on our results.

SPE cannot give higher utility to citizens than the maximization of the citizen's utility (1), subject to feasibility (6), implementability (9), and sustainability (10). This can be achieved with no replacement of the politician, with no default and with the required public good provision at all dates.

We next prove that actions $\rho_t^* = 0$, $\gamma_t^* = \iota_t^* = 1$ for all t are necessarily part of the best SPE. To do this, let us first suppose that there exists a best SPE that implements the maximization of (1), subject to (6), (9), and (10). Let this allocation be denoted by $\{\tau_{k,t}^*, \tau_{l,t}^*, b_t^*, \iota_t^*, \gamma_t^*, \rho_t^*, x_t^*, c_t^*, l_t^*, k_{t+1}^*, q_{t+1}^*\}_{t=0}^\infty$. We will then show that $\rho_t^* = 0$, $\gamma_t^* = \iota_t^* = 1$, so the best SPE involves no political replacement, no default, and involves public good provision along the equilibrium path.

Now, to obtain a contradiction, suppose that the best SPE involves politician replacement along the equilibrium path. Then, the initial politician must be replaced after some equilibrium-path history \hat{h}^t (even though he has not deviated). At time t this politician is in power and pursues a policy that maximizes (1), subject to (6), (9), and (10). This implies that at t , $\gamma_t^* = \iota_t^* = 1$ and the politician's sustainability constraint, (10), holds. Hence, the utility of the politician at time t must be at least $v(F(k_t^*, l_t^*))$. In particular, let us write the utility of this politician as

$$V(k_t^*) \equiv v(x_t^*) + \delta V(k_{t+1}^*) \geq v(F(k_t^*, l_t^*)), \quad (11)$$

where the first relation is just a definition, and the inequality is imposed by (10). Here $V(k_{t+1}^*)$ is the continuation utility of this politician, but since there is replacement in equilibrium (by hypothesis), $V(k_{t+1}^*) = 0$. After replacement, the next politician must be given continuation utility

$$V^R(k_{t+1}^*) \equiv \sum_{s=0}^{\infty} \delta^s v(x_{t+s}^*) \geq v(F(k_{t+1}^*, l_{t+1}^*)) > 0$$

so that the sustainability constraint (10) for this new politician is satisfied. Now consider the following variation: do not replace the initial politician at \hat{h}^t and provide him with exactly the same continuation allocation as the new politician. By construction (and by the fact that all politicians are identical), this variation satisfies (10) after \hat{h}^t . Now, the time t utility of the initial politician after this variation is given as

$$V^A(k_t^*) \equiv v(x_t^*) + \delta V^R(k_{t+1}^*) > v(F(k_t^*, l_t^*)),$$

where the strict inequality follows from (11) combined with the fact $V^R(k_{t+1}^*) > V(k_{t+1}^*) = 0$. But this implies that with this variation, the sustainability constraint,

(10), for the initial politician at time t holds as strict inequality, thus x_t^* can be reduced and c_t^* can be increased, implying that $\{\tau_{k,t}^*, \tau_{l,t}^*, b_t^*, \iota_t^*, \gamma_t^*, \rho_t^*, x_t^*, c_t^*, l_t^*, k_{t+1}^*, q_{t+1}^*\}$ could not have been a solution to the problem of maximizing (1), subject to (6), (9), and (10), yielding a contradiction and establishing the claim that the best SPE must involve $\rho_t^* = 0$ for all t .

To see that the best SPE involves no default, suppose that $\iota_t^* = 0$ and $b_t^* > 0$ (if $b_t^* \leq 0$, $\iota_t = 0$ is not allowed). Then, there exists no price q_t at which individuals would buy bonds in the previous period $t - 1$, thus the allocation must have zero bonds, $b_t^* = 0$, which implies that $\iota_t^* = 1$. This contradiction establishes that $\iota_t^* = 1$ for all t . That the best SPE involves public good provision at all dates is also straightforward by the hypothesis of the proposition (that any solution to maximizing (1), subject to (6), (9), and (10) involves $\gamma_t = 1$).

To complete the proof, we only need to show that the maximization of (1), subject to (6), (9), and (10) is a SPE. This follows straightforwardly from Proposition 1 and the fact that replacing a politician that has deviated from the implicitly-agreed tax sequence is a best response for the citizens given the history h^t up to that point. To see this, consider the following strategy profile; after a deviation the politician will always play $\tau'_{k,t} = \tau'_{l,t} = 1$ for all t . This is a best response for the politician anticipating replacement at each date after deviation, and given this strategy by politicians, replacement after deviation is indeed a best response for the citizens. ■

We now can state and prove our main result, which characterizes the time path of taxes corresponding to the best SPE.

Proposition 4 *Suppose that the maximization of (1), subject to the feasibility constraint, (6), the implementability constraint (9), and the political sustainability constraint (10) involves $\gamma_t = 1$ for all t , that $\{g_t\}_{t=0}^\infty$ converges to some $g^S > 0$, and the best SPE equilibrium $\{\tau_{k,t}^*, \tau_{l,t}^*, b_t^*, \iota_t^*, \gamma_t^*, \rho_t^*, x_t^*, c_t^*, l_t^*, k_{t+1}^*, q_{t+1}^*\}$ is such that the equilibrium allocation $\{c_t^*, l_t^*, b_t^*, k_{t+1}^*\}_{t=0}^\infty$ converges to a steady state (c^S, l^S, b^S, k^S) . Then we have that:*

1. *if the politicians are as patient as, or relatively more patient than, the citizens, i.e., if $\delta \geq \beta$, then the sustainability constraint (10) becomes slack as $t \rightarrow \infty$, and we have that $\lim_{t \rightarrow \infty} \tau_{k,t}^* = 0$;*
2. *if the politicians are relatively less patient than the citizens, i.e., if $\delta < \beta$, then the sustainability constraint (10) binds as $t \rightarrow \infty$, and $\lim_{t \rightarrow \infty} \tau_{k,t}^* > 0$.*

Proof. The proposition follows from the fact that the sequence $\{c_t^*, l_t^*, b_t^*, k_{t+1}^*\}_{t=0}^\infty$ is a solution to maximization of (1) subject to (6), (9) and (10). Write the Lagrangian for this problem and let $\beta^t \lambda_t \geq 0$ be the Lagrange multiplier on the feasibility constraint (6), η on the implementability constraint (9) and $\psi_t \geq 0$ on the participation constraint (10).

Differentiating the Lagrangian implies that the first-order necessary conditions with respect to c_t , l_t , k_{t+1} , and x_t , are

$$u'(c_t^*) + \eta (u'(c_t^*) + u''(c_t^*)c_t^*) = \lambda_t, \quad (12)$$

$$h'(l_t^*) + \eta (h'(l_t^*) + h''(l_t^*)l_t^*) + \beta^{-t}\psi_t v'(F(k_{t+1}^*, l_{t+1}^*)) = \lambda_t F_L(k_{t+1}^*, l_{t+1}^*), \quad (13)$$

$$\lambda_t = \lambda_{t+1}\beta F_K(k_{t+1}^*, l_{t+1}^*) - \beta^{-t}\psi_{t+1}v'(F(k_{t+1}^*, l_{t+1}^*)) F_K(k_{t+1}^*, l_{t+1}^*), \quad (14)$$

$$\lambda_t \beta^t = \sum_{s=0}^t \delta^{t-s} \psi_s v'(x_t^*). \quad (15)$$

Note that by definition, the multiplier on the implementability constraint, η , must be finite. From (12) it follows that there exists $\lim_{t \rightarrow \infty} \lambda_t = \lambda^S < \infty$, because $\lim_{t \rightarrow \infty} c_t^*$ is assumed to exist, and Inada conditions ensure that it is finite since the steady-state output is finite, and $u(\cdot)$ is twice continuously differentiable.

(Part 1) First, suppose that the discount factors of the politician and the citizens are equal, $\delta = \beta$. Then, (15) implies

$$\lambda_t = \sum_{s=0}^t \beta^{-s} \psi_s v'(x_t^*).$$

Suppose, to obtain a contradiction, that $\beta^{-t}\psi_t$ does not converge to zero. We know that $x_t^* \rightarrow x^S$ from the feasibility constraint (6), which in a best SPE must be satisfied with equality: indeed, by hypothesis $\{c_t^*, l_t^*, b_t^*, k_{t+1}^*\}_{t=0}^\infty$ converges to some steady state (c^S, l^S, k^S, b^S) and $\{g_t\}_{t=0}^\infty$ converges to some steady state g^S . Moreover, clearly, $\lim_{t \rightarrow \infty} \sum_{s=0}^t \beta^{-s} \rightarrow \infty$. Then it must be the case that $\lambda_t/v'(x^S) \rightarrow \infty$. Since we proved that $\lim_{t \rightarrow \infty} \lambda_t = \lambda^S < \infty$, this is only possible if $x^S \rightarrow 0$. This implies that the sustainability constraint (10) is violated for sufficiently large t , unless $F(k_t^*, l_t^*) \rightarrow 0$ (i.e., $F(k^S, l^S) = 0$). But the latter would imply that γ_t goes to 0 in finite time (since $g^S > 0$). By hypothesis, the maximization of (1) subject to (6), (9) and (10) yields a solution with $\gamma_t > 0$ for all t . Consequently, the above-described allocation cannot be a best SPE, yielding a contradiction. We therefore conclude that $\beta^{-t}\psi_t \rightarrow 0$. Thus, as $t \rightarrow \infty$, (10) becomes asymptotically slack.

Let us next take the limit as $t \rightarrow \infty$ in (12), (13) and (14). Using the fact that $\beta^{-t}\psi_t \rightarrow 0$, these imply

$$u'(c^S) + \eta(u'(c^S) + u''(c^S)c^S) = \lambda^S, \quad (16)$$

$$h'(l^S) + \eta(h'(l^S) + h''(l^S)l^S) = \lambda^S F_L(k^S, l^S), \quad (17)$$

$$\lambda^S = \lambda^S \beta F_K(k^S, l^S). \quad (18)$$

Equations (16) and (17) imply that $\lambda^S > 0$. To see this, recall that $\lambda^S \geq 0$, because it is the multiplier on the resource constraint. To obtain a contradiction to the claim that $\lambda^S > 0$, suppose that $\lambda^S = 0$. Then, since $h' > 0$ and $h'' > 0$, (17) implies that $\eta \in (-1, 0)$. However, since $u' > 0$ and $u'' < 0$, (16) cannot be satisfied with $\eta \in (-1, 0)$ and $\lambda^S = 0$. This yields a contradiction and establishing that $\lambda^S > 0$. In view of this, (18) implies that

$$\beta F_K(k^S, l^S) = \lim_{t \rightarrow \infty} \beta F_K(k_t^*, l_t^*) = 1. \quad (19)$$

Then, (7) combined with (19) implies that $\lim_{t \rightarrow \infty} \tau_{k,t}^* = 0$, completing the proof of Part 1 when $\delta = \beta$.

Next consider the case where $\delta > \beta$. Since $\sum_{s=0}^t \delta^{t-s}/\beta^t \psi_s > \sum_{s=0}^t \beta^{-s} \psi_s$, the same argument as above establishes that $\beta^{-t}\psi_t \rightarrow 0$ and therefore (19) must hold and thus $\lim_{t \rightarrow \infty} \tau_{k,t}^*$ exists and is equal to 0. This completes the proof of Part 1.

(Part 2). Now consider the case where $\delta < \beta$. By the hypothesis that a steady state exists, (12) implies that $\lambda_t \rightarrow \lambda^S$. First, to obtain a contradiction, suppose that $\lambda^S = 0$. From (15), we have

$$\begin{aligned} \lambda^S &= \lim_{t \rightarrow \infty} \frac{1}{\beta^t} \sum_{s=0}^t \delta^{t-s} \psi_s \\ &= \lim_{t \rightarrow \infty} \left\{ \psi_0 \left(\frac{\delta}{\beta} \right)^t + \psi_1 \left(\frac{\delta}{\beta} \right)^{t-1} + \dots + \psi_t \beta^{-t} \right\}. \end{aligned}$$

Since $\psi_s \geq 0$ for all s , $\lambda^S = 0$ implies that each term in the summation in the second line must go to zero as $t \rightarrow \infty$. Therefore, $\beta^{-t}\psi_t \rightarrow 0$. Then, as $t \rightarrow \infty$, (16) and (17) again hold with $\lambda^S = 0$, and the same argument as in Part 1 yield a contradiction and establishes that $\lambda^S > 0$. By the hypothesis that a steady state exists, we also have $v'(x_t) \rightarrow v'(x^S) > 0$ (since $v'(x) > 0$ for all x). Combining these two observations with (15), we conclude that $\sum_{s=0}^t \delta^{t-s} \psi_s / \beta^t$ must converge to a strictly positive constant (that is, $\lim_{t \rightarrow \infty} \sum_{s=0}^t \delta^{t-s} \psi_s / \beta^t = \Psi > 0$).

Next, suppose, to obtain a contradiction, that $\beta^{-t}\psi_t \rightarrow 0$. This means that for any $\varepsilon > 0$ there exists $T < \infty$ such that for all $t \geq T$, we have $\beta^{-t}\psi_t < \varepsilon$. Take $t > T$ and note that

$$\begin{aligned} & \frac{1}{\beta^t} \sum_{s=0}^t \delta^{t-s} \psi_s \\ & < \psi_0 \left(\frac{\delta}{\beta}\right)^t + \dots + \psi_T \beta^{-T} \left(\frac{\delta}{\beta}\right)^{t-T} + \varepsilon \left[\left(\frac{\delta}{\beta}\right)^{t-T-1} + \left(\frac{\delta}{\beta}\right)^{t-T-2} + \dots + 1 \right] \\ & < \left\{ \psi_0 \left(\frac{\delta}{\beta}\right)^t + \dots + \psi_T \beta^{-T} \left(\frac{\delta}{\beta}\right)^{t-T} \right\} + \varepsilon \frac{1}{1 - \delta/\beta}, \end{aligned}$$

where the first inequality exploits the fact that $\beta^{-t}\psi_t < \varepsilon$ for all $t > T$ and the second line uses the fact that the sum in square brackets is less than $1/(1 - \delta/\beta)$. Next, observe that for t sufficiently large, the expression in the curly brackets is arbitrarily small. Therefore, for sufficiently large t , we have $\sum_{s=0}^t \delta^{t-s}\psi_s/\beta^t < 2\varepsilon/(1 - \delta/\beta)$. Since ε is arbitrary, we have $\sum_{s=0}^t \delta^{t-s}\psi_s/\beta^t \rightarrow 0$, which yields a contradiction to the hypothesis that $\lim_{t \rightarrow \infty} \sum_{s=0}^t \delta^{t-s}\psi_s/\beta^t = \Psi > 0$. This establishes that $\beta^{-t}\psi_t$ does not converge to 0. Then, combining (12), (14) and (19) implies that $\lim_{t \rightarrow \infty} \tau_{k,t}^*$ also exists and $\lim_{t \rightarrow \infty} \tau_{k,t}^* > 0$, completing the proof of Part 2. ■

This proposition is the main result of our paper. The intuition for this result is that, when $\delta = \beta$ or when $\delta > \beta$, the political sustainability constraints are present, but the best SPE involves backloading of the payments to politicians.⁹ This backloading (defined in the right sense) ensures that the sustainability constraint of the politician will eventually become slack. As this happens, distortions, and in particular distortions in saving decisions, disappear, and the corresponding competitive equilibrium converges to zero capital taxes. Therefore, the first part of this proposition shows that the Chamley-Judd results on zero capital taxes generalize to political economy environments where politicians are sufficiently patient.

The second part of the proposition, on the other hand, shows how positive capital taxes can arise as part of the best SPE when politicians are more impatient than the citizens, that is, when $\delta < \beta$. As a result, the sustainability constraint, (10), remains binding asymptotically. A binding sustainability constraint implies that higher output must be associated with greater rents to politicians. This raises the opportunity cost of

⁹See Acemoglu, Golosov and Tsyvinski (2008a) for further discussion of backloading in political economy environments and Ray (2002) for a general treatment of backloading results in principal-agent models.

increasing output for the citizens. In particular, reducing the capital stock away from the “first-best” level weakens the deviation temptations of the politician and reduces the rents that needs to be paid in order to ensure sustainability. Consequently, the best SPE involves lower savings than the first-best (the undistorted neoclassical growth model). It is also important to note that these lower saving levels are *decentralized* by positive long-run capital taxes. This follows from (7); if the economy had $\tau_k = 0$, each individual would choose the undistorted level of savings, leading either to the violation of the sustainability constraint or to higher rents for politicians. Thus positive capital taxes are necessary to ensure the appropriate level of capital accumulation and emerge as a tool useful in maximizing the ex ante utility of the citizens in the presence of political economy distortions.

The result in the first part of Proposition 4 is surprising. It suggests that the conclusions of the existing literature that the capital tax is zero may have a wider applicability than the framework with a benevolent government typically considered in the literature and applies, as in our paper, to a class of circumstances in which the government is controlled by self-interested politicians without the ability to commit to future taxes. Nevertheless, the second part of the proposition might ultimately be the more important result, since politicians being more impatient (short-sighted) than the citizens is arguably a better approximation of reality, particularly if there are exogenous reasons for which politicians lose power (even if they do not deviate from the prescribed sequence of actions). In this light, Proposition 4 suggests that considerable caution is necessary in using the normative benchmark of zero capital taxes emerging from models that ignore political economy constraints.

3 Quantitative Investigation

In this section, we provide an illustrative quantitative investigation of the theoretical results presented in the previous sections. Our purpose is not to undertake a quantitatively plausible calibration, but to give further intuition for the theoretical results derived in the previous section, and also provides some simple insights about convergence to the steady state and the structure of taxes before such convergence takes place.

We choose standard functional forms. In particular, the instantaneous utility of consumption for the citizens is assume to take the iso-elastic form

$$u(c) = \frac{1}{1 - \sigma} c^{1 - \sigma},$$

with $\sigma = 2$, while the disutility of labor is given by

$$h(l) = \frac{1}{1 + \varphi} l^{1+\varphi},$$

where $\varphi = 1$. The discount factor of the citizens is taken as $\beta = 0.95$.

The production function takes the standard Cobb-Douglas form (with full depreciation)

$$F(k, l) = Ak^\alpha l^{1-\alpha},$$

where we normalize $A = 1$, and set $\alpha = 1/3$ to be consistent with a capital share of approximately 1/3 in national income. We set the initial amount of capital to $k_0 = 0.1$.

The instantaneous utility function of politicians is given by

$$v(x) = x^{\sigma_g} / \sigma_g,$$

where $\sigma_g = 0.75$. This implies that politicians has a larger intertemporal elasticity of substitution than the citizens. We adopt this specification, since, otherwise, deviations are not sufficiently attractive for politicians (without introducing the ability to save and borrow for the politicians).

We consider two values for the discount factor of the politician $\delta = 0.95$ and $\delta = 0.9$. Government expenditure is set equal to $g = 0.1$ in each period. Figure 3 shows the results of this numerical example. It depicts the path of capital taxes in the best SPEs for the two different values of δ and the path of capital taxes in the corresponding Ramsey economy (without political economy constraints).¹⁰ In the Ramsey economy, the optimal tax is positive in the first period and then is equal to zero.

The two solid lines in Figure 3 depict the best SPE corresponding to $\delta = 0.95$ and to $\delta = 0.9$. In the first case, the tax on capital converges to zero as predicted by Proposition 4. However, the convergence is slower than in the corresponding Ramsey economy, where there is only one period of positive taxation. In fact, in the best SPE, capital taxes are at first as high as 20% compared to taxes less than 10% in the Ramsey economy.

When $\delta = 0.9$, so that the politician is more impatient than the citizens, capital taxes again start relatively high and decline over time, but do not converge to zero. In this case, the limiting value of capital taxes is about 3.5%. This computation therefore shows that

¹⁰To make the Ramsey economy comparable to the setup with political sustainability constraints, we take the amount of government expenditure to be $x_t + g$ at time t , where the sequence $\{x_t\}$ is the one generated by the best SPE for the same parameter values. This is the reason why Ramsey equilibria are different depending on the value of δ .

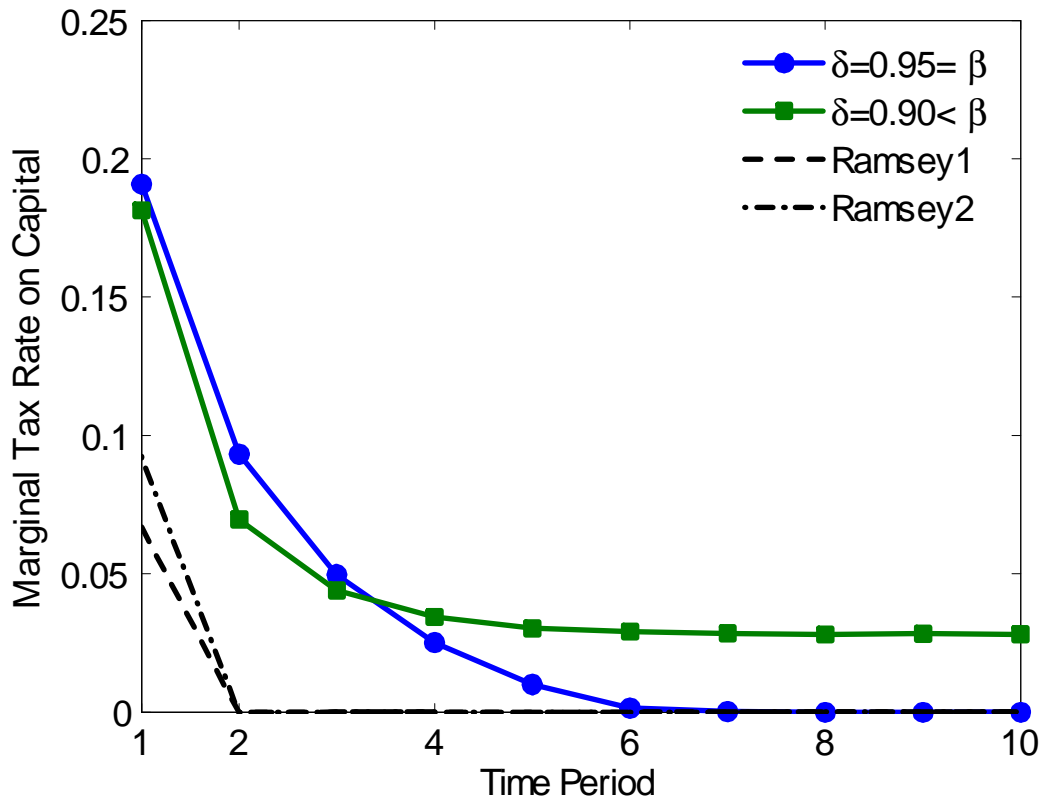


Figure 3: The best SPE and Ramsey equilibria for different values of δ .

a relatively small difference between the discount factors of politicians and citizens leads to positive long-run capital taxes, which is again consistent with the patterns implied by Proposition 4. It is also useful to note that a lower discount factor for the politician does not necessarily imply that capital taxes will be uniformly higher. The figure shows that with $\delta = 0.95$, capital taxes start out higher than in the economy with $\delta = 0.9$, and only fall below those in the $\delta = 0.9$ economy in later periods.

4 Conclusion

The main result of the Ramsey paradigm of dynamic optimal taxation, first arrived by Chamley (1985) and Judd (1985) is that long-run capital taxes should be equal to zero. In practice, most societies have positive taxes on capital income. One perspective, implicitly adopted for example by Atkeson, Chari and Kehoe (1999), is that this is a “bad idea”—a result of bad policy design or incorrect understanding of economic theories.

In this paper, we took an alternative perspective and attempted to understand whether positive taxes on capital income may result from political economy considerations, that is, not as a bad idea, but as a necessary cost to be borne because the government is not perfect agent of the citizens.

Formally, we studied the dynamic taxation of capital and labor in the neoclassical growth model under the assumption that taxes are controlled by self-interested politicians who cannot commit. Politicians, in turn, can be removed from power by citizens via elections. As in the standard (Ramsey) dynamic taxation models, our environment only allows linear taxes on capital and labor income. The celebrated Chamley-Judd result shows that, with benevolent governments with full commitment power, long-run capital taxes should be equal to zero. Since this result relies on the existence of a benevolent government that is able to commit to a complete sequence of (future) tax policies, one may conjecture that the presence of self-interested politicians unable to commit to future taxes will lead to positive long-run capital taxes.

We showed that the long-run capital tax is indeed positive when politicians are more impatient than the citizens. In this case, the marginal cost of additional savings for the citizens is higher in equilibrium than in the undistorted allocation, because a greater level of the capital stock of the economy will increase the politician’s temptation to deviate and thus necessitates greater rents to the politician to satisfy the political sustainability constraint. However, perhaps somewhat surprisingly, when politicians are as patient as, or

more patient than, the citizens, we established that the political sustainability constraint eventually becomes slack and long-run capital taxes converge to zero. Our analysis, therefore, shows that the standard dynamic fiscal policy results may have wider applicability than previously recognized. Perhaps more importantly, they also suggest considerable caution in using these results in more realistic environments without a benevolent, all-powerful social planner. If, as many studies of political economy suggest, politicians are more short-sighted than citizens, the best subgame perfect equilibrium involves positive taxes on capital, even in the long run.

Several research directions for future research are highlighted by our results in the current paper. First, we characterized the structure of “best equilibria”—from the viewpoint of the citizens. An interesting question is whether such equilibria will arise in practice and what types of institutions make their emergence more likely. For example, one may study whether certain specific types of institutions lead to (support) such equilibria, while others make allocations that are within the constrained Pareto frontier more likely. Second, we focused on the specific type of political economy considerations, resulting from self-interested rulers. In practice, in addition to the self-interest of politicians and parties, there are also issues related to conflict between different groups of citizens, and the two sets of issues interact in a rich manner. How these richer political economy interactions affect the structure of optimal dynamic taxation in general, and capital taxation in particular, is another interesting area for future research.

References

Abreu, Dilip. 1988. "On the Theory of Repeated Games with Discounting." *Econometrica*, 56(2):383-396.

Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski. 2008a. "Political Economy of Mechanisms." *Econometrica*, 76(3): 619-642.

Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski. 2008b. "Dynamic Mirrlees Taxation under Political Economy Constraint." MIT Department of economics, working paper.

Aguiar, Mark, Manuel Amador, and Gita Gopinath. 2007a. "Investment Cycles and Sovereign Debt Overhang." Unpublished.

Aguiar, Mark, Manuel Amador, and Gita Gopinath. 2007b. "Efficient Expropriation: Sustainable Fiscal Policy in a Small Open Economy." Unpublished.

Albanesi, Stefania and Roc Armenter 2007a. "Understanding Capital Taxation in Ramsey Models". Unpublished.

Albanesi, Stefania and Roc Armenter. 2007a. "Intertemporal Distortions in the Second Best." NBER Working Paper 13629, CEPR Discussion Paper 6577.

Atkeson, Andrew, V.V. Chari, and Patrick Kehoe. 1999. "Taxing Capital Income: A Bad Idea." *Federal Reserve Bank of Minneapolis Quarterly Review*, 23: 3-17.

Atkinson, Anthony and Joseph Stiglitz. 1980. "Lectures in Public Economics." London: McGraw-Hill.

Barro, Robert. 1973. "The Control of Politicians: An Economic Model." *Public Choice*, 14(1): 19-42.

Battaglini, Marco and Stephen Coate. 2008. "A Dynamic Theory of Spending, Taxation, and Debt." forthcoming *American Economic Review*.

Becker, Gary and Casey B. Mulligan. 2003. "Deadweight Costs and the Size of Government." *Journal of Law and Economics*, 46(2): 293-340.

Benhabib, Jess, and Aldo Rustichini. 1997. "Optimal Taxes without Commitment." *Journal of Economic Theory*, 77(2): 231-259.

Besley, Timothy and Michael Smart. 2007. "Fiscal Restraints and Voter Welfare." *Journal of Public Economics*, 91(3-4): 755-773.

Besley, Timothy and Stephen Coate. 1998. "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis." *American Economic Review*, 88(1): 139-156.

Brennan, Geoffrey, and James Buchanan. 1980. "The Power to Tax." Cambridge

University Press.

Buchanan, James M. and Gordon Tullock. 1962. *The Calculus of Consent*. Michigan: University of Michigan Press.

Caballero, Ricardo and Pierre Yared. 2008. "Inflating the Beast: Political Incentives under Uncertainty". Unpublished.

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium and Infinite Lives." *Econometrica*, 54(3): 607-622.

Chari, V.V. and Patrick Kehoe. 1990. "Sustainable Plans." *Journal of Political Economy*, 98(4): 783-802.

Chari, V.V. and Patrick Kehoe. 1993. "Sustainable Plans and Mutual Default." *Review of Economic Studies*, 60(1): 175-195.

Chari, V.V., Lawrence Christiano, and Patrick Kehoe. 1994. "Optimal Fiscal Policy in a Business Cycle Model." *Journal of Political Economy*, 102(4): 617-652.

Conesa, Juan Carlos, Kitao, Sagiri, and Krueger, Dirk. 2008. "Taxing Capital? Not a Bad Idea After All!", forthcoming in *American Economic Review*.

Erosa, Andres and Martin Gervais. 2002. "Optimal Taxation in Life-Cycle Economies." *Journal of Economic Theory*, 105(2): 338-369.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice*, 50(1-3): 5-25.

Garriga, Carlos. 2001. "Optimal Fiscal Policy in Overlapping Generations Models." Working Papers in Economics 66, Universitat de Barcelona.

Golosov, Michael, Aleh Tsyvinski and Ivan Werning. 2006. "New Dynamic Public Finance: A User's Guide." In *NBER Macroeconomics Annual 2006*. Cambridge, MA: MIT Press.

Golosov, Michael, Narayana Kocherlakota and Aleh Tsyvinski. 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies*, 70(3): 569-587.

Hassler, John, Per Krusell, Kjetil Storesletten and Fabrizio Zilibotti. 2005. "The Dynamics of Government." *Journal of Monetary Economics*, 52:7, pp 1331-1358.

Hassler, John, Per Krusell, Kjetil Storesletten and Fabrizio Zilibotti. 2008. *Journal of Monetary Economics*, 55:4, May, 2008, pp. 692-709.

Jones, Larry E., Rodolfo E. Manuelli and Peter E. Rossi. 1997. "On the Optimal Taxation of Capital Income." *Journal of Economic Theory*, 73(1): 93-117.

Judd, Kenneth. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, 78(1): 59-83.

- Klein, Paul, Per Krusell and José-Víctor Ríos-Rull. 2007. "Time Consistent Public Expenditures," *Review of Economic Studies*, Forthcoming.
- Kocherlakota, Narayana. 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation." *Econometrica*, 73(5): 1587-1621.
- Ljungqvist, Lars and Thomas J. Sargent. 2004. *Recursive Macroeconomic Theory*. Cambridge, MA: MIT Press.
- Marcet, Alberto and Katharina Greulich, 2009. "Pareto-Improving Optimal Capital and Labor Taxes", CSIC Working paper.
- Persson, Torsten and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press.
- Persson, Torsten and Guido Tabellini. 2004. "Constitutional Rules and Fiscal Policy Outcomes." *American Economic Review*, 94(1): 25-45,
- Phelan, Christopher, and Ennio Stacchetti. 2001. "Sequential Equilibria in a Ramsey Tax Model." *Econometrica*, 69(6): 1491–1518.
- Ramsey, Frank. 1927. "A Contribution to the Theory of Taxation." *Economic Journal*, 37(145): 47-61.
- Song, Zheng, Kjetil Storesletten and Fabrizio Zilibotti. 2009. "Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt", Working paper.
- Ray, Debraj. 2002. "Time Structure of Self-Enforcing Agreements." *Econometrica*, 70(2): 547-82.
- Reis, Catarina. 2007. "Taxation without Commitment." Unpublished.
- Wilson, John D. 1989. "On the Optimal Tax Base for Commodity Taxation." *American Economic Review*, 79(5):1196-1206.
- Yared, Pierre. 2008. "Politicians, Taxes, and Debt.", forthcoming in *Review of Economic Studies*.
- Zhu, Xiaodong. 1992. "Optimal Fiscal Policy in a Stochastic Growth Model." *Journal of Economic Theory*, 58(2): 250-289.